

# APPLIED STATISTICS MODULE

VERSION 2

SCIENCE ⚙️ ENGINEERING ⚙️ TECHNOLOGY

SITI ZANARIAH SATARI   NORYANTI MUHAMMAD   SITI ROSLINDAR YAZIZ  
NOR HAFIZAH MOSLIM   ROSLINAZAIRIMAH ZAKARIA   FARAHAANIM MISNI



# APPLIED STATISTICS MODULE

VERSION 2



# **APPLIED STATISTICS MODULE**

**VERSION 2**

**SCIENCE ⚙ ENGINEERING ⚙ TECHNOLOGY**

**SITI ZANARIAH SATARI   NORYANTI MUHAMMAD   SITI ROSLINDAR YAZIZ  
NOR HAFIZAH MOSLIM   ROSLINAZAIRIMAH ZAKARIA   FARAHANIM MISNI**

Publisher  
Universiti Malaysia Pahang  
Kuantan  
2017

Copyright © Universiti Malaysia Pahang, 2017

First Published, 2017

All right reserved.

Apart from fair dealing for the purpose of study, research, criticism or review, as permitted under the Copyright Act, no part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher. Enquiries to be made to the author and the publisher Penerbit Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang Darul Makmur. Negotiation is subject to royalty arrangement or honorarium.

Perpustakaan Negara Malaysia

Cataloguing-in-Publication Data

Applied statistics module new version: science, engineering, technology/

Siti Zanariah Satari ... (et al.)

ISBN 978-967-0691-64-0

1. Statistics 2. Science-Statistical methods.

3. Engineering-Statistical methods. 4. Technology-Statistical methods.

I. Siti Zanariah Satari

519.5

Published By:

**Publisher**

Universiti Malaysia Pahang,

Lebuhraya Tun Razak,

Pahang Darul Makmur.

Tel: 09-549 3320 Fax: 09-549 3381

Printing:

**Syarikat Percetakan Inderapura Sdn. Bhd. (8763 H)**

Jalan Tanjong Api Off Jalan Telok Sisek

25050 Kuantan, Pahang Darul Makmur.

Tel: 09-5177225, 5177031 & 5177025

Fax: 09-5139434

# CONTENT

---

## CHAPTER 1: INTRODUCTION TO STATISTICS

|       |   |    |
|-------|---|----|
| 1.1   | Statistical Terminologies                           | 2  |
| 1.1.1 | What is Statistics?                                 | 2  |
| 1.1.2 | Why We Need Statistics?                             | 2  |
| 1.1.3 | Population and Sample                               | 3  |
| 1.1.4 | Descriptive and Inferential Statistics              | 6  |
| 1.1.5 | Role of Computer in Statistics                      | 7  |
| 1.2   | Statistical Problem-Solving Methodology             | 9  |
| 1.2.1 | Identifying the Problem or Opportunity              | 9  |
| 1.2.2 | Deciding on the Method of Data Collection           | 10 |
| 1.2.3 | Collecting the Data (Sampling Techniques)           | 12 |
| 1.2.4 | Classifying and Summarising the Data                | 18 |
| 1.2.5 | Presenting and Analysing the Data                   | 27 |
| 1.2.6 | Making the Decision and Conclusion                  | 27 |
| 1.3   | Reviews on Descriptive Statistics                   | 28 |
| 1.3.1 | Measures of Central Tendency                        | 29 |
| 1.3.2 | Measures of Variation (Dispersion)                  | 32 |
| 1.3.3 | Measures of Position                                | 36 |
| 1.3.4 | Descriptive Statistics using <i>Microsoft Excel</i> | 39 |
| 1.4   | Exploratory Data Analysis                           | 42 |
| 1.4.1 | Stem and Leaf Plot                                  | 43 |
| 1.4.2 | Outliers  | 46 |
| 1.4.3 | Boxplot   | 47 |
| 1.5   | Normal Probability Plot                             | 52 |
|       | Summary   | 55 |
|       | Mind Expanding Exercises                            | 57 |
|       | Answers to Exercises                                | 63 |
|       | Tutorial 1  | 65 |
|       | Answers to Tutorial 1                               | 77 |

## CHAPTER 2: SAMPLING DISTRIBUTIONS AND CONFIDENCE INTERVAL

|       |  |    |
|-------|--|----|
| 2.1   | Sampling Distribution and Sampling Error | 84 |
| 2.1.1 | Definition                               | 84 |
| 2.1.2 | Sampling Distribution of Means           | 86 |
| 2.1.3 | Sampling Distribution of Proportions     | 89 |
| 2.2   | Estimate, Estimation, Estimator          | 92 |
| 2.2.1 | Definition                               | 92 |
| 2.2.2 | Properties of Good Point Estimator       | 92 |
| 2.2.3 | Confidence Interval                      | 93 |
| 2.3   | Confidence Interval for Population Mean  | 95 |

|       |   |     |
|-------|---|-----|
| 2.4   | Confidence Interval for Difference Between Two Population Means               | 101 |
| 2.4.1 | Independent Samples   | 101 |
| 2.4.2 | Dependent Samples   | 106 |
| 2.5   | Confidence Interval for Population Proportion                                 | 109 |
| 2.6   | Confidence Interval for Difference Between Two Population Proportions         | 112 |
| 2.7   | Confidence Interval for Population Variance and Population Standard Deviation | 114 |
| 2.8   | Confidence Interval for Ratio of Two Population Variances                     | 116 |
|       | Summary   | 121 |
|       | Mind Expanding Exercises  | 126 |
|       | Answers to Exercises  | 128 |
|       | Tutorial 2  | 129 |
|       | Answers to Tutorial 2   | 140 |

### CHAPTER 3: HYPOTHESIS TESTING

|       |  |     |
|-------|--|-----|
| 3.1   | Introduction to Hypothesis Testing                                   | 144 |
| 3.1.1 | Terms and Definitions  | 145 |
| 3.1.2 | Procedure in Hypothesis Testing                                      | 152 |
| 3.2   | Hypothesis Testing for a Population Mean                             | 153 |
| 3.3   | Hypothesis Testing for the Difference Between Two Population Means   | 157 |
| 3.3.1 | Independent Samples  | 157 |
| 3.3.2 | Dependent Samples  | 163 |
| 3.4   | Hypothesis Testing for Population Proportion                         | 166 |
| 3.5   | Hypothesis Testing for Difference Between Two Population Proportions | 168 |
| 3.6   | Hypothesis Testing for Population Variance                           | 171 |
| 3.7   | Hypothesis Testing for the Ratio of Two Population Variances         | 173 |
| 3.8   | Hypothesis Testing Using Confidence Interval Approach                | 177 |
| 3.9   | <i>P</i> -Value in Hypothesis Tests                                  | 179 |
| 3.9.1 | Finding the <i>P</i> -Value  | 180 |
| 3.9.2 | Procedure of Hypothesis Testing Using <i>P</i> -Value                | 181 |
| 3.9.3 | <i>P</i> -Value Using <i>Microsoft Excel</i>                         | 183 |
|       | Summary  | 192 |
|       | Mind Expanding Exercises   | 199 |
|       | Answers to Exercises   | 200 |
|       | Tutorial 3   | 202 |
|       | Answers to Tutorial 3  | 215 |

### CHAPTER 4: ANALYSIS OF VARIANCE

|     |                                      |     |
|-----|--------------------------------------|-----|
| 4.1 | Introduction to Analysis of Variance | 220 |
| 4.2 | One-Way ANOVA                        | 222 |
| 4.3 | Two-Way ANOVA                        | 230 |
| 4.4 | ANOVA using <i>Microsoft Excel</i>   | 240 |
|     | Summary                              | 246 |
|     | Mind Expanding Exercises             | 248 |
|     | Answers to Exercises                 | 250 |
|     | Tutorial 4                           | 251 |
|     | Answers to Tutorial 4                | 264 |

## CHAPTER 5: LINEAR REGRESSION AND CORRELATION

|       |  |     |
|-------|--|-----|
| 5.1   | Correlation  | 268 |
| 5.1.1 | Scatter Diagram  | 269 |
| 5.1.2 | Correlation Coefficient  | 270 |
| 5.1.2 | Coefficient of Determination                                     | 274 |
| 5.2   | Simple Linear Regression Model                                   | 275 |
| 5.2.1 | Least-squares Estimation of the Model Parameters                 | 276 |
| 5.3   | Hypothesis Testing for Simple Linear Regression Model            | 281 |
| 5.3.1 | Hypothesis Testing for Intercept                                 | 281 |
| 5.3.2 | Hypothesis Testing for Slope                                     | 283 |
| 5.3.3 | Hypothesis Testing for Slope using ANOVA                         | 285 |
| 5.4   | Regression Analysis Using <i>Microsoft Excel</i>                 | 287 |
| 5.5   | Multiple Linear Regression Analysis                              | 291 |
| 5.5.1 | Multiple Linear Regression Model                                 | 291 |
| 5.5.2 | Multiple Linear Regression Analysis Using <i>Microsoft Excel</i> | 293 |
| 5.6   | Model Selection  | 299 |
|       | Summary  | 306 |
|       | Mind Expanding Exercises   | 308 |
|       | Answers to Exercises   | 309 |
|       | Tutorial 5   | 312 |
|       | Answers to Tutorial 5  | 333 |

## CHAPTER 6: GOODNESS OF FIT AND CONTINGENCY TABLE

|       |   |     |
|-------|---|-----|
| 6.1   | Goodness of Fit Test  | 342 |
| 6.1.1 | Goodness of Fit Test for Categorical Data                             | 342 |
| 6.1.2 | Fitting of the Distribution   | 347 |
| 6.2   | Contingency Table   | 354 |
| 6.2.1 | Testing for Independence Between Two Variables                        | 356 |
| 6.2.2 | Test of Homogeneity Proportions                                       | 359 |
| 6.3   | Goodness of Fit Test & Contingency Table Using <i>Microsoft Excel</i> | 362 |
|       | Summary   | 365 |
|       | Mind Expanding Exercises  | 366 |
|       | Answers to Exercises  | 367 |
|       | Tutorial 6  | 368 |
|       | Answers to Tutorial 6   | 381 |

## APPENDIX A: FREQUENCY DISTRIBUTION, HISTOGRAM & BOXPLOT USING MICROSOFT EXCEL

|     |  |     |
|-----|--|-----|
| A.1 | Frequency Distribution                 | 384 |
| A.2 | Histogram Using <i>Microsoft Excel</i> | 389 |
| A.3 | Boxplot Using <i>Microsoft Excel</i>   | 391 |

## **APPENDIX B: COMMON PROBABILITY DISTRIBUTION**

|      |  |     |
|------|--|-----|
| B.1  | Introduction to Statistical Probability Distribution | 398 |
| B.2  | Binomial Distribution                                | 399 |
| B.3  | Poisson Distribution                                 | 403 |
| B.4  | Poisson Approximation for Binomial Distribution      | 406 |
| B.5  | Normal Distribution                                  | 407 |
| B.6  | Central Limit Theorem                                | 415 |
| B.7  | Normal Approximation for Binomial Distribution       | 417 |
| B.8  | Normal Approximation for Poisson Distribution        | 418 |
| B.9  | The $t$ -Distribution                                | 419 |
| B.10 | The Chi-Square Distribution                          | 420 |
| B.11 | The $F$ -Distribution                                | 421 |
| B.12 | Uniform Distribution                                 | 422 |

|                   |     |
|-------------------|-----|
| <b>REFERENCES</b> | 423 |
|-------------------|-----|