GOVERNMENT OF ARUNACHAL PRADESH
DEPARTMENT OF PLANNING
ITANAGAR.

SYLLABUS FOR DIRECT RECUITMENT OF RESEARCH OFFICER

STATISTICS – PAPER –I (Pg. No. 2-6) & II (Pg. No. 7-14)

(PAPER- I)

MATHEMATICAL ANALYSIS

UNIT I: Recap of elements of set theory; Introduction to real umbers. Introduction to n-dimensional Euclidian space; open and closed intervals (rectangles), compact sets, Bolzano – Weirstrass theorem, Heine – Borel theorem.

UNIT II: Sequences and series; their convergence. Real valued function; continuous functions, uniform continuity. Differentiation; maxima – minima of functions, functions of several variables, constrained maxima – minima of functions.

UNIT III: Multiple integrals and their evaluation by repeated integration, change of variables in multiple integration, Improper integrals. Differentiation under the sign of integral – Leibnitz rule.

LINEAR ALGEBRA

UNIT I: Vector spaces over fields of scalars, subspaces, linear independence of vectors, basis and dimension of a vector space, completion theorem, orthogonality of vectors and subspaces, Vector spaces with an inner product, gram-Schmidt ortghogonalization process, orthonormal basis, linear transformations and projections and their representation by matrices.

UNIT II: Non-singular matrices and their inversion, determinants, ranks, row and column rank of a matrix, Idempotent matrix, its properties, trace, invariance theorems, Sylvestor and Frobenious inequalities, elementary matrices, partitioned matrices, G – inverse, Kronecker product. Systems of homogeneous and non-homogeneous linear equations, their consistency and maximal linearly independent solutions, minimal and characteristic polynomials of a square matrix, characteristic roots and vectors, Cayley – Hamilton theorem, similarity and diagonalization of square matrices. Real quadratic forms and their value classes, canonical reductions and simultaneous reducibility of quadratic forms.

PROBABILITY THEORY

UNIT I: Classes of sets, fields, sigma-field, Borel sigma-field in RK, sequence of sets, limsup and liminf of a sequence of sets. Measure, Probability measure, properties of a measure, Caratheodory extension theorem (statement only).

UNIT II: Measurable functions as limit of simple functions, Random variables, sequence of random variables, almost sure convergence, convergence in

probability (and in measure). Integration of a measurable function with respect to a measure. Expectation and moments. (statements of) Monotone convergence theorem, Fatou's lemma, and Dominated convergence theorem (and discussion).

UNIT III: Probability inequalities (Tchebyshef, Markov, Jensen). Independence Borel-Cantelli Lemma, Kolmogorov zero-one law and Borel zero-one law. Kolmogorov's Strong Law of Large numbers for iid sequences.

UNIT IV: Convergence in distribution, characteristic functions and their elementary properties, Inversion and Uniqueness theorem (statement and discussion), Polya's theorem and Levy's continuity theorem (statement only), de Moivre-Laplace Central Limit Theorem (CLT), Lindeberg-Levy's CLT, statement and discussion of Lindeberg-Feller's CLT.

DISTRIBUTION THEORY I

UNIT I: Joint, marginal and conditional pmfs and pdfs. Computations of probability, expectations and variances by conditioning, Generating functions (m.g.f and p.g.f) of random variables, their properties and applications.

UNIT II: Some continuous distributions (Cauchy, pareto, Weibull, lognormal), Bivariate normal and bivariate exponential distributions and their properties, multinomial distribution.

UNIT III: Functions of random variables and their distributions using Jacobian and other tools, convolution and compound distributions, truncated and mixture distributions.

UNIT IV: Sampling distributions from normal population central and non-central Chi-square, t and F distributions.

DISTRIBUTION THEORY II

UNIT I: Order statistics and their distributions and properties. Joint and marginal distributions of order statistics. Extreme values and their asymptotic distribution (statement only) with applications, Asymptotic distribution of median, distribution of quantities.

UNIT II: Multivariate normal distribution, p.d.f and c.d.f moments, marginasl and conditional distributions.

UNIT III: Distribution of linear and quadratic forms in normal variables, expectations, variances and covariances, characteristic functions, independence of

quadratic forms, conditions for a quadratic form to be distributed as chi-square and non-central chi-square, decomposition of quadratic forms, Cochran's theorem and James theorem.

SAMPLE SURVEY

UNIT I: Simple random sampling – estimation based on distinct units in srswor. Systematic sampling (circular, population with trend), domain estimation in srs.

UNIT II: Unequal probability sampling; pps wr and wor methods (including Lahiri's scheme) and related estimators of a finite population mean. Hansen – Hurwitz and Desraj estimators for a general sample size and Muthy's estimator for a sample of size 2. Horvitz – Thompson Estimator (~~THE~~). (HTE).

UNIT Iii: Stratified sampling – allocation problem and construction of strata.

UNIT IV: Cluster sampling. Two-stage sampling. Ratio and regression estimators based on srswor method of sampling, Double sampling.

UNIT V: Non-sampling errors, modeling observational errors, application to longitudinal studies. Randomized response technique: Warner's related question model, unrelated question model.

STATISTICAL INFERENCE - I

UNIT I: Parametric models: Identifiable (indexing) parametric set up, estimation (point and interval) and testing of hypotheses, joint distribution of a sample and induced sampling distribution of a statistic; examples form standard discrete and continuous models. Likelihood function and information in datat about the parameter, concept of non-information, sufficiency, Neyman factorizability criterion, likelihood equivalence, minimal sufficient statistic, exponential family, invariance property of sufficiency, Fisher information for one and several parameters model.

UNIT II: Methods of estimation; maximum likelihood method, method of moments method of minimum chi-square, method of scoring. Choice of estimators based on unbiasedness, minimum variance, mean squared error, minimum variance unbiased estimator, Rao-Backwell theorem, completeness, Lehmann-Scheffe theorem, necessary and sufficient conditions for MVUE, Cramer – Rao Scheffe inequality. Consistency and CAN (statement only)

UNIT III: Tests of Hypotheses: concepts of critical regions, test functions, two kinds of errors, size function, power function, level, MP tet, Neyman-Pearson

Lemma and likehood ratio test, asymptolic distribution of L.R. statistic (statements only)

UNIT IV: Interval estimation; confidence level, construction of confidence intervals using pivots, shortest expected length confidence interval, construction of confidence interval using tests of hypothesis.

LINEAR MODELS AND REGRESSION ANALYSIS

UNIT I: Gauss-Markov set-up, estimability of parameters, normal equations and least squares estimates, error and estimation spaces, variances and covariance of least squares estimates, estimation of error variance, estimation with correlated observations, least squares estimates with restriction on parameters, simultaneous estimates of linear parametric functions. Tests of hypotheses, linear models with restricted hypothesis, confidence intervals and regions, Analysis of Variance.

UNIT II: Simple linear regression fit of polynomials and use of orthogonal polynomials, multiple regression, logistic regression.

UNIT III: Residual and their plots as tests for departure from assumptions such as fitness of the model, normality, homogeneity of variances, correlated structure and detection of outliers; Remedies; Transformation; Power transformations for dependent and independent variables, problems of multicolinearity.

INTRODUCTION TO PROBABILITY THEORY AND DISTRIBUTIONS

UNIT I: Important Concepts in Probability: Definition of probability – classical relative frequency – subjective and axiomatic approach to probability, merits and demerits of these approaches (only general ideas to be given). Random Experiment: Trial, sample point and sample space, definition of an event, operation of events, mutually exclusive and exhaustive events. Discrete sample space, combinatorics, properties of probability based on axiomatic approach, conditional probability, independence of events, Bayes' theorem and its applications.

UNIT II: Random Variables: Definition of discrete random variables, probability mass function, idea of continuous random variable, probability density function, illustrations of random variables and its properties, expectation of a random variable and its properties – moments, measures of location, dispersion, skewness and kurtosis. Chebyshev's inequality and applications, statements and applications of weak law of large numbers, central limit theorems.

UNIT III: Standard univariate discrete distributions and their properties: Discrete Uniform, Binomial, Poisson, Hypergeometric, and Negative Binomial distributions. Continuous univariate distributions – uniform, normal, exponential, gamma and beta distributions.

INTRODUCTION TO SAMPLE SURVEY

UNIT I: Sample surveys, Concepts of population and sample, need for sampling, census and sample survey, basic concepts in sampling, organizational aspects of survey sampling, sample selection and sample size.

UNIT II: Some basic sampling methods – simple random sampling (SRS) with and without replacement.

UNIT III: Stratified random sampling, systematic sampling (linear only), introduction to ratio and regression methods of estimation under SRS.

UNIT IV: Non sampling errors, acquaintance with the working (questionnaires, sampling design, methods followed in field investigation, principal findings etc.) of NSSO and other agencies undertaking sample surveys.

LINEAR PROGRAMMING

UNIT I: Introduction to Linear Programming (LP). Mathematical Formulation of Linear Programming Problem (LPP). Graphical solution to LPP.

UNIT II: General LPP, Canonical and Standard forms of General LPP, Duality in LPP, Simplex Method. Big-M method and Two-phase method.

UNIT III: Transportation and Assignment problems. (Including Traveling Salesman's Problem).

PAPER -II

DESIGN OF EXPERIMENTS

UNIT I: Block Design, multiple comparison, Von Ferroni, Turkey and Scheffe, simultaneous confidence interval. Incomplete Block Design, Balance Incomplete Block Design (BIBD), introduction to Partially Balanced Incomplete Block Design (PBIBD), Analysis of covariance in a general Gauss-Markov model.

UNIT II: General factorial experiments, factorial effects, best estimates and testing the significance of factorial effects, study of 2 and 3 factorial experiments in randomized blocks, complete and partial confounding. Fractional replication for symmetric factorials.

UNIT III: Application areas, response surface experiments, clinical trials, longitudinal data.

STATISTICAL INFERENCE II

UNIT I: Unbiased test, UMP and UMPU tests, Wald's SPRT with prespecified errors of two kinds

UNIT II: One sample location problem, sign test and signed rank test, one and two sample Kolmogorov Smirnov tests. Two sample location problems. Wilcoxon-Mann-Whitney test, normal score test, ARE of various tests based on linear rank statistics. Kruskal-Wllis K sample test, one and two sample U statistics, asymptolic of U statistics.

UNIT III: Basic concepts of decision theory; inference problems viewed as decision problem. Problem of classification, minimax approach and Baye's approach, structure of Baye's rule, complete class of rules, construction minimax rule.

UNIT IV: Concepts and evaluation of subjective probability of an event; subjective prior distribution of a parameter. Baye's theorem and computation of posterior distribution. Natural conjugate family of prior for a model. Loss function, Baye's risk. Bayesian estimation of parameters of binomial, poisson, normal and exponential distributions.

STATISTICS FOR NATIONAL DEVELOPMENT & ENVIRONMENTAL STATISTICS

UNIT I: Indices of development, human development index, Estimation of national income – product approach, income approach and expenditure approach.

UNIT II: Population projection using Leslie matrix, Measuring inequality in incomes. Gini's coefficient, Theil's measure. Poverty measurement – different issues, measures of incidence and intensity, indices due to Kakwani, Sen etc.

UNIT III: Ecological diversity, Species abundance curve, indices of diversity, richness and evenness.

UNIT IV: Harvesting renewable biological resources – maximum sustainable yield, Bionomic equilibrium.

INTRODUCTION TO ECONOMETRICS

UNIT I: Nature of econometrics. The general linear model (GLM) and its extension. Ordinary least squares (OLS) estimation and prediction. Use of dummy variables and seasonal adjustment. Generalized least squares (GLS) estimation and prediction. Heteroscedastic disturbances. Pure and mixed estimation. Grouping of observations and of equations.

UNIT II: Auto correlation, its consequences and tests. Theil BLUS procedure. Estimation and prediction. Multicollinearity problems, its implications and tools for handling the problem. Ridge regression.

UNIT III: Linear regression with stochastic regressors. Instrumental variable estimation. Errors in variables. Autoregressive linear regression. Distributed lag models. Use of principal components, canonical correlations and discriminant analyses in econometrics.

BIOMETRY

UNIT I: An introduction to Biometry and Statistics: data collection and data presentation, frequency distribution, graphical representation, measures of central tendency, dispersion, skewness and kurtosis. Probability distribution: Binomial, Poisson and Normal distribution.

UNIT II: Introduction to bivariate frequency data and its measurement: covariance, correlation, scatter diagram. Regession analysis: Liner regression, regression coefficient, fitting of regression equation by least square method.

UNIT III: Population, sample. Statistic, standard error, estimation, confidence interval and confidence level, confidence interval estimate of proportion and mean. Hypothesis and its types, errors, level of significance. Test statistics: Student's chi-square, F and Z-Statistics and their applications in testing of hypothesis.

UNIT IV: An introduction to Analysis of Variance (ANOVA), its definition, assumptions and uses. One way classification and statistical analysis of the model involved in it.

SURVIVAL ANALYSIS

UNIT I: Concepts of time, order and random censoring, Types of censoring, likelihood in these cases survival Functions and Hazard rates, Life time distributions – Exponential, Gamma, Weibull, Lognormal, Linear Failure rate, Parametric Inference (point estimation, confidence intervals, Scores, LR, MLE tests) for these distributions.

UNIT II: Life tables, mean residential life and their elementary properties, Aging classes and their properties, Bathtub Failure jute, Estimation of survival function – Actuarial method, Product-limit (Kaplan-Meier) estimator, Hazard function estimator.

UNIT III: Two sample problem: Gehan test, Log rank test, Mantel – Hazenszel test, Tarone-Ware class of tests, Efron test, K-sample problems: Generalized Gehion test, Generalized Mantel – Haewszel test, Introduction to regression for failure rate.

APPLIED MUTLIVARIATE ANALYSIS

UNIT I: Multiple regression, multiple and partial correlation coefficients. MLEs of the parameters of multivariate normal distribution and their sampling distributions. Wishart distribution and its properties. Tests of hypothesis about the mean vector of a multinormal population, Hotelling's T^2 - statistic, its distribution and applications.

UNIT II: Classification and discrimination for two known populations: Bayes, Minimax and Likelihood Rati procedures. Mahabolis D^2 - Statistic and its application, Sample discriminant function and discrimination based on Fisher's method. Cluster Analysis and evaluation of clusters.

UNIT III: Introduction to principal component analysis, canonical correlation analysis, factor analysis. Methods and applications of MANOVA (without derivation of the distribution of Wilk's λ)

TIME SERIES ANALYSIS & FORECASTING

UNIT I: Stationary time Series, Auto correlation and Partial auto correlation function, correlogram analysis, Spectral properties of stationary models, periodogram analysis, spectral density function.

UNIT II: Exponential & moving average smoothing and forecasting, Detail study of stationary process: moving average, autoregressive, autoregressive moving average and autoregressive integrated moving average process, Box-Jenkins models.

UNIT III: Discussion (without proof) of estimation of mea, auto covariance and auto correlation function under large sample theory, choice of AR and MA periods, Estimation of ARIMA model parameters, forecasting with Box-Jenkins model, Residual analysia and diagnostic checking.

STOCHASTIC PROCESSES

UNIT I: Introduction to stochastic processes (sp's): classification of sp's according to state space and time domain. Countable state Markov chains (MC's), Chapman-Kolmogorov equations; calculation of n-step transition probability and its limit. Stationary distribution, classification of states; transient MC; random walk and gambler's ruin problem; Applications from social, biological and physical sciences.

UNIT II: Discrete state space continuous time MC: Kolmogorov – Feller differential equations; Poisson process, birth and death process; applications to queues and storage problems. Introduction to Wiener process.

UNIT III: Inference in Markov chains, estimation of transition probabilities, testing for order of a Markov chain, estimation of functions of transition probabilities, Parametric models and their goodness of fit.

OPERATIONS RESEARCH

UNIT I: Definition and scope of operations research; phases in operations research, models and their solutions, decision-making under uncertainty and risk, use of different criteria.

UNIT II: Review of linear programming (LP) problems – duality theorem, transportation and assignment problems; sensitivity analysis; non-linear programming; Kuhn Tucker conditions, Wolfe's and Beale's algorithms for solving quadratic programming problems.

UNIT III: Analytical structure of inventory problems; Economic order quantity (EOQ) formula of Harris, its sensitivity analysis and extension allowing quantity discounts and shortages. Multi-item inventory subject to constraints. P and Q – systems with constant and random lead items.

UNIT IV: Queueing models; specifications and effectiveness measures. Steady-state solutions of M/M/1 and M/M/c models with associated distributions of queue length and waiting time. M/G/1 queue and Pollazcek Khinchin result. Steady state solutions of M/Ek/1 Simulation.

RELIABILITY THEORY

UNIT I: Reliability concepts and measures; components and systems, coherent systems; reliability coherent systems, cuts and paths, modular decomposition, bounds and system reliability; structural and reliability importance components.

UNIT II: Life distributions; reliability function, hazard rate, common life distributions; exponential, weibull, gamma etc. Estimation of parameters and tests in these models.

UNIT III: Notions of ageing; increasing failure rate (IFR), increasing failure rate average (IFRA), not better than used (NBU), decreasing mean residual life (DMRL) and not better than used in expectation (NBUE). Classes and their duals; loss of memory property of the exponential distribution; closures of these classes under formation of coherent systems, convolutions and mixtures.

UNIT IV: Univariate shock models ad the distributions arising out of them; bivariate shock models, common bivariate exponential distributions and their properties. Reliability estimation based on failure times in variously censored life tests and tests with replacement of failed items; stress-strength reliability and its estimation.

STATISTICAL ECOLOGY

UNIT I: Introduction to Ecology and evolution. Population dynamics: Single species – monomolecular, logistic and Gompertz models, Leslie matrix model for age and stage structured population.

UNIT II: Survivorship curves – Constant, monotone and bathtub shaped hazard rates. Abundance estimation: nearest neighbor, line transect sampling, forest sampling with satellite remote sensing.

UNIT III: Ecological diversity: Species abundance curve, indices of diversity (Simpson's index, Shannon – Wiener index). Diversity as average ratity. Harvesting renewable biological resources – Maximum sustainable yield, Usher's approach. Forestry management – Faustmann model.

UNIT IV: Game theory in ecology – Evolutionarily stable strategy, its properties, simple games such as Hawk – Dove game. Foraging theory: Optimal foraging, diet choice, mean variance trade-off.

STATISTICAL COMPUTING

UNIT I: Introduction to object-oriented programming in C++, simple syntax, data types and operations, functions and parameters, classes, input/output, control statements, loops, pointers and arrays, libraries, linking to databases.

UNIT II: Introduction to S-PLUS/R environment, basics of S language, objects and classes, connections, data manipulations, expressions/assignments, functions, control structures, array and matrix operations, graphics and their control.

UNIT III: Writing programs in C++/S/R for: matrix computations, solutions of linear system of equations, computation of eigen values and eigen vectors, simple hypothesis tests, analysis of variance, iinear regression, random number generation, simulation and Monte-Carlo technique.

UNIT IV: Exposure to statistical packages S-PLUS/R/SPSS/SYSTAT, analysis of interesting data sets using one or more of packages, graphics, descriptive statistics, representation of multivariate data, simple hypothesis tests, analysis of variance, linear and nonlinear regression, principal component, discriminant and cluster analysis, analyses of time series data.

POPULATION DYNAMICS AND DEMOGRAPGHY

UNIT I: Introduction to Population: Meaning of Population, Size, structure, distribution of population, the structure of demographic rates. Age-sex pyramids. Demographic data: Census, Registration system, Indian SRS, and surveys. NFSH- 1, 2. Evaluation of Quality of demographic data: Chandrasekaran-Deming formula, accuracy of data on sex and age: Whipple's, Myer's and UN indices.

UNIT II: Mortality: concepts and rates, measures of infant mortality rate. Force of mortality. Mortality laws-Gompertz and Makeham. Life table and its

construction: complete and abridged. Greville's, reed-Merrel's and Chiang's methods.

UNIT III: Fertility and Reproduction: Period and cohort measures. P/F ratio and own children method. Reproductive measures. Nuptiality rates. Gross and net nuptiality tables. Internal and international Migration: concept and rates. Uses of place of birth and duration of residence data.

UNIT IV: Theory of stable population model (one sex), quasi and stationary population. Lotks's stable population model. The equations characterizing a stable population, the effect of changes in fertility and mortality on age structure, growth rates, birth rates and death rates. Momentum of population growth. Population, projection: Mathematical curves viz., growth curves, modified exponential, logistic curves and its properties, and their fitting, component method and matrix method of population projection.

ADVANCED SAMPLE SURVEY

UNIT I: General notions of sampling designs and sampling schemes and their equivalence [statement only]; fixed effective sample size and variable effective sample size sampling designs; inclusion probabilities of first and second order and their inter-relations for fixed effective sample size sampling designs; mean and variance of varying effective sample sizes and their relationship with the inclusion probabilities of first two orders; notion of design-unbiasedness; unbiased estimation linear and quadratic functions of population quantities; unbiased estimation of a finite population total and finite population variance. Illustrative examples.

UNIT II: Horvitz-Thompson Estimator [HTE] of a finite population total; expression for variance of the HTE for a general sampling design and its alternative expression for fixed effective sample size sampling designs; necessary and sufficient conditions on the sampling designs for existence of unbiased variance estimators and study of their non-negativity. Illustrative examples. Concept of a super-population; model-expectation and model-variance. Godambe's result on optimality of the HTE. Optimum allocation under super population model.

UNIT III: Stratified sampling, systematic sampling, cluster sampling and two-stage sampling designs viewed as special cases of general sampling designs; discussions on existence/ non-existence of variance estimators for unbiased estimator of finite population totals; post-census survey methodologies.

UNIT IV: Large-scale surveys conducted by various rounds of NSSO related to (i) agricultural production, (ii) industrial production, (iii) trade, (iv) services

250
(12)

sectors, (v) socio-economic statistics, (vi) price statistics and (vii) national accounts statistics – with special reference to coverage and periodicity and responsible organizations.