Tutorial letter 101/3/2018

Statistical Inference II

STA2602

Semesters 1 & 2

Department of Statistics

IMPORTANT INFORMATION:

This tutorial letter contains important information about your module and includes the assignment questions for both semesters.

CONTENTS

Page

1 INTRODUCTION

Dear Student

We are pleased to welcome you to this module and hope that you will find it both interesting and rewarding. We shall do our best to make your study of this module successful. You will be well on your way to success if you start studying early in the year and resolve to do the assignment(s) properly. All second-level modules are presented in the semester system from 2011 onwards.

You will receive a number of tutorial letters during the year. A tutorial letter is our way of communicating with you about teaching, learning and assessment.

Tutorial Letter 101 contains important information about the scheme of work, resources and assignments for this module. We urge you to read it carefully and to keep it at hand when working through the study material, preparing the assignment(s), preparing for the examination and addressing questions to your lecturers.

In Tutorial Letter 101, you will find the assignments and assessment criteria as well as instructions on the preparation and submission of the assignments. This tutorial letter also provides all the information you need with regard to the prescribed study material and other resources and how to obtain it. Please study this information carefully and make sure that you obtain the prescribed material as soon as possible.

We have also included certain general and administrative information about this module. Please study this section of the tutorial letter carefully.

Right from the start we would like to point out that you must read all the tutorial letters you receive during the semester immediately and carefully, as they always contain important and, sometimes, urgent information.

We hope that you will enjoy this module and wish you all the best!

1.1 Tutorial matter

Take note that every tutorial letter you will be receiving is important and you have to read them all immediately and carefully. Some information contained in these tutorial letters may be urgent, while others may, for example, contain examination information. So, it is wise to keep them all in a file!

Some of this tutorial matter may not be available when you register. Tutorial matter that is not available when you register will be posted to you as soon as possible, but is also available on myUnisa.

At the time of registration, you will receive an inventory letter that will tell you what you have received in your study package and also show items that are still outstanding. Also see the brochure entitled *my Studies @ Unisa*.

Check the study material that you have received against the inventory letter. You should have received all the items listed in the inventory, unless there is a statement like “out of stock” or “not

available". If any item is missing, follow the instructions on the back of the inventory letter without delay.

Shortly after registration The Department of Despatch should supply you with the following tutorial matter for this module:

- **Tutorial letter 101**. Read this carefully and save it for future reference as it contains important information *as well as your assignments for the semester.*
- **A study guide** written by a lecturer to guide you through the relevant sections in the prescribed book. Use it together with the textbook as the guide indicates the relevant prescribed sections, explaining difficult concepts in more detail, giving additional examples and exercises, etc.
- Other tutorial letters to further assist you with your studies, will be dispatched to you throughout the year.

If you have access to the Internet, you can view the study guide and tutorial letters for the modules for which you are registered on the University's online campus, myUnisa, at http://my.unisa.ac.za

There are two types of tutorial letters:

- The 100-series (e.g. Tutorial letter 101, 102, 103, etc.) containing general information, assignment questions, information about your lecturer or the examination, a trial paper, etc.
- The 200-series (e.g. Tutorial letter 201, 202, 203, etc.) containing the solutions to the assignments and the trial paper.

2 PURPOSE OF AND OUTCOMES FOR THE MODULE

2.1 Purpose

This module enables students to gain insight in statistical inference using different properties of estimation and methods of estimation. Included are linear models and estimation by least square as well as designing experiments and analysis of variance procedures. This module will support further studies and application in the sector of statistics theory in the field of statistics, as part of the Bachelor of Science and Bachelor of commerce qualifications. This module will be an illustration of Mathematical Statistics as a theory of information to contribute to the development of communities and of research in Southern Africa, Africa or globally utilizing mathematics extensively, but only as a tool.

2.2 Outcomes

Qualifying students will be able to:

- State confidently what theory of statistical estimation is all about, define the properties of estimators: efficiency, consistency and sufficiency.
- Define the linear model in matrix notation, and calculate the least squares regression coefficients of the model using matrices.

- Introduce more advanced tools for data analysis.
- Two factors affect the quantity of information in a sample, the variation in the data and the sample size, concerned with methods to reduce the variation in a sample, at minimum cost.
- Analyse the variation in a set of responses and assign portions of this variation to each variable in a set of independent variables.

3 LECTURER(S) AND CONTACT DETAILS

3.1 Lecturer(s)

The lecturer responsible for this module is:

Prof P Ndlovu
Room: 6-22, Florida Science Campus
GJ Gerwel Building
Tel: 011 670 9250
Email: ndlovp@unisa.ac.za

You are most welcome to contact me if you encounter any difficulties with the tutorial matter. You are also welcome to come and see me, but please telephone beforehand so that we can arrange a time which is convenient for all of us. If you cannot reach me, the department's secretary may be able to locate me or take a message. Her telephone number is (011) 670 9255.

PLEASE NOTE: Letters to lecturers may not be enclosed with or inserted into assignments.

3.2 University

If you need to contact the University about matters not related to the content of this module, please consult the publication *My Studies @ Unisa* that you received with your study material. This brochure contains information on how to contact the University (e.g. to whom you can write for different queries, important telephone and fax numbers, addresses and details of the times certain facilities are open).

Always have your student number at hand when you contact the University.

4 MODULE RELATED RESOURCES

4.1 Prescribed books

The prescribed book for this module is

Wackerly, Dennis D; Mendenhall, William III and Scheaffer, Richard L;
Mathematical statistics with applications (2008), 7th ed

You have to buy this book. Please consult the list of official booksellers and their addresses listed in my Studies @ Unisa. Prescribed books can be obtained from the University's official booksellers. If you have difficulty locating your book(s) at these booksellers, please contact the Prescribed Books Section at 012 429 4152 or e-mail vospresc@unisa.ac.za.

For shorter reference, we use **WMS** when we need to draw your attention to the prescribed book.

4.2 Recommended books

There are no recommended books for this module.

4.3 Electronic Reserves (e-Reserves)

There are no e-Reserves for this module.

4.4 Library services and resources information

For brief information go to : http://www.unisa.ac.za/contents/studies/docs/myStudies-at-Unisa2018-brochure.pdf
For more detailed information, go to the Unisa website: http://www.unisa.ac.za/, click on Library
For research support and services of Personal Librarians, go to:
http://www.unisa.ac.za/Default.asp?Cmd=ViewContent&ContentID=7102
The Library has compiled numerous library guides:

- find recommended reading in the print collection and e-reserves -http://libguides.unisa.ac.za/request/undergrad
- request material - http://libguides.unisa.ac.za/request/request
- postgraduate information services - http://libguides.unisa.ac.za/request/postgrad
- finding , obtaining and using library resources and tools to assist in doing research http://libguides.unisa.ac.za/Research_Skills
- how to contact the Library/find us on social media/frequently asked questions - http://libguides.unisa.ac.za/ask

5 STUDENT SUPPORT SERVICES FOR THE MODULE

For information on the various student support systems and services available at Unisa (e.g. student counseling, tutorial classes, language support), please consult the publication *my Studies @ Unisa* that you received with your study material.

5.1 Contact with Fellow Students

5.1.1 Study Groups

It is advisable to have contact with fellow students. One way to do this is to form study groups. **Please consult the publication *my Studies@Unisa* to find out how to obtain the addresses of students in your region.**

5.1.2 myUnisa

If you have access to a computer that is linked to the internet, you can quickly access resources and information at the University. The *myUnisa* learning management system is Unisa's online campus that will help you to communicate with your lecturers, with other students and with the administrative departments of Unisa - all through the computer and the internet.

To go to the *myUnisa* website, start at the main Unisa website, http://www.unisa.ac.za, and then click on the "Login to *myUnisa*" link on the right-hand side of the screen. This should take you to the *myUnisa* website. You can also go there directly by typing in http://my.unisa.ac.za.

Please consult the publication *my Studies @ Unisa* which you received with your study material for more information on *myUnisa*.

5.1.3 Discussion classes

There are no discussion classes offered in this module. Should the need for discussion classes arise in future, students will be informed well in advance about actual dates and venues.

5.2 Free computer and internet access

Unisa has entered into partnerships with establishments (referred to as Telecentres) in various locations across South Africa to enable you (as a Unisa student) free access to computers and the Internet. This access enables you to conduct the following academic related activities: registration; online submission of assignments; engaging in e-tutoring activities and signature courses; etc. Please note that any other activity outside of these are for your own costing e.g. printing, photocopying, etc. For more information on the Telecentre nearest to you, please visit www.unisa.ac.za/telecentres.

6 MODULE-SPECIFIC STUDY PLAN

SEMESTER 1	Chapters for preparing your assignments	From	To
Assignment 1	Chapters 8 (8.1, 8.2), 9 and 11	Registration	5 March 2018
	Start writing your assignment	5 March 2018	12 March 2018
Assignment 2	Chapters 12 and 13	12 March 2018	9 April 2018
	Start writing your assignment	9 April 2018	16 April 2018

SEMESTER 2	Chapters for preparing your assignments	From	To
Assignment 1	Chapters 8 (8.1, 8.2), 9 and 11	Registration	20 August 2018
	Start writing your assignment	20 August 2018	27 August 2018
Assignment 2	Chapters 12 and 13	27 August 2018	21 September 2018
	Start writing your assignment	21 September 2018	27 September 2018

7 MODULE PRACTICAL WORK AND WORK-INTEGRATED LEARNING

There are no practicals for this module.

8 ASSESSMENT

8.1 Assessment criteria

The outcomes of this module are given in Section 2.2 of this tutorial letter. These outcomes describe what you should be able to do in order to successfully pass this module. Assignments, examinations, and in some modules projects and portfolios are the ways we use to assess whether you have reached the outcomes.
The criteria we use to assess your work can be summarised as follows:

- You must apply the correct and appropriate formulas, presentations, methods, rules, laws, values from tables, and so on, as required in the question.
- Applying of formulas, methods etc. must be done correctly.
- Results, tests, computer printouts etc. should be interpreted correctly, when you are asked to do so.
- Calculations must be correct and accurate.
- You may be asked to prove/show/verify certain theoretical results.

The following general comments are valid to all our modules. In some cases the lecturers will give further instructions to keep in mind when completing your work; these will be given in the tutorial letters for that particular module.

8.1.1 Written assignment and examination questions

Please keep the following in mind when answering questions.

- Read the question carefully – you will get zero marks if you end up answering what was not asked for!
- Give full calculations, marks will usually not be given for the end results only.
- Present your solutions clearly. A collection of disjointed formulas and numbers is not the right way to answer questions, please use words to explain what you are doing and why. Use correct mathematical notation and remember that lines of mathematical equations must always be linked to each other – for example with the = sign if they are a series of continuing calculations, or otherwise maybe by the signs for “equals” or “therefore”. See your textbooks and/or study guides for examples.

8.1.2 Multiple choice questions

- Only one of the given answers is correct. If you believe several to be correct, check your work again!
- We suggest you keep copies of your calculations, so that when you get the results, you can check where you went wrong.

8.2 Assessment plan

The assessment in this module consists of two assignments and an examination.

Your final mark for the module is determined from your year mark and your examination mark. The year mark forms 20% and the examination mark 80% of the final mark. The year mark is the average of the marks you receive for assignments 1 and 2. An assignment submitted late or not at all will give you 0%. If you do well in your assignments you have a good year mark and that can make all the difference between a pass or fail or between a distinction or simply a pass!

The two assignments prescribed for this module must be seen as part of the learning process. The typical assignment question is a reflection of a typical examination question. There are fixed submission dates for the assignments and each assignment is based on specific chapters in the prescribed book. You have to adhere to these dates as assignments are only marked if they are received on or before the due dates.

You will only get examination admission if you submit the first assignment by its due date. You should complete both assignments as well as you can, since

- they are the sole contributors towards your year mark,
- they form an integral part of the learning process and indicate the form and nature of the questions you can expect in the examination.

Assignments and Learning

Assignments are seen as part of the learning material for this module. As you do the assignment, study the reading texts, consult other resources, discuss the work with fellow students or tutors or do research, you are actively engaged in learning. Looking at the assessment criteria given for each assignment, and the feedback you receive in your marked assignment, will help you to understand what is required of you more clearly.

8.3 General assignment numbers

The two assignments are numbered 01 and 02 for each semester.

8.3.1 Unique assignment numbers

Please note that each assignment has its unique six-digit assignment number which has to be written on the cover of your assignment upon submission. The unique numbers are given later on in this tutorial letter; you will find them in the heading of each set of assignment questions.

8.3.2 Due dates for assignments

Assignment for SEMESTER 1	Sections from the following Chapters are covered	Due Date
1	**WMS**: Chapters 8 (8.1, 8.2), 9 and 11 Study Guide: Units 1 to 4	13 March 2018
2	**WMS**: Chapters 12 and 13 Study Guide: Units 5 and 6	17 April 2018

Assignment for SEMESTER 2	Sections from the following Chapters are covered	Due Date
1	**WMS**: Chapters 8 (8.1, 8.2), 9 and 11 Study Guide: Units 1 to 4	28 August 2018
2	**WMS**: Chapters 12 and 13 Study Guide: Units 5 and 6	28 September 2018

8.4 Submission of assignments

For detailed information on assignments, please refer to the *my Studies @ Unisa* brochure, which you received with your study package.
To submit an assignment via myUnisa:

- Go to myUnisa.
- Log in with your student number and password.
- Select the module.
- Click on assignments in the menu on the left-hand side of the screen.
- Click on the assignment number you wish to submit.
- Follow the instructions.

For general information and requirements as far as assignments are concerned, see the brochure *my Studies @ Unisa* which you received with your study material.

8.5 Assignments

This tutorial letter 101 contains the assignments for both semesters, so select the semester you are enrolled for and do the set of assignments for that semester only. The assignments for Semester 1 are in Appendix A, pages 14–19. The assignments for Semester 2 are in Appendix B, pages 20–25. Solutions to the assignments will be available on *myUnisa* to ALL students registered for this module a while after the closing date of the relevant assignment.

9 OTHER ASSESSMENT METHODS

There are no other assessment methods for this module.

10 EXAMINATION

10.1 Examination Admission

You need to have a final mark of 50% to pass this module and 75% to obtain a distinction.

In this module a maximum of 20 marks is added to your examination mark (out of 80) to form your final mark. This 20% contribution comes from the marks you obtained for the two assignments and is called your year mark. If you do well in your assignments you have a good year mark and that can make all the difference between a pass or fail or between a distinction or simply a pass!

Currently admission to the examination is only based on the proof that you are actively involved in your studies. This proof is based on the **submission of your first assignment** before a fixed given date. Admission therefore does not rest with the department and if you do not submit that particular assignment in time, we can do nothing to give you admission. Although you are most probably a part time student with many other responsibilities, work circumstances will not be taken into consideration for exemption from assignments or the eventual admission to the examination.

No concession will be made to students who do not qualify for the examination.

10.2 Examination Period

This module is offered in a semester period of fifteen weeks. This means that

- if you are registered for the first semester, you will write the examination in May/June 2018 and should you fail and qualify for a supplementary examination, that supplementary examination will be written in October/November 2018.
- if you are registered for the second semester, you will write the examination in October/November 2018 and should you fail and qualify for a supplementary examination, that supplementary examination will be written in May/June 2019.

The examination section will provide you with information regarding the examination in general, examination venues, examination dates and examination times. Eventually, your results will also be processed by them and sent to you.

10.3 Examination Paper

Your examination will be a **2 hour examination.** The questions will be similar to the assignment questions, but there will also be questions on theory. Should you have a final mark of less than 50%, it implies that you failed the module STA2602. However, should your results be within a specified percentage (usually from 40% to 49%), you will be given a second chance in the form of a *supplementary* examination on the dates as specified in 10.2. If you fail the examination with less than 40%, the year mark will not count to help you pass.

10.4 Previous Examination Papers

Previous examination papers are not available to students. However, you will receive a *trial* paper towards the end of the semester that you can use as an indication of typical examination questions. Solutions to this trial paper is also sent out in a follow-up tutorial letter. Remember

that the examples, exercises, activities in the guide as well as your assignment questions are also indicators of typical examination questions.

10.5 Tutorial Letter with Information on the Examination

As mentioned before, you will receive a tutorial letter containing a trial paper. Should the lecturer want to discuss any matter about the examination, it will be included in this tutorial letter. In the study guide you are given clear indications of the sections in the textbook that you have to know and can be tested on in the examination. Remember that you have to work continuously and do not treat statistics as any other subject, where it may be possible to study only selected sections of the work. All the topics are interlinked and you will definitely run into trouble if you skip sections!

You are automatically admitted to the exam on the submission of Assignment 01 by a specific date – see Section 8.1. Please note that lecturers are not responsible for exam admission, and ALL enquiries about exam admission should be directed by e-mail to exams@unisa.ac.za.

11 FREQUENTLY ASKED QUESTIONS

The my Studies @ Unisa brochure contains an A-Z guide of the most relevant study information. Please refer to this brochure for any other questions.

12 SOURCES CONSULTED

No books other than the prescribed book was consulted in preparing this tutorial letter.

13 CONCLUSION

Remember that there are no "short cuts" to studying and understanding statistics. You need to be dedicated, work consistently and practise, practise and practise some more! We trust that you will find a depth of knowledge in STA2602 that you can apply in many aspects of your life.
Be positive, determined and eager to learn and you will be successful!

Prof P Ndlovu

ADDENDUM A: FIRST SEMESTER ASSIGNMENTS

A.1 Assignment 01

ONLY FOR SEMESTER 1 STUDENTS
ASSIGNMENT 01
Unique Nr.: 840133
Fixed closing date: 13 March 2018

QUESTION 1 **[25]**

(a) Let $\hat{\theta}$ be an estimator of θ. Using the identity

$$(\hat{\theta} - \theta)^2 = [(\hat{\theta} - E(\hat{\theta})) + (E(\hat{\theta}) - \theta)]^2,$$

show that

$$\begin{aligned} MSE(\hat{\theta}) &= E[(\hat{\theta} - E(\theta))^2] + (E(\hat{\theta}) - \theta)^2 \\ &= V(\hat{\theta}) + [B(\hat{\theta})]^2 \end{aligned}$$

where $B(\hat{\theta})$ is the bias of $\hat{\theta}$. **(6)**

(b) Suppose that $\hat{\theta}_1$ and $\hat{\theta}_2$ are **independent unbiased** estimators of θ. Consider another estimator

$$\hat{\theta}_3 = \alpha\hat{\theta}_1 + (1 - \alpha)\hat{\theta}_2$$

where the real number α is in the interval $[0, 1]$.

(i) Show that $\hat{\theta}_3$ is also an unbiased estimator of θ. **(3)**

(ii) If $V(\hat{\theta}_1) = V(\hat{\theta}_2) = \sigma^2$, what is $V(\hat{\theta}_3)$ in terms of σ^2 and α? **(4)**

(iii) Refer to part (ii) above. Find the value of α which minimizes $V(\hat{\theta}_3)$. **(5)**

(iv) If $V(\hat{\theta}_1) = \sigma^2$ and $V(\hat{\theta}_2) = 2\sigma^2$, what is the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_3$ in terms of α? **(5)**

(v) Evaluate the efficiency in part (iv) above using the value of α you found in part (iii). **(2)**

QUESTION 2 **[8]**

Suppose that $\hat{\theta}_n$ is an estimator of θ based on a random sample of size n. Another equivalent definition of the consistency of $\hat{\theta}_n$ as an estimator of θ is that

$$\lim_{n\to\infty} V(\hat{\theta}_n) = 0 \text{ and } \lim_{n\to\infty} E(\hat{\theta}_n) = \theta.$$

Let $X_1, X_2, ..., X_n$ be a random sample from the uniform distribution on the interval $(0, \theta)$, and let

$$Y_{(n)} = \max\{X_1, X_2, ..., X_n\}.$$

It can be shown that

$$E(Y_{(n)}) = \left(\frac{n}{n+1}\right)\theta \text{ and } V(Y_{(n)}) = \left(\frac{n}{(n+2)(n+1)^2}\right)\theta^2.$$

Show that $\hat{\theta}_n = \left(\frac{n+1}{n}\right) Y_{(n)}$ is a consistent estimator of θ.

QUESTION 3 **[7]**

Let $X_1, X_2, ..., X_n$ be a random sample from a distribution with probability density function

$$f(x, \beta) = \begin{cases} \beta e^{-\beta x} & \text{for } x \geq 0 \\ 0 & \text{elsewhere.} \end{cases}$$

(a) What is the likelihood ($L(\beta) = L(x_1, x_2, ..., x_n|\beta)$) of the sample? **Simplify it.** **(3)**

(b) Use the factorization criterion/theorem to show that $\sum_{i=1}^{n} X_i$ is a sufficient statistic for β. **(4)**

QUESTION 4 **[10]**

Refer to QUESTION 3 above. $E(X_i) = \frac{1}{\beta}$.

(a) Find the method-of-moments estimator of β. **(3)**

(b) Find the maximum likelihood estimator of β. **(7)**

QUESTION 5 **[20]**

Suppose that $Y_1, Y_n, ..., Y_n$ independent variables such that

$$Y_i = \beta_0 + \epsilon_i, \ i = 1, 2, ..., n;$$

where β_0 is an unknown parameter, and $\epsilon_1, \epsilon_2, ..., \epsilon_n$ are independent random errors each with a normal distribution with mean 0 and variance σ^2.

(a) Find $E(Y_i)$ and hence the method-of-moments estimator of β_0. What is the variance of the estimator? **(2+3+1)**

(b) Give that the probability density function of Y_i is

$$f(x, \beta_0, \sigma^2) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y-\beta_0)^2} & \text{for } -\infty < x < \infty \\ 0 & \text{elsewhere,} \end{cases}$$

show that the maximum likelihood estimator of β_0 is the same as the method-of-moments estimator of β_0. **(7)**

(c) The least squares estimator of β_0 is the value of β_0 which minimizes

$$SSE = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (Y_i - \beta_0)^2.$$

Show that the least squares estimator of β_0 is the same as the maximum likelihood estimator of β_0. **(7)**

QUESTION 6 **[30]**

An appliance store conducted a 5-month experiment to determine the relationship between advertising expenditure (x) and sales revenue (y). The data are given in the table below.

Month i	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	1	1	1	1	1
2	2	1	4	1	2
3	3	2	9	4	6
4	4	2	16	4	8
5	5	4	25	16	20
Totals	$\sum_{i=1}^{5} x_i = 15$	$\sum_{i=1}^{5} y_i = 10$	$\sum_{i=1}^{5} x_i^2 = 55$	$\sum_{i=1}^{5} y_i^2 = 26$	$\sum_{i=1}^{5} x_i y_i = 37$

(a) Calculate S_{xx}, S_{yy}, and S_{xy}. **(6)**

(b) Calculate the correlation coefficient (r) and interpret it. **(3+2)**

(c) Fit the simple linear regression model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \ i = 1, 2, ..., 5;$$

(where $\epsilon_1, \epsilon_2, ..., \epsilon_5$ are assumed to be independent random errors each with a normal distribution with mean 0 and variance σ^2) to the data. That is, obtain the least squares estimates of β_0 and β_1. **(3+3)**

(d) What is the estimate of σ^2? **(3)**

(e) What is the estimate of $V(\hat{\beta}_1)$? **(2)**

(f) Test the hypotheses H_0 $\beta_1 = 0$ versus H_a $\beta_1 \neq 0$ at the 0.05 level of significance. **(3+3+2)**

Total: [100]

A.2 Assignment 02

ONLY FOR SEMESTER 1 STUDENTS
ASSIGNMENT 02
Unique Nr.: 831048
Fixed closing date: 17 April 2018

QUESTION 1 [10]

Suppose that it is desired to conduct an experiment to compare the effects of three types of fertiliser (A, B, C) on the **yield** of a certain variety of maize grown under the natural conditions of South Africa. **Three uniform one-acre plots** are available for experimentation in each of the **three research stations**: $Limpopo$, $Mpumalanga$, KZN. In terms of the suitability of the environmental conditions for maize growth, KZN has the best followed by $Mpumalanga$.

(a) What is the response variable in the experiment? **(2)**

(b) What is the treatment factor to be studied in the experiment? **(2)**

(c) What are the experimental units? **(2)**

(d) Suggest a blocking factor for the experiment. **Justify your answer.** **(4)**

QUESTION 2 [35]

An experiment was conducted to compare **three** different insecticides (A, B, C) on the **number** of seedlings, of a particular variety of string bean, that emerged per **subplot**. **Four different plots** (in terms of moisture content, fertility, etc.) of the same size were prepared, with each plot divided into **three subplots of the same size.** A suitable distance was maintained between the subplots within each plot. Each subplot was planted with 100 seeds and then maintained under the insecticide randomly assigned to it.

(a) Identify:

(i) the response variable; **(2)**
(ii) treatment factor; **(2)**
(iii) the experimental units; and **(2)**
(iv) the blocking factor (**with justification**) in the experiment. **(4)**

(b) The data collected from the experiment are given in the table below.

	Plot				
Insecticide	1	2	3	4	Mean
A	56	49	65	60	57.50
B	84	78	94	93	87.25
C	80	72	83	85	80.00

(i) Complete the ANOVA tables below. **(8)**

ANOVA Table I

Source	df	SS	MS	F
Insecticide	??	1925.17	??	??
Error	??	??	??	
Total	??	2334.92		

ANOVA Table II

Source	df	SS	MS	F
Insecticide	??	1925.17	??	??
Plot	??	386.25	??	??
Error	??	??	??	
Total	??	2334.92		

(ii) Why are the estimates of experimental errors $(MSE's)$ in the two tables different? **(3)**

(iii) Use an **appropriate** $ANOVA$ (I or II?) to compare the effects of the three insecticides on the number of seedlings that emerge per subplot. **Use the 0.05 level of significance.** **(5)**

(iv) Use an **appropriate** estimate of the experimental error (MSE) to construct a 95% confidence interval for the difference between the B and C insecticide means. **(9)**

QUESTION 3 **[15]**

The data in the table below was collected for the purpose of comparing the mean daily sales of two restaurants in the same city.

Day	Restaurant I	Restaurant II	(Restaurant I - Restaurant II)
1 (Monday)	759	678	81
2 (Tuesday)	981	933	48
3 (Wednesday)	1005	918	87
4 (Thursday)	1449	1302	147
5 (Friday)	1905	1782	123
6 (Saturday)	2073	1971	102
7 (Monday)	693	639	54
8 (Tuesday)	873	825	48
9 (Wednesday)	1074	999	75
10 (Thursday)	1338	1281	57
11 (Friday)	1932	1827	105
12 (Saturday)	2106	2049	57
Sample mean	1349	1267	82
Sample variance	280978.9	266294.2	1023.3

(a) Which between the $pooled-t$ and the $paired-t$ tests is appropriate for comparing the mean daily sales of the two restaurants? **(4)**

(b) Is the there sufficient evidence to conclude that the mean daily sales of Restaurant I is greater than the mean daily sales of Restaurant II? **Use the 0.05 level of significance.** **(11)**

QUESTION 4 **[25]**

Data in the table below was collected for the purpose of comparing the mean distances traveled by four brands (A, B, C, D) of golf balls when struck with a driver by the same robotic golfer.

	Brand A	Brand B	Brand C	Brand D
	251.2	263.2	269.7	251.6
	245.1	262.9	263.2	248.6
	248.0	265.0	277.5	249.4
	251.1	254.5	267.4	242.0
	265.5	264.3	270.5	246.5
	250.0	257.0	265.5	251.3
	253.9	262.8	270.7	262.8
	244.6	264.4	272.9	249.0
	254.6	260.6	275.6	247.1
	248.8	255.9	266.5	245.9
Mean	251.3	261.1	270.0	249.4

(a) Complete the one-way $ANOVA$ table (of the above data) below. **(15)**

Source	df	SS	MS	F
Brand	??	??	??	??
Error	??	??	??	
Total	??	??		

(b) Do the data provide sufficient evidence to indicate a difference among the brand means? **Use the 0.05 level of significance.** **(5)**

(c) Under what assumptions about the errors in the data are your conclusions in part (b) valid? **(5)**

Total: [85]

ADDENDUM B: SECOND SEMESTER ASSIGNMENTS

B.1 Assignment 01

ONLY FOR SEMESTER 2 STUDENTS
ASSIGNMENT 01
Unique Nr.: 741040
Fixed closing date: 28 August 2018

QUESTION 1 **[25]**

(a) Let $\hat{\theta}$ be an estimator of θ. Using the identity

$$(\hat{\theta} - \theta)^2 = [(\hat{\theta} - E(\hat{\theta})) + (E(\hat{\theta}) - \theta)]^2,$$

show that

$$\begin{aligned} MSE(\hat{\theta}) &= E[(\hat{\theta} - E(\theta))^2] + (E(\hat{\theta}) - \theta)^2 \\ &= V(\hat{\theta}) + [B(\hat{\theta})]^2 \end{aligned}$$

where $B(\hat{\theta})$ is the bias of $\hat{\theta}$. **(6)**

(b) Suppose that $\hat{\theta}_1$ and $\hat{\theta}_2$ are **independent unbiased** estimators of θ. Consider another estimator

$$\hat{\theta}_3 = \alpha\hat{\theta}_1 + (1 - \alpha)\hat{\theta}_2$$

where the real number α is in the interval $[0, 1]$.

(i) Show that $\hat{\theta}_3$ is also an unbiased estimator of θ. **(3)**

(ii) If $V(\hat{\theta}_1) = V(\hat{\theta}_2) = \sigma^2$, what is $V(\hat{\theta}_3)$ in terms of σ^2 and α? **(4)**

(iii) Refer to part (ii) above. Find the value of α which minimizes $V(\hat{\theta}_3)$. **(5)**

(iv) If $V(\hat{\theta}_1) = \sigma^2$ and $V(\hat{\theta}_2 = 3\sigma^2$, what is the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_3$ in terms of α? **(5)**

(v) Evaluate the efficiency in part (iv) above using the value of α you found in part (iii). **(2)**

QUESTION 2 **[8]**

Suppose that $\hat{\theta}_n$ is an estimator of θ based on a random sample of size n. Another equivalent definition of the consistency $\hat{\theta}_n$ as an estimator of θ is that

$$\lim_{n\to\infty} V(\hat{\theta}_n) = 0 \text{ and } \lim_{n\to\infty} E(\hat{\theta}_n) = \theta.$$

Let $X_1, X_2, ..., X_n$ be a random sample from a normal distribution with mean 0 and variance θ, and let

$$\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

It can be shown that

$$E\left(\frac{n\hat{\theta}_n}{\theta}\right) = n \text{ and } V\left(\frac{n\hat{\theta}_n}{\theta}\right) = 2n.$$

Show that $\hat{\theta}_n$ is a consistent estimator of θ.

QUESTION 3 **[7]**

Let $X_1, X_2, ..., X_n$ be a random sample from a distribution with probability mass function

$$p(x, \beta) = \begin{cases} \beta(1-\beta)^{x-1} & \text{for } 0 < \beta < 1 \text{ and } x = 1, 2, ... \\ 0 & \text{elsewhere.} \end{cases}$$

(a) What is the likelihood ($L(\beta) = L(x_1, x_2, ..., x_n|\beta)$) of the sample? **Simplify it.** **(3)**

(b) Use the factorization criterion/theorem to show that $\sum_{i=1}^{n} X_i$ is a sufficient statistic for β. **(4)**

QUESTION 4 **[10]**

Refer to QUESTION 3 above. $E(X_i) = \frac{1}{\beta}$.

(a) Find the method-of-moments estimator of β. **(3)**

(b) Find the maximum likelihood estimator of β. **(7)**

QUESTION 5 **[20]**

Suppose that $Y_1, Y_n, ..., Y_n$ independent variables such that

$$Y_i = \beta x_i + \epsilon_i, \ i = 1, 2, ..., n;$$

where β_0 is an unknown parameter, $x_1, x_2, ..., x_n$ are **known** real numbers, and $\epsilon_1, \epsilon_2, ..., \epsilon_n$ are independent random errors each with a normal distribution with mean 0 and variance σ^2.

(a) Show that $\frac{\bar{Y}}{\bar{x}}$ is an unbiased estimator of β. What is the variance of the estimator? **(4+2)**

Hint: $\bar{Y} = \beta\bar{x} + \bar{\epsilon}$ where $\bar{\epsilon} = \frac{1}{n}\sum_{i=1}^{n} \epsilon_i$.

(b) Give that the probability density function of Y_i is

$$f(x, \beta, \sigma^2) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y-\beta x_i)^2} & \text{for } -\infty < x < \infty \\ 0 & \text{elsewhere,} \end{cases}$$

show that the maximum likelihood estimator of β is **not** the same as the estimator of β in part (a). **(7)**

(c) The least squares estimator of β is the value of β which minimizes

$$SSE = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (Y_i - \beta x_i)^2.$$

Show that the least squares estimator of β is the same as the maximum likelihood estimator of β. **(7)**

QUESTION 6 **[30]**

A firm collected 10-year data to determine the relationship between the yearly number of sales people (x) and the yearly sales revenue (y). The data are given in the table below.

Year i	x_i	y_i
1981	15	1.35
1982	18	1.63
1983	24	2.33
1984	22	2.41
1985	25	2.63
1986	29	2.93
1987	30	3.41
1988	32	3.26
1989	35	3.63
1990	38	4.15

$$\sum_{i=1}^{10} x_i = 268, \sum_{i=1}^{10} y_i = 27.73, \sum_{i=1}^{10} x_i^2 = 7668, \sum_{i=1}^{10} y_i^2 = 83.8733, \sum_{i=1}^{10} x_i y_i = 800.62$$

(a) Calculate S_{xx}, S_{yy}, and S_{xy}. **(6)**

(b) Calculate the correlation coefficient (r) and interpret it. **(3+2)**

(c) Fit the simple linear regression model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \ i = 1, 2, ..., 5;$$

(where $\epsilon_1, \epsilon_2, ..., \epsilon_5$ are assumed to be independent random errors each with a normal distribution with mean 0 and variance σ^2) to the data. That is, obtain the least squares estimates of β_0 and β_1. **(3+3)**

(d) What is the estimate of σ^2? **(3)**

(e) What is the estimate of $V(\hat{\beta}_1)$? **(2)**

(f) Test the hypotheses H_0 $\beta_1 = 0$ versus H_a $\beta_1 \neq 0$ at the 0.05 level of significance. **(3+3+2)**

Total: [100]

B.2 Assignment 02

ONLY FOR SEMESTER 2 STUDENTS
ASSIGNMENT 02
Unique Nr.: 869711
Fixed closing date: 28 September 2018

QUESTION 1 [10]

Suppose that it is desired to investigate the effects of the **Rate of Application** of a certain fertilizer (at levels 0 kg/ha; 100 kg/ha; 300 kg/ha), and the **Watering Frequency** (at frequencies once/week; twice/week) on the yield of a certain variety of maize grown under the natural conditions of South Africa. **Six uniform one-acre plots** are available for experimentation in each of the **three research stations**: $Limpopo$, $Mpumalanga$, KZN. In terms of the suitability of the environmental conditions for maize growth, KZN has the best followed by $Mpumalanga$.

(a) What is the response variable in the experiment? **(2)**

(b) What are the treatment factors to be studied in the experiment? **(2)**

(c) What are the experimental units? **(2)**

(d) Suggest a blocking factor for the experiment. **Justify your answer.** **(4)**

QUESTION 2 [35]

An experiment was conducted to compare the mileage (km) per litre obtained by competing brands of petrol (I, II, III). Three new Mazda 323's, 3 new Toyota Camrys, and 3 new Nissan Sentras were used in the experiment. During the experiment the cars were operated under the same conditions in order to eliminate the effects of external variables on the distance travelled per assigned brand of petrol.

(a) Identify:

(i) the response variable; **(2)**
(ii) treatment factor; **(2)**
(iii) the experimental units; and **(2)**
(iv) the blocking factor (**with justification**) in the experiment. **(4)**

(b) The data collected from the experiment are given in the table below.

	Type of Car		
Brand	Mazda	Toyota	Nissan
I	10.6	12.0	11.0
II	9.0	15.0	12.0
III	12.0	17.4	13.0

(i) Complete the ANOVA tables below. **(8)**

ANOVA Table I

Source	df	SS	MS	F
Brand	??	13.80	??	??
Error	??	??	??	
Total	??	49.34		

ANOVA Table II

Source	df	SS	MS	F
Brand	??	13.80	??	??
Car	??	28.20	??	??
Error	??	??	??	
Total	??	49.34		

(ii) Why are the estimates of experimental errors $(MSE's)$ in the two tables different? **(3)**

(iii) Use an **appropriate** $ANOVA$ (I or II?) to compare the mean mileage of the brands. **Use the 0.05 level of significance.** **(5)**

(iv) Use an **appropriate** estimate of the experimental error (MSE) to construct a 95% confidence interval for the difference between the brands I and III means. **(9)**

QUESTION 3 **[15]**

The data in the table below is of starting annual salaries of randomly chosen 10 pairs of male and female university graduates. Each pair has the same degree, and similar average marks of the core modules for the awarded degree.

Pair	Male I	Female II	(Male - Female)
1	24300	23800	500
2	26500	26600	-100
3	25400	24800	600
4	23500	23500	0
5	28500	27600	900
6	22800	23000	-200
7	24500	24200	300
8	26200	25100	1100
9	23400	23200	200
10	24200	23500	700
Sample mean	24930	24530	400
Sample variance	3009000.00	2331222.22	188888.89

(a) Which between the pooled$-t$ and the paired$-t$ tests is appropriate for comparing the mean starting annual salaries of the male and female graduates? **(4)**

(b) Is the there sufficient evidence to conclude that the mean starting annual salary of male graduates is greater than the mean starting annual salary of female graduates? **Use the 0.05 level of significance.** **(11)**

QUESTION 4 **[25]**

A randomized design was used to compare the mileage (km) per litre obtained by competing brands of petrol. Fifteen identical cars were randomly assigned to brands I, II and III, with each brand assigned five cars. The cars were operated under the same conditions and the distance travelled by each car per liter of the assigned brand of petrol recorded. The data are given in the table below.

	Brand I	Brand II	Brand III
	10.5	11.5	11.0
	12.0	13.0	11.5
	14.0	15.0	14.5
	16.0	17.0	15.5
	19.0	11.5	11.05
Mean	14.3	13.6	12.7

(a) Complete the one-way $ANOVA$ table (of the above data) below. **(15)**

Source	df	SS	MS	F
Brand	??	??	??	??
Error	??	??	??	
Total	??	??		

(b) Do the data provide sufficient evidence to indicate a difference among the brand means? **Use the 0.05 level of significance.** **(5)**

(c) Under what assumptions about the errors in the data are your conclusions in part (b) valid? **(5)**

Total: [85]