

Department of Statistics

Central University of Jharkhand, Ranchi

Programme Name: Ph.D. in Statistics

A Ph.D. in Statistics is aimed at advancing the knowledge in the field of statistical theory and its applications, preparing graduates for careers in academia, research and industries.

Vision

- To make the Department a multidisciplinary teaching and research hub in the state of Jharkhand.
- To foster statistical excellence coupled with analytical acumen among students coming forward to pursue Post Graduation in Statistics with a view to strengthening statistical-base in Indian line with Bhartiya culture.
- To cultivate a community of statisticians who can contribute significantly to the advancement of knowledge and betterment of the society.

Mission

- To serve as a catalyst for positive transformation within the realm of statistical education and research.
- To foster a vibrant knowledge community characterized by multi-disciplinary learning via cutting edge and impactful research.
- To contribute meaningfully to the holistic development and self-sufficiency of future generations.
- To advance knowledge, address societal challenges, and empower individuals for an ever-evolving world.

Programme Objectives:

1. **Develop Advanced Knowledge:** To equip students with deep theoretical understanding and practical skills in statistical methods, mathematical foundations, and data analysis techniques.
2. **Promote Research Skills:** To prepare students to conduct original, high-quality research that contributes to the advancement of the field. This includes designing experiments, analyzing complex data, and interpreting results.

3. **Encourage Methodological Innovation:** To foster the development of new statistical methodologies and techniques, addressing contemporary challenges and applying them to various domains such as biostatistics, social sciences, and data science.
4. **Enhance Problem-Solving Abilities:** To train students to approach and solve complex problems in diverse areas using statistical reasoning and methodologies.
5. **Prepare for Leadership:** To develop leadership skills in statistical consulting, teaching, and project management to prepare graduates for influential roles in academia, industry, or government.

Programme Outcomes

1. **Original Research Contributions:** Graduates will produce significant and original research after successful completion of Ph.D. program.
2. **Advanced Analytical Skills:** Proficiency in applying advanced statistical techniques and tools to real-world data, including designing experiments, performing data analysis, and interpreting complex results.
3. **Teaching and Communication Skills:** Ability to effectively teach statistical concepts and methods, and communicate complex ideas clearly to both specialized and general audiences.
4. **Professional Expertise:** Expertise in specialized areas of statistics, such as computational statistics, statistical genetics, or Bayesian methods, depending on their research focus and interests.
5. **Career Readiness:** Preparation for various career paths, including academic positions (professor, researcher), industry roles (data scientist, statistical consultant), and positions in government or non-profit organizations.

Overall, a Ph.D. in Statistics aims to develop leading experts who can push the boundaries of statistical knowledge and apply their expertise to solve complex problems in a variety of fields.

Curriculum Structure

Compulsory Course

Sl. No.	Course No.	Course Title	L-T-P	Credits
1.		Research Methodology	3-0-1	4
2.		Research and Publication Ethics	1-0-1	2

Based on the suggestions of Ph.D. Supervisor, student has to take at least two papers from the following list of Papers.

Sl. No.	Course No.	Course Title	L-T-P	Credits
1.		Advanced Sample Survey	4-0-0	4
2.		Survival Analysis	4-0-0	4
3.		Bayesian Inference	4-0-0	4

4.		Reliability Theory	4-0-0	4
5.		MOOCs		

L - Lecture hrs/week; T - Tutorial hrs/week; P-Practical/Lab hrs/week

DETAILED SYLLABUS

1. Research Methodology

Credit: 4

Unit	Content	Hours
I	Introduction to Research methodology: Introduction to Research: Meaning of research, types of research, role of research, process of research, Questionnaire: Construction and administration. Different Measurement scales like Guttman Scales, Likert Scale etc, Research problem and review of literature, various sampling techniques, Determination of sample size for different sampling schemes, research designs for comparative treatments.	10
II	Examination of data: dealing with missing data and outliers.	10
III	Testing the assumption of: Normality, homoscedasticity, linearity, Absence of correlated factors, Data transformation.	10
IV	Use of Statistical Softwares like R, R Studio, and SPSS: random number generation and simulation using R. Latex and Beamer	10

References:

1. Kothari, C.R. (2013). *Research methodology: Methods and techniques*. (2nd Ed.). New Delhi: New Age International Publishers.
2. Oliver, P. (2004). *Writing Your Thesis*, New Delhi: Vistaar Publications, 2004.
3. Day, R.A., (1992). *How To Write and Publish a Scientific Paper*, Cambridge University Press, London.

2. Research and Publications Ethics

Credit: 2

Unit	Content	Hours
I	Theory: Module One: Philosophy and Ethics <ol style="list-style-type: none"> 1) Introduction to Philosophy: definition, nature and scope, concept, branches. 2) Ethics: definition, moral philosophy, nature of moral judgments and reactions. 	10

II	<p>Module Two: Scientific Conduct</p> <ol style="list-style-type: none"> 1) Ethics with respect to science and research 2) Intellectual honest and research integrity 3) Scientific misconducts: Falsification, Fabrication, and Plagiarism (FFP) 4) Redundant publications: Duplicate and overlapping publications, salami slicing 5) Selective reporting and misrepresentation of data 	10
III	<p>Module Three: Publication Ethics</p> <ol style="list-style-type: none"> 1) Publication ethics: definition, introduction and importance 2) Best practices/standards setting initiatives and guidelines: COPE, WAME, etc. 3) Conflict of interest 4) Publication misconduct: definition, concept, problems that lead to unethical behaviour and vice versa, types. 5) Violation of publication ethics, authorship and contributorship 6) Identification of publication misconduct, complaints and appeals 7) Predatory publishers and journals 	10
IV	<p>PRACTICE</p> <p>Module Four: Open Access Publication</p> <ol style="list-style-type: none"> 1) Open access publications and initiatives 2) SHERPA/RoMEO online resource to check publisher copyright & self-archiving policies 3) Software tool to identify predatory developed by SPPU 4) Journal finder/Journal suggestion tools viz., JANE, Elsevier Journal Finder, Springer Journal Suggester, etc. <p>Module Five: Publication Misconduct</p> <p>A. Group Discussions</p> <ol style="list-style-type: none"> 1) Subject specific ethical issues, FFP, authorship 2) Conflict of interest 3) Complaints and appeals: examples and fraud from India and abroad <p>B. Software Tools</p> <ol style="list-style-type: none"> 1) Use of plagiarism software like Turnitin, Urkund and other open source software tools <p>Module Six: Databases and Research Metrics</p> <p>A. Databases</p> <ol style="list-style-type: none"> 1) Indexing databases 2) Citation databases: Web of Science, Scopus etc. <p>B. Research Metrics</p> <ol style="list-style-type: none"> 2) Impact Factor of Journal as per Journal Citation Report, SNIP, SJR, IPP, Cite Score <p>Metrics: h-index, g-index, i10index, altmetrics</p>	10

Suggested Readings:

1. A. Bird, *Philosophy of Science* (Routledge, 2006);
2. Alasdair MacIntyre, *A Short History of Ethics* (London, 1967);
3. P. Chaddah, *Ethics in Competitive Research: Do not get Scooped, do not get Plagiarized*, 2018, ISBN: 978-9387480865;
4. National Academy of Sciences, National Academy of Engineering and Institute of Medicine, *On Being a Scientist: A Guide to Responsible Conduct in Research*, Third Edition, National Academies Press, 2009;
5. D. B. Resnik, "What is Ethics in Research & Why is it Important, *National Institute of Environmental Health Sciences*, 2011, pp.1-10. Retrieved from <https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm>;
6. J. Beall, "Predatory Publishers are Corrupting Open Access, *Nature*, Vol.489, No.7415, 2012, pp.179—179. <https://doi.org/10.1038/489179a>;
7. Indian National Science Academy (INSA), *Ethics in Science Education, Research and Governance*, 2019, ISBN: 978-81-939482-1-7. <https://www.insaindia.res.in/pdf/Ethics.Book;pdf>;

1. Advance Sample Survey

Credit: 4

Unit	Content	Hours
I	Review of SRS, Stratified random sampling, systematic sampling, ratio, product and regression methods of estimation, cluster sampling, two stage sampling, repetitive sampling, ranked set sampling, sources of errors in sampling, Non response, Hansen & Hurwitz approach.	10
II	Randomized response techniques, Warner's model, unrelated question model, use of binary responses, small area estimation.	10
III	Survey error evaluation methods, variance estimation in the presence of measurement error for binary and polytomous response variables, response probability model, Hui-Walter model,	10
IV	Latent Class models for evaluating classification errors with model estimation and evaluation.	10

References:

1. Chaudhury, A. & Mukherjee (1988). *Randomized response: Theory and Techniques*, New York: Marcel Dekkar Inc.
2. Cochran, W. G. (1997). *Sampling Techniques*, Wiley.
3. Des Raj and Chandhok (1998). *Sampling Theory*, Narosa.

4. Singh, D.&Chaudhary, F.S. (1986). Theory and Analysis of sample survey Designs, John Wiley & Sons.
5. Mukhopadhyay, P. (2008). Theory and Methods of Survey Sampling. PHI Learning Pvt Ltd. New Delhi, India.
6. Paul P. Biemer (2011). Latent Class Analysis of Survey error, John Wiley & Sons
7. Alan Agresti (2007). An introduction to Categorical Data Analysis, second edition, A John Wiley & Sons.

2. Survival Analysis

Credits: 4

Unit	Content	Hours
I	Censoring: Introduction, Concepts of Type-I (time), Type-II (order) and random censoring, likelihood in these cases. Life time distributions, exponential Gamma, Weibull, lognormal, Pareto, linear failure rate. Inference for exponential, Gamma, Weibull distributions under censoring.	10
II	Failure rate, mean residual life and their elementary properties. Ageing classes and their properties, bathtub failure rate.	10
III	Estimation of survival function – Actuarial estimator, Kaplan Meier estimator, Tests of exponentiality against non-parametric classes: Total time on Test, Deshpande Test.	10
IV	Two sample problem: Gehan test, Log rank test. Mantel-Haenszel test, Cox's proportional hazards model, competing risks model.	10

References:

1. Cox, D.R. and Oakes, D. (1984). Analysis of Survival Data, Chapter 1,2,3,4, Taylor and Francis.
2. Crowder, M.J. (2001). Classical Competing Risks, Chapman & Hall, CRC London.
3. Miller, R.G. (1998). Survival Analysis, Second Edition, Wiley Interscience.
4. Deshpande, J.V. and Purohit, S. G (2016). Life Time Data: Statistical Models and Methods 2nd Edition, World Scientific.
5. Gross, A.J. and Clark, V.A. (1976). Survival Distributions-Reliability applications in Bio-Medical Sciences, Chapters 3,4 John Wiley and Sons.
6. Kalbfleisch, J.D. and Prentice, R.L. (1980). The Statistical Analysis of Failure time Data, John Wiley and Sons.
7. Mahesh K. B. Parmar and Machin David (1999). Survival Analysis: A Practical Approach, Wiley.

3. Bayesian Inference

Credits: 4

Unit	Content	Hours
I	Subjective interpretation of probability in terms of fair odds; Subjective prior distribution of a parameter; Bayes theorem and computation of posterior distribution. Natural conjugate family of priors for a model. Conjugate families for exponential family models, and models admitting sufficient statistics of fixed dimension. Mixtures from conjugate family, Jeffreys' invariant prior. Maximum entropy priors.	10
II	Utility function, expected utility hypothesis, construction of utility function, Loss functions: (i) bilinear, (ii) squared error, (iii) 0-1 loss, and (iv) Linex. etc. Elements of Bayes Decision Theory, Bayes Principle, normal and extensive form of analyses.	10
III	Generalized maximum likelihood estimation. Bayes estimation under various loss functions. Evaluation of the estimate in terms of the posterior risk, Pre posterior analysis and determination of optimal fixed sample size. Linear Bayes estimates. Empirical and Hierarchical Bayes Methods of Estimation. Bayesian interval estimation: Credible intervals, HPD intervals, Comparison with classical confidence intervals.	10
IV	Bayesian testing of hypotheses, specification of the appropriate form of the prior distribution for a Bayesian testing of hypothesis. Prior and posterior odds. Bayes factor for various types of testing hypothesis problems. Lindley's method for Significance tests, two sample testing problem for the parameters of a normal population. Finite action problem and hypothesis testing under "O-Ki" loss, function. Large sample approximation for the posterior distribution. Lindley's approximation of Bayesian integrals.	10

References:

- Aitchison, J. and Dunsmore, I.R. (1975). Statistical Prediction Analysis, Cambridge University Press.
- Bansal, A. K. (2007). Bayesian Parametric Inference, Narosa Publishing House, New Delhi.
- Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis, Springer Verlag, New York.
- Box, G.E.P. and Tiao, G.C. (1973). Bayesian Inference in Statistical Analysis, Addison & Wesley.
- DeGroot, M.H. (2004). Optimal Statistical Decisions, McGraw Hill.
- Leonard, T. and Hsu, J.S.J. (1999). Bayesian Methods, Cambridge University Press.
- Lee, P. M. (1997). Bayesian Statistics: An Introduction, Arnold Press.

15. Robert, C.P. (2001). The Bayesian Choice: A Decision Theoretic Motivation, Second Edition, Springer Verlag, New York.

4. Reliability Theory

Credits: 4

Unit	Content	Hours
I	Reliability concepts and measures: components and systems, coherent systems, reliability of coherent systems, cuts and paths, modular decomposition, bounds on system reliability, structural and reliability importance of components.	10
II	Life distributions and associated survival, conditional survival and hazard rate functions. Exponential, Weibull, gamma life distributions and estimation of their parameters, Concepts of Type-I (time), Type-II (order) and random censoring.	10
III	Life distributions, exponential Gamma, Weibull, lognormal, Pareto, linear failure rate. Inference for exponential, Gamma, Weibull distributions under censoring.	10
IV	Failure rate, mean residual life and their elementary properties. Ageing classes and their properties, bathtub failure rate.	10

References:

1. Barlow, R.E. and Proschan, F. (1985). Statistical Theory of Reliability and Life Testing, Holt, Rinehart and Winston.
2. Lawless, J.F. (1982). Statistical Models and Methods of Life Time Data, John Wiley Models, Marcel Dekker.
3. Shaked, M. and Shanthikumar, J.G.(2007). Stochastic Orders & Their Applications, Academic Press.
4. Nelson, W. (1982). Applied Life Data analysis; John Wiley Stochastic orders and their Application, Academic Press.
5. Zacks, S. (1992). Introduction to Reliability Analysis: Probability Models and Statistical Methods, Springer Text in Statistics.
6. Deshpande, J.V. and Purohit, S.G. (2016) : Life Time Data: Statistical Models and Methods 2nd Edition, World Scientific.