***MATHEMATICS SECTION CURRICULUM REVIEW ALLOCATION 10TH JULY, 2024***

***[MATH 142, 241,242,244, 245, 261, 263, 341]***

***JULY, 2024***

***````````````````````````````````````````````````````````````````````````````````````````````````````````````````````````````````````````````***

## MATH 142: EXPLORATORY DATA ANALYSIS (L/P 45/0; CF 3.0)

Course Purpose

The purpose of this course is to enable leaners to apply statistical data visualization and summary statistics to make informed decisions based on data and use appropriate statistical packages.

Expected Learning Outcomes

By the end of the course, the learner should be able to:

1. Discuss fundamental statistical concepts used in Exploratory Data Analysis, such as measures of central tendency, variability, and probability distributions.
2. Promote the importance of data literacy and statistical reasoning in decision-making processes.
3. Explore different methods of data collection, including surveys, experiments, and observational studies.
4. Recognize different data types (e.g., categorical, numerical) and their implications for analysis.
5. Create and interpret graphical displays (e.g., box plots, histograms, scatter plots) to reveal patterns and outliers.
6. Construct frequency distribution tables and summary statistics (e.g., mean, median, variance) using statistical software tools (R, GensStat, Excel).
7. Analyze data, draw meaningful conclusions, and consider practical implications.
8. Understand measures of association for continuous, ordinal, and nominal variables (e.g., correlation, contingency tables).
9. Recognize the role of exploratory data analysis in scientific research, business, and everyday decision-making.

Course Content

Some basic concepts. Populations and samples, randomness and independence**.** Data sources, Data types, methods of data collection, numerical summaries of data, data exploratory techniques: data displays; box plot, stem and leaf, histogram, charts, diagrams. Use of R/GensStat/Excel software to generate; Frequency distribution tables, graphical displays, scatter plots, frequency graphs, Summary statistics, measures of location and dispersion, skewness and Kurtosis. Categorical data analysis; Measures of association for; continuous variables, ordinal variables and nominal variables. Chi-square, loglinear analysis. Use of statistical package.

Teaching and Learning Methods

Lectures, Tutorials, Blended Learning (Online Educational Materials with Face-To-Face Instruction), Computer Assisted Instruction), Question/ Answer Approach, Project-Based Approaches, Group Discussions & Assignments, Presentations, Cooperative Learning, Case Studies and Experiential Learning Where Applicable.

Instructional Materials and Equipment

Overhead Projector, Power Point, Hand-Outs, Charts, Felt Pens, Computer Laboratories, Computers, Learning Resource Centre, and Core Reading Resources.

**Course Assessment**

CATs and Assignments 40%, Final examinations 60%, Total marks 100%

**References**

Aarya, D. D., Rajoria, Y. K., & Sabharwal, M. A. (2023). *Probability And Statistics*. Academic Guru Publishing House.

Andrews, M. Data Visualization. <https://www.mjandrews.org/book/ddsr/assets/chapter_04_visualization.pdf>

Collett, P. (2021). Graphical Data Analysis. *Communities of teachers & learners*, 1-11. <https://jalt-publications.org/sites/default/files/pdf-article/jalt2020-pcp-001_0.pdf>

Larson-Hall, J., & Mizumoto, A. (2019). Using statistical analysis software (R, SPSS). In *The Routledge Handbook of Research Methods in Applied Linguistics* (pp. 385-397). Routledge. eBook ISBN 9780367824471

Sandfeld, S. (2023). Exploratory Data Analysis. In *Materials Data Science: Introduction to Data Mining, Machine Learning, and Data-Driven Predictions for Materials Science and Engineering* (pp. 179-206). Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-031-46565-9_9>

Yu, C. H. A. (2022). *Data mining and exploration: From traditional statistics to modern data science*. CRC Press. <https://doi.org/10.1201/9781003153658>

Mazhar, S. A., Anjum, R., Anwar, A. I., & Khan, A. A. (2021). Methods of data collection: A fundamental tool of research. *Journal of Integrated Community Health (ISSN 2319-9113)*, *10*(1), 6-10.

Rubin, A. (2022). Learning to reason with data: How did we get here and what do we know?. In *Situating Data Science* (pp. 154-164). Routledge. eBook ISBN: 9781003165354

Wolff, A., Wermelinger, M., & Petre, M. (2019). Exploring design principles for data literacy activities to support children’s inquiries from complex data. *International Journal of Human-Computer Studies*, *129*, 41-54. <https://doi.org/10.1016/j.ijhcs.2019.03.006>

## MATH 241: PROBABILITY AND STATISTICS I (L/P 45/0; CF 3.0)

Course Purpose

The purpose of this course unit is to enable learners to apply probability and statistics skills in management and decision making.

Expected Learning Outcomes

By the end of this course unit, the learner should be able to:

1. Discuss the axioms of univariate probability distribution functions and their practical implications.
2. Distinguish between different types of discrete and continuous random variables in data sets and solve related problems
3. Explore the impact of expectations on decision-making and statistical modeling.
4. Construct and interpret confidence intervals for population parameters using sample data.
5. Relate hypothesis testing to real-world scenarios, such as medical trials or quality control.
6. Appreciate the importance of probability distributions (normal, t-, and chi-square) in solving real-world scientific research problems and decision-making.

Course Content

Random variable and probability distributions. Moments and moment generating functions. Statistical Independence of random variables. Confidence intervals for population parameter. Special univariate distributions; discrete and continuous cases. Expectations and variances of special univariate distributions. Approximations of univariate distributions. Applications of univariate distributions.

Teaching and Learning Methods

Lectures, Tutorials, Blended Learning (Online Educational Materials with Face-To-Face Instruction), Computer Assisted Instruction), Question/ Answer Approach, Project-Based Approaches, Group Discussions & Assignments, Presentations, Cooperative Learning, Case Studies and Experiential Learning Where Applicable.

Instructional Materials and Equipment

Overhead Projector, Power Point, Hand-Outs, Charts, Felt Pens, Computer Laboratories, Computers, Learning Resource Centre, and Core Reading Resources.

**Course Assessment**

CATs and Assignments 40%, Final examinations 60%, Total marks 100%

References

Foadoddini, M., & Mofrad, S. A. (2020). Effect of Aloe vera extract on depression in people with prediabetes. *Modern Care Journal*, *17*(2). <https://doi.org/10.5812/modernc.100927>

Grami, A. (2019). *Probability, random variables, statistics, and random processes: Fundamentals & applications*. John Wiley & Sons. ISBN: 9781119300816

Jafari, A., Nejatian, M., Momeniyan, V., Barsalani, F. R., & Tehrani, H. (2021). Mental health literacy and quality of life in Iran: a cross-sectional study. *BMC psychiatry*, *21*, 1-11. <https://doi.org/10.1186/s12888-021-03507-5>

Linde, W. (2024). *Probability Theory: A first course in probability theory and statistics*. Walter de Gruyter GmbH & Co KG. 10.1515/9783111325064-201

Linde, W. (2024). *Probability Theory: A first course in probability theory and statistics*. Walter de Gruyter GmbH & Co KG. e-ISBN (PDF). ISBN: 978-3-11-132506-04

Mishra, P., Singh, U., Pandey, C. M., Mishra, P., & Pandey, G. (2019). Application of student's t-test, analysis of variance, and covariance. *Annals of cardiac anaesthesia*, *22*(4), 407-411. <https://doi.org/10.4103/aca.ACA_94_19>

Roch, S. (2024). *Modern discrete probability: An essential toolkit*. Cambridge University Press. <https://doi.org/10.1017/9781009305129>

## MATH 242: PROBABILITY AND STATISTICS II (45/0 C.F.3.0)

**Course Purpose**

To enable learners to apply the knowledge of getting probability distribution and problem solving and statistical decision making.

**Expected learning outcomes**

By the end of the course, the learner should be able to

1. Understand bivariate probability distributions functions of random variables, bivariate expectations, and characteristic functions and explore their relevance in statistical theory and applications in studying the relationships between random variables.
2. Promote an understanding statistical independence and its implications in bivariate data analysis.
3. Identify the importance of conditional expectations and their role in decision-making and prediction.
4. Explore techniques such as cumulative distribution functions (CDFs) and transformations.
5. Construct probability distributions arising from the normal distribution family, including the Chi-square, t-distribution, and F-distribution and to solve related problems.
6. Appreciate the importance of bivariate normal distribution in solving real world problems.

**Course content**

1. Bivariate Probability Distributions

* Joint probability distributions (J.p.d.f) and Joint cumulative probability distributions(J.c.d.f)
* Marginal distribution functions
* Conditional distribution functions
* Statistical independence

1. Bivariate Expectations

* Bivariate moments, covariance and correlation
* Conditional expectation and variance
* Joint moment generating functions

1. Distribution Functions of Random Variables

* CDF Technique
* Method based on transformation of variables
* Method based on moment generating functions

1. Distributions Arising from Normal Distribution

* The Chi-square distribution
* The t-distribution
* The F- distribution

1. Characteristic Functions

* Discrete and Continuous cases
* Mean and variance applying characteristics function
* Applications in univariate discrete and continuous distributions

1. Bivariate Normal Distribution

* Density function
* Marginal and conditional densities
* Stochastic independence

1. Limiting Distribution Theory

* Limit theory
* Convergence theorem

Teaching and Learning Methods

Lectures, Tutorials, Blended Learning (Online Educational Materials with Face-To-Face Instruction), Computer Assisted Instruction), Question/ Answer Approach, Project-Based Approaches, Group Discussions & Assignments, Presentations, Cooperative Learning, Case Studies and Experiential Learning Where Applicable.

Instructional Materials and Equipment

Overhead Projector, Power Point, Hand-Outs, Charts, Felt Pens, Computer Laboratories, Computers, Learning Resource Centre, and Core Reading Resources.

**Course Assessment**

CATs and Assignments 40%, Final examinations 60%, Total marks 100%

**References**

Abebe, T. H. (2019). The Derivation and choice of appropriate test statistic (z, t, f and chi-square test) in research methodology. *J. Math. Lett*, *5*(3), 33-40. <https://doi.org/10.11648/j.ml.20190503.11>

Agresti, A. (2021). The foundations of statistical science: A history of textbook presentations. *Brazilian Journal of Probability and Statistics*, *35*(4), 657-698. <https://doi.org/10.1214/21-BJPS518>

Astrauskas, A. (2023). Some Bounds for the Expectations of Functions on Order Statistics and Their Applications. *Journal of Theoretical Probability*, *36*(2), 1116-1147. <https://doi.org/10.1007/s10959-022-01179-9>

Casella, G., & Berger, R. (2024). *Statistical inference*. CRC Press. <https://doi.org/10.1201/9781003456285>

Foadoddini, M., & Mofrad, S. A. (2020). Effect of Aloe vera extract on depression in people with prediabetes. *Modern Care Journal*, *17*(2). <https://doi.org/10.5812/modernc.100927>

Hoorijani, F., Manouchehri, B., Faridi, H., & Minaee, S. (2023). The Effect of Music on The Level of Happiness in Primigravida Women. *The Egyptian Journal of Hospital Medicine*, *90*(2), 3298-3305. <https://dx.doi.org/10.21608/ejhm.2023.291346>

Kotz, S., Balakrishnan, N., & Johnson, N. L. (2019). *Continuous multivariate distributions, Volume 1: Models and applications* (Vol. 334). John wiley & sons. ISBN: 0-471-1837-3

Linde, W. (2024). *Probability Theory: A first course in probability theory and statistics*. Walter de Gruyter GmbH & Co KG. e-ISBN (PDF) 978-3-11-132506-4

Mavrakakis, M. C., & Penzer, J. (2021). *Probability and statistical inference: from basic principles to advanced models*. Chapman and Hall/CRC. ISBN: 9781315366630 (ebk)

OBI, J. C. Foundation Probability and Basic Distribution Theory. ISBN 978-1-119-30081-6

Potters, M., & Bouchaud, J. P. (2020). *A first course in random matrix theory: for physicists, engineers and data scientists*. Cambridge University Press. <https://doi.org/10.1017/978110876900>

## MATH 244: INTRODUCTION TO STATISTICAL INFERENCE (L/P 45/0; CF 3.0)

Course Purpose

The purpose of this course is to enable the learner to be able to apply the knowledge of applied statistics in computations and decision making.

Expected Learning Outcomes

By the end of the course, the learner should be able to:

1. Describe the challenges of statistical inference.
2. State and explain key properties of point estimators (such as unbiasedness, efficiency) and interval estimators (confidence intervals).
3. Gain practical skills in obtaining random samples and using random number tables.
4. Apply hypothesis testing techniques to real-world situations, making informed decisions based on sample evidence.
5. Understand how regression models provide insights into relationships between variables and make predictions.
6. Conduct goodness of fit tests (e.g., chi-square tests) and test for independence appropriately, considering specified parameters or unknown parameters.

**Course Content**

Meaning and role of statistics, objectives of statistical research or investigation. Statistical decision problems, basic concepts of inference: sampling, estimation, and hypothesis testing. Role of normal distribution in statistical inference. Random samples, use of random number tables. Inference about population means: point estimates and confidence intervals, simple analysis of frequency data. Simple nonparametric tests: sign test/Wilcoxon Signed-Rank Test for paired data, run test for randomness and trends.

**Teaching and Learning Methods**

Lectures, Tutorials, Blended Learning (Online Educational Materials with Face-To-Face Instruction), Computer Assisted Instruction), Question/ Answer Approach, Project-Based Approaches, Group Discussions & Assignments, Presentations, Cooperative Learning, Case Studies and Experiential Learning Where Applicable.

Instructional Materials and Equipment

Overhead Projector, Power Point, Hand-Outs, Charts, Felt Pens, Computer Laboratories, Computers, Learning Resource Centre, and Core Reading Resources.

**Course Assessment**

CATs and Assignments 40%, Final examinations 60%, Total marks 100%

**References**

Alita, D., Putra, A. D., & Darwis, D. (2021). Analysis of classic assumption test and multiple linear regression coefficient test for employee structural office recommendation. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *15*(3), 295-306. <https://journal.ugm.ac.id/ijccs/article/view/65586/31843>

Bruce, P., Bruce, A., & Gedeck, P. (2020). *Practical statistics for data scientists: 50+ essential concepts using R and Python*. O'Reilly Media. ISBN 9781492072942

Casella, G., & Berger, R. (2024). *Statistical inference*. CRC Press. <https://www.academia.edu/download/63689011/Statistical_Inference_2nd_Edition20200620-89838-dmtd1q.pdf>

Cevallos-Torres, L., Botto-Tobar, M., Cevallos-Torres, L., & Botto-Tobar, M. (2019). Process sampling. *Problem-Based Learning: A Didactic Strategy in the Teaching of System Simulation*, 13-31. <https://doi.org/10.1007/978-3-030-13393-1_2>

Franke, B., Plante, J. F., Roscher, R., Lee, E. S. A., Smyth, C., Hatefi, A., ... & Reid, N. (2016). Statistical inference, learning and models in big data. *International Statistical Review*, *84*(3), 371-389. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/insr.12176>

Makar, K., & Rubin, A. (2018). *Learning about statistical inference* (pp. 261-294). Springer International Publishing. <https://doi.org/10.1007/978-3-319-66195-7_8>

Orcan, F. (2020). Parametric or non-parametric: Skewness to test normality for mean comparison. *International Journal of Assessment Tools in Education*, *7*(2), 255-265. <https://doi.org/10.21449/ijate.656077>

van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E. J. (2020). Bayesian rank-based hypothesis testing for the rank sum test, the signed rank test, and Spearman's ρ. *Journal of Applied Statistics*, *47*(16), 2984-3006. <https://doi.org/10.1080/02664763.2019.1709053>

## MATH 245: INTRODUCTION TO EPIDEMIOLOGICAL METHODS

**Course Purpose**

The purpose of this course is to enable the learner to describe primary applications of epidemiology in public health practice.

**Expected Learning Outcomes**

By the end of the course, the learner should be able to:

1. Explain the fundamental principles and scope of epidemiology, emphasizing its role in public health research and practice.
2. Summarize the historical evolution of epidemiology.
3. Describe a public health problem according to three essential characteristics: person, place, and time.
4. Discuss the key features and uses of descriptive epidemiology.
5. Apply descriptive methods to summarize and interpret disease frequency, morbidity, and mortality patterns.
6. Describe a public health problem using epidemiologic triangle
7. Explore the key features and applications of analytic epidemiology
8. Identify the key steps is solving a public health problem
9. Identify types of data sources, methods of data collection, analysis and interpretation, and their relevance in epidemiological research.
10. Enumerate and explain how epidemiological methods contribute to public health practice, policy, and decision-making.

**Course Content**

Introduction to epidemiologic methods, types of epidemiology, epidemiologic triangle, epidemiologic/the public health cycle, measures of disease frequency/occurrence: Morbidity (incidence and prevalence), mortality, ratios, proportions, rates, Measures of effect: relative risk/ratio and odds ratio; Measures of central tendencies and dispersion: mean, median, mode, ranges, variance, standard deviation. Organizing Epidemiologic Data: tables, graphs: scatter plots, pie chart, and box plot, bar plots. Types of epidemiological studies: descriptive (correlational studies, case reports, case series, cross-sectional surveys, etc.) and analytical studies (case control (Unmatched or matched) and cohort studies, experimental/ clinical trials, case-crossover and case-only designs) {definitions advantages and disadvantages}. Sources of bias in epidemiological studies (confounding, selection bias, information bias). Controlling confounding by stratifying data with application to HIV and AIDS, malaria, TB and other contagious disease. Hypothesis testing: solving a public health problem, data sources, practical and statistical significance, Simple regression models in epidemiological analysis, Confidence interval estimation for means and OR. Developing clinical prediction rules, analyses of community-level associations or interventions, propensity scores. Use of statistical package (MS. Excel, R.).

Teaching and Learning Methods

Lectures, Tutorials, Blended Learning (Online Educational Materials with Face-To-Face Instruction), Computer Assisted Instruction), Question/ Answer Approach, Project-Based Approaches, Group Discussions & Assignments, Presentations, Cooperative Learning, Case Studies and Experiential Learning Where Applicable.

Instructional Materials and Equipment

Overhead Projector, Power Point, Hand-Outs, Charts, Felt Pens, Computer Laboratories, Computers, Learning Resource Centre, and Core Reading Resources.

**Course Assessment**

CATs and Assignments 40%, Final examinations 60%, Total marks 100%

References

Alexander, L. K., Lopes, B., Ricchetti-Masterson, K., & Yeatts, K. B. (2020). Common measures and statistics in epidemiological literature. *ERIC notebook*. <https://sph.unc.edu/wp-content/uploads/sites/112/2015/07/nciph_ERIC3.pdf>

Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American journal of epidemiology*, *163*(12), 1149-1156. <https://doi.org/10.1093/aje/kwj149>

Goldsteen, R. L., Goldsteen, K., & Dwelle, T. (2024). Introduction to public health: promises and practices. <https://doi.org/10.1891/9780826186157>

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, *31*(4), 337-350. <https://doi.org/10.1007/s10654-016-0149-3>

Jager, K. J., Tripepi, G., Chesnaye, N. C., Dekker, F. W., Zoccali, C., & Stel, V. S. (2020). Where to look for the most frequent biases?. *Nephrology*, *25*(6), 435-441. <https://doi.org/10.1111/nep.13706>

Jager, K. J., Zoccali, C., Kramar, R., & Dekker, F. W. (2007). Measuring disease occurrence. *Kidney international*, *72*(4), 412-415. <https://doi.org/10.1038/sj.ki.5002341>

John, T. J., & Kompithra, R. Z. (2023). Eco-epidemiology triad to explain infectious diseases. *Indian Journal of Medical Research*, *158*(2), [https://doi.org/107-112. 10.4103/ijmr.ijmr\_3031\_21](https://doi.org/107-112.%2010.4103/ijmr.ijmr_3031_21)

Lesko, C. R., Fox, M. P., & Edwards, J. K. (2022). A framework for descriptive epidemiology. *American journal of epidemiology*, *191*(12), 2063-2070. <https://doi.org/10.1093/aje/kwac115>

Loke, Y. K., & Mattishent, K. (2020). Propensity score methods in real‐world epidemiology: a practical guide for first‐time users. *Diabetes, Obesity and Metabolism*, *22*, 13-20. <https://doi.org/10.1111/dom.13926>

Magarey, R. D., & Trexler, C. M. (2020). Information: A missing component in understanding and mitigating social epidemics. *Humanities and Social Sciences Communications*, *7*(1), 1-11. <https://doi.org/10.1057/s41599-020-00620-w>

Merrill, R. M. (2024). *Introduction to epidemiology*. Jones & Bartlett Learning. ISBN 9781284280388

Nahar, P., van Marwijk, H., Gibson, L., Musinguzi, G., Anthierens, S., Ford, E., ... & Bastiaens, H. (2020). A protocol paper: community engagement interventions for cardiovascular disease prevention in socially disadvantaged populations in the UK: an implementation research study. *Global Health Research and Policy*, *5*, 1-9. <https://doi.org/10.1186/s41256-020-0131-1>

Omair, A. (2015). Selecting the appropriate study design for your research: Descriptive study designs. *Journal of health specialties*, *3*(3), 153. <https://doi.org/10.4103/1658-600X.159892>

Reilly, M. (2023). *Controlled Epidemiological Studies*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429197574>

Rezigalla, A. A. (2020). Observational study designs: Synopsis for selecting an appropriate study design. *Cureus*, *12*(1). <https://doi.org/10.7759%2Fcureus.6692>

Rossi, R. J. (2022). *Applied biostatistics for the health sciences*. John Wiley & Sons. ISBN 9781119722670 (PDF) 9781119722700 (EPUB)

Shiba, K., & Kawahara, T. (2021). Using propensity scores for causal inference: pitfalls and tips. *Journal of epidemiology*, *31*(8), 457-463. <https://doi.org/10.2188/jea.JE20210145>

Sullivan, L. M. (2022). *Essentials of biostatistics for public health*. Jones & Bartlett Learning. ISBN 9781284231977

Tripepi, G., Jager, K. J., Dekker, F. W., Wanner, C., & Zoccali, C. (2007). Measures of effect: relative risks, odds ratios, risk difference, and ‘number needed to treat’. *Kidney international*, *72*(7), 789-791. <https://doi.org/10.1038/sj.ki.5002432>

Yan, A. F., Chen, Z., Wang, Y., Campbell, J. A., Xue, Q. L., Williams, M. Y., ... & Egede, L. E. (2022). Effectiveness of social needs screening and interventions in clinical settings on utilization, cost, and clinical outcomes: a systematic review. *Health equity*, *6*(1), 454-475. <https://doi.org/10.1089/heq.2022.0010>

Zhao, Q. Y., Luo, J. C., Su, Y., Zhang, Y. J., Tu, G. W., & Luo, Z. (2021). Propensity score matching with R: conventional methods and new features. *Annals of translational medicine*, *9*(9). <https://doi.org/10.21037%2Fatm-20-3998>

## MATH 261: INTRODUCTION TO COMPUTER INTERACTIVE STATISTICS (30/30 C.F.3.0)

**Course Purpose**

To present statistical research tools from survey to analysis using various statistical software to make valid, viable and objective conclusions.

**Expected Learning Outcomes**

By the end of the course, the learner should be able to:

1. Identify types of variables and data
2. Perform data entry, import, pre-process, and analyze data using a statistical software.
3. Comprehend and perform data preprocessing techniques
4. Present data using graphical and tabular representation
5. Identify the presence of errors or misleading quantitative information and remedies
6. Summarize numeric and categorical data
7. Conduct hypothesis testing for means, proportions, and variance
8. Apply computational and analysis skills/ techniques to solve data-related problems and make evidenced based decisions.
9. Plan and conduct a survey.
10. Write statistical reports

**Course Content**

Basic concepts of modern statistics: types of Variables and Measurement Scales, Population, Sample and sampling techniques. Data collection, Big data quality management framework, Darta curation, cleaning/ cleansing. R: windows / user interface, Data entry, importation, structures, types, Objects, session, functions, packages, and libraries, and arguments. Descriptive Statistics: Data visualization, recognition of accuracy or misleading quantitative information using statistical graphics, Data distributions (normal and bimodal), Data preprocessing/ preparation steps (filtering, removal of sparse features, discretization, continuization, normalization, randomization, etc.); Proportions (One-way/Two-way tables), Measures Central Tendency and Dispersion (Ungrouped and Grouped). Inferential Statistics: Point and confidence interval estimation; estimating a mean, estimating a proportion, estimating a variance, estimating a variance ratio, sampling techniques and variability, bootstrap estimation. Hypothesis testing: Types of statistical errors, level of significance, power of test, hypothesis testing for means and proportions. Understanding of statistical reports: Exporting r-output: Figures to PNG, JPEG, PDF etc.; formatted tables to excel, word, and knitting as HTML; Writing summary statistical reports. Introduction to r-markdown

Teaching and Learning Methods

Lectures, Tutorials, Blended Learning (Online Educational Materials with Face-To-Face Instruction), Computer Assisted Instruction), Question/ Answer Approach, Project-Based Approaches, Group Discussions & Assignments, Presentations, Cooperative Learning, Case Studies and Experiential Learning Where Applicable.

Instructional Materials and Equipment

Overhead Projector, Power Point, Hand-Outs, Charts, Felt Pens, Computer Laboratories, Computers, Learning Resource Centre, and Core Reading Resources.

**Course Assessment**

CATs and Assignments 40%, Final examinations 60%, Total marks 100%

**References**

Brown, D. S. (2021). *Statistics and Data Visualization Using R: The Art and Practice of Data Analysis*. SAGE Publications.

Dunleavy, D. J. (2020). *Appraising contemporary social work research: Meta-research on statistical reporting, statistical power, and evidential value*. The Florida State University.

Hahs-Vaughn, D. L., & Lomax, R. G. (2020). *An introduction to statistical concepts*. Routledge. <https://doi.org/10.4324/9781315624358>

Huang, F. (2022). Data processing. In *Encyclopedia of big data* (pp. 312-316). Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-32010-6_314>

Jones, E., Harden, S., & Crawley, M. J. (2022). *The R book*. John Wiley & Sons. ebook ISBN 9781119634324

Jung, K., Lee, J., Gupta, V., & Cho, G. (2019). Comparison of bootstrap confidence interval methods for GSCA using a Monte Carlo simulation. *Frontiers in psychology*, *10*, 2215. <https://doi.org/10.3389/fpsyg.2019.02215>

Kang, H. (2021). Sample size determination and power analysis using the G\* Power software. *Journal of educational evaluation for health professions*, *18*. <https://doi.org/10.3352/jeehp.2021.18.17>

Kattan, M. W., & Vickers, A. J. (2020). Statistical analysis and reporting guidelines for CHEST. *Chest*, *158*(1), S3-S11. <https://doi.org/10.1016/j.chest.2019.10.064>

Khan, A. (2022). A computational framework for Data-Driven Decision Support and Knowledge Augmented Process Analytics. <https://ro.uow.edu.au/theses1/1735>

Kopyt, M. (2020). Basic operations in the R software. In *Applied Spatial Statistics and Econometrics* (pp. 1-35). Routledge. eBook ISBN9781003033219

Lane D.M., (2013). Introduction to Statistics: An Interactive e-book, 3rd Edition, David M. Lane

Serdar, C. C., Cihan, M., Yücel, D., & Serdar, M. A. (2021). Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia medica*, *31*(1), 27-53. <https://doi.org/10.11613/BM.2021.010502>

Taleb, I., Serhani, M. A., Bouhaddioui, C., & Dssouli, R. (2021). Big data quality framework: a holistic approach to continuous quality management. *Journal of Big Data*, *8*(1), 76. <https://doi.org/10.1186/s40537-021-00468-0>

Xie, Y., Dervieux, C., & Riederer, E. (2020). *R markdown cookbook*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781003097471>

## MATH 263: STATISTICAL PROGRAMMING, I (30/30 C.F.3.0)

**Course Purpose**

The purpose of this course is to introduce students to object-oriented concepts in computer programming using Java programming language.

**Expected Learning Outcomes**

By the end of the course, the learner should be able to:

1. Discuss the principles of statistical programming.
2. Write programs using different programming languages commonly used in statistical analysis (e.g., R, Python).
3. Design and implement classes of objects given desired attributes for efficient code organization and reuse in statistical projects.
4. Applying Monte Carlo methods for statistical modeling and uncertainty assessment.
5. Explore statistical power and its relevance in hypothesis testing.
6. Estimate confidence intervals for population parameters using a statistical software
7. Develop application software that uses a simple database for storing and retrieving statistical data.

**Course Content**

Basic Knowledge of High-Level Programming Languages: Introduction to programming languages such as C++, S-plus, and R, understanding syntax, data types, and control structures; Computer Arithmetic: Numeric representation and precision in programming, handling arithmetic operations and potential pitfalls; Algorithm for Mean and Standard Deviation: Implementing algorithms to calculate mean and standard deviation, efficiency and accuracy; Error Analysis: Identifying common errors in programming, Strategies for debugging and error handling; Pseudo-Random Number Generators: Generating random numbers for simulations, Evaluating the quality of random number sequences; Generation of Random Variates: Understanding discrete and continuous probability distributions, Simulating random variates from these distributions; Monte Carlo Simulation; Power of Tests: Confidence Intervals: Estimating confidence intervals for population parameters.

Teaching and Learning Methods

Lectures, Tutorials, Blended Learning (Online Educational Materials with Face-To-Face Instruction), Computer Assisted Instruction), Question/ Answer Approach, Project-Based Approaches, Group Discussions & Assignments, Presentations, Cooperative Learning, Case Studies and Experiential Learning Where Applicable.

Instructional Materials and Equipment

Overhead Projector, Power Point, Hand-Outs, Charts, Felt Pens, Computer Laboratories, Computers, Learning Resource Centre, and Core Reading Resources.

**Course Assessment**

CATs and Assignments 40%, Final examinations 60%, Total marks 100%

**References**

Abt, G., Boreham, C., Davison, G., Jackson, R., Nevill, A., Wallace, E., & Williams, M. (2020). Power, precision, and sample size estimation in sport and exercise science research. *Journal of Sports Sciences*, *38*(17), 1933-1935. <https://doi.org/10.1080/02640414.2020.1776002>

Braun, W. J., & Murdoch, D. J. (2021). *A first course in statistical programming with R*. Cambridge University Press. ISBN: 978-1-108-99514-6

Carrasco, J., García, S., Rueda, M. M., Das, S., & Herrera, F. (2020). Recent trends in the use of statistical tests for comparing swarm and evolutionary computing algorithms: Practical guidelines and a critical review. *Swarm and Evolutionary Computation*, *54*, 100665. <https://doi.org/10.1016/j.swevo.2020.100665>

Chowdhary, S., Lim, J. P., & Nagarakatte, S. (2020, June). Debugging and detecting numerical errors in computation with posits. In *Proceedings of the 41st ACM SIGPLAN Conference on Programming Language Design and Implementation* (pp. 731-746). <https://doi.org/10.1145/3385412.3386004>

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European journal of epidemiology*, *31*(4), 337-350. <https://doi.org/10.1007/s10654-016-0149-3>

Isukapalli, S. S., & Roy, A. (2007). General Principles of Programming: Computer and Statistical. *Pharmacometrics: The Science of Quantitative Pharmacology*, 23-52. <https://doi.org/10.1002/9780470087978.ch2>

Jung, K., Lee, J., Gupta, V., & Cho, G. (2019). Comparison of bootstrap confidence interval methods for GSCA using a Monte Carlo simulation. *Frontiers in psychology*, *10*, 2215. <https://doi.org/10.3389/fpsyg.2019.02215>

Kang, H. (2021). Sample size determination and power analysis using the G\* Power software. *Journal of educational evaluation for health professions*, *18*. <https://doi.org/10.3352/jeehp.2021.18.17>

McNamara, A. (2019). Key attributes of a modern statistical computing tool. *The American Statistician*. <https://doi.org/10.1080/00031305.2018.1482784>

Morris, Tim P., Ian R. White, and Michael J. Crowther. "Using simulation studies to evaluate statistical methods." *Statistics in medicine* 38, no. 11 (2019): 2074-2102. <https://doi.org/10.1002/sim.8086>

Rizzo, M. L. (2019). *Statistical computing with R*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429192760>

Scott, S. L., Blocker, A. W., Bonassi, F. V., Chipman, H. A., George, E. I., & McCulloch, R. E. (2022). Bayes and big data: The consensus Monte Carlo algorithm. In *Big Data and Information Theory* (pp. 8-18). Routledge. eBook ISBN 9781003289173

Serdar, C. C., Cihan, M., Yücel, D., & Serdar, M. A. (2021). Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia medica*, *31*(1), 27-53. <https://doi.org/10.11613/BM.2021.010502>

Tollefson, M. (2019). *R Quick Syntax Reference: A Pocket Guide to the Language, APIs and Library*. Apress. ISBN-13 (pbk): 978-1-4842-4404-3 <https://doi.org/10.1007/978-1-4842-4405-0>

Vana-Gür, L. 105.731 AKSTA Statistical Computing. <https://vowi.fsinf.at/images/2/23/TU_Wien-AKSTA_Statistical_Computing_VU_%28Vana_G%C3%BCr%29_-_StatComp_Script2022S.pdf>

## MATH 341: SAMPLING METHODS I (L/P 45/0; CF 3.0)

Course Purpose

The purpose of this course is to help learners design effective surveys, estimate population parameters, and make informed decisions based on sampled data..

Expected Learning Outcomes

By the end of this course, the learner should be able to:

1. Discuss the importance of sampling methods and how sampling impacts data quality and statistical inference.
2. Identify sources of errors in surveys (sampling bias, non-response, measurement error, etc.) and understand their impact on data quality and survey results.
3. Explore adequate sample size determination techniques to ensure reliable statistical inference.
4. Construct sampling designs using methods like simple random sampling, stratified random sampling, and systematic selection.
5. Value the role of statistical approximation and accuracy assessment in solving mathematical and scientific problems.

Course Content

General principles of a sample survey, Sources of errors in survey, Confidence interval, Simple size determination, Sampling techniques; Simple random sampling; with replacement and without replacement. stratified random sampling; proportional and optimal allocations. Systematic sampling. Sampling theory. Sample selection with probability proportional to size. Modern sampling techniques and applications.

Teaching and Learning Methods

Lectures, Tutorials, Blended Learning (Online Educational Materials with Face-To-Face Instruction), Computer Assisted Instruction), Question/ Answer Approach, Project-Based Approaches, Group Discussions & Assignments, Presentations, Cooperative Learning, Case Studies and Experiential Learning Where Applicable.

Instructional Materials and Equipment

Overhead Projector, Power Point, Hand-Outs, Charts, Felt Pens, Computer Laboratories, Computers, Learning Resource Centre, and Core Reading Resources.

**Course Assessment**

CATs and Assignments 40%, Final examinations 60%, Total marks 100%

**References**

Berndt, A. E. (2020). Sampling methods. *Journal of Human Lactation*, *36*(2), 224-226. <https://doi.org/10.1177/0890334420906850>

Lehdonvirta, V., Oksanen, A., Räsänen, P., & Blank, G. (2021). Social media, web, and panel surveys: using non‐probability samples in social and policy research. *Policy & internet*, *13*(1), 134-155. <https://doi.org/10.1002/poi3.238>

Lohr, S. L. (2021). *Sampling: design and analysis*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429298899>

Sharma, S. K., Mudgal, S. K., Thakur, K., & Gaur, R. (2020). How to calculate sample size for observational and experimental nursing research studies. *National Journal of Physiology, Pharmacy and Pharmacology*, *10*(1), 1-8. <https://doi.org/10.5455/njppp.2020.10.0930717102019>

Turner, D. P. (2020). Sampling Methods in Research Design. *Headache: The Journal of Head & Face Pain*, *60*(1). <https://doi.org/10.1111/head.13707>

Wang, X., & Ji, X. (2020). Sample size estimation in clinical research: from randomized controlled trials to observational studies. *Chest*, *158*(1), S12-S20. <https://doi.org/10.1016/j.chest.2020.03.010>

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social science & medicine*, *292*, 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>