# UNIVERSITY OF ALLAHABAD

## COMBINED RESEARCH ENTRANCE TEST 2015 (CRET 2015)

**SYLLABUS: Statistics**

**Note for Level 1 B and Level 2 tests:**

**Level 1 B:**

There will be 50 multiple objective type questions covering ENTIRE syllabus. 3 marks will be awarded for correct answer and 1 mark will be deducted for each wrong answer.

Duration: $1\frac{1}{2}$ Hours (90 minutes).

**Level 2: There will be in all 14 Questions divided into 3 Parts.**

There will be three sections.

A. In Section A, there will be 10 Short answer (50 words) type questions covering entire syllabus (all compulsory). Total marks of this part will be 100 *i.e.* $10 \times 10$.
B. In Section B, there will be 3 Medium answer (200 words) type questions with internal choice/s considering specializations (all compulsory). Total marks for this part will be 60 *i.e.* $20 \times 3$.
C. In Section C, there will be 1 Long answer (400 words) type question on research methodology and research aptitude with internal choice/s. total marks of this part will be 40.

Duration 3 Hours (180 minutes)

Total marks will be 200.

**Probability Theory:** Probability space of a random experiment, probability measures, random variables as a measurable function, $\sigma$-field induced by a sequence of random variables, decomposition of distribution functions in purely discrete, absolutely continuous and singular components, $C_r$.-inequality, Cauchy-Schwartz inequality, Hölder inequality, Minkowski inequality, Jensen inequality, Lyapunov inequality, Kolmogorov inequality, Hajek-Rényki inequality, Sequences of distribution functions, Helly Bray theorem, Defferent types of convergence of sequence of random variables, distribution function of random vectors, Weak and strong law of large numbers, Khinchin, Borel and Kolmogorov theorems, Borel-Cantelli lemmas and Zero-one law, Characteristic function, Inversion theorem, Continuity theorem, One dimensional central limit problem: Lindeberg-Levy, Lyapunov, Lindeberg-Feller theorems.

**Time Series Analysis:** Time series as a stationary or nonstationary stochastic process, time domain analysis based on correlogram, sample autocovariance function (acvf) and autocorrelation function (acf) at lag k, AR(p) process, MA(q) process, mixed ARMA(p,q) process, stationarity and invertibility conditions, ARIMA(p,d,q) model, estimation of parameters, tests for stationarity, frequency domain analysis based on the spectral density function, spectra of AR(1) and MA(1) models, periodogram and its relationship with acvf, forecasting by exponential smoothing and Box-Jenkins procedures.

**Multivariate Analysis:** Multivariate normal distribution, Characteristic function, Maximum likelihood estimators of the mean vector and covariance matrix, Multiple and partial correlation coefficients and their null sampling distributions, Wishart distribution, Hotteling's , Mahalanobis' and their applications, Discriminant analysis.

**Statistical Inference:** Sufficiency, Fisher- Neyman – Halmos – Savage factorization criterion, minimal sufficiency, Completeness, Bounded completeness, Ancillary statistics, Basu's theorem on independence of Statistics, exponential family, Bhattacharya bound, Chapman Robbins and Kiefer (CRK) bound, Generalized Rao Cramér bound for the multiparameter case, Maximum likelihood estimation, Zehna theorem for invariance, Cramér theorem for weak consistency, asymptotic normality, BAN and CAN estimators, asymptotic efficiency, equivariant estimation,Generalized Neyman Pearson lemma, UMP tests for distributionswith MLR, LR, tests and their properties, UMPU tests, similar regions, Neyman structure, Invariant tests. SPRT, Fundamental identity, OC and ASN functions, Sequential estimation, Wald's equation, Wolfowitz generalization of FRC bound, Stein's two stage procedure, asymptotic theory of sequential estimation, sequential estimation of normal mean.

**Analysis of Variance and Design of experiments:** Two-way classification with equal number of observations per cell and Tukey's test, general two-way classification, Analysis of covariance, $2^n$ , $3^2$ and $3^3$ factorial experiments, complete and partial confounding, Balanced Incomplete Block Design (BIBD), construction of BIBD, intra block and inter block analysis, Partially Balanced Incomplete Design (PBIBD), split plot design.

**Sampling Theory:** Varying probability sampling with and without replacement, cumulative total and Lahiri's methods of selection, Estimation of population mean, Desraj ordered estimates, Horwitz-Thompsom estimator, Midzuno, and Narain system of sampling, post-stratification and deep stratification, double sampling in ratio and regression estimation, two stage and multi-stage sampling, basic idea of randomized response technique, non sampling errors.

**Nonparametric Inference:** Asymptotic distribution of an order statistic, Sufficiency and completeness of n-tuple of order statistic, nonparametric estimation of distribution function and Glivenko-Cantelli fundamental theorem of statistics, one sample and two sample location tests, Application of U-statistic to rank tests, One sample and two sample Kolmogorov-Smirnov tests, The Mann- Whitney U-test, Kruskal-Wallis one-way ANOVA test, Friedman's two-way ANOVA by ranks, Efficiency criteria, Pitman ARE.

**Econometrics:** Linear regression model, assumptions, estimation of parameters by least squares and maximum likelihood methods, tests of hypothesis and confidence estimation for regression coefficients, $R^2$ and adjusted $R^2$, use of extraneous information in terms of exact and stochastic linear restrictions, restricted restriction and mixed regression methods and their properties, point and interval predictors, tests for structural change, use of dummy variables, problem of multicollinearity, consequences and solutions, estimation of parameters by generalized least squares in models with non-spherical disturbances, heteroscedasticity of disturbances, estimation under heteroscedasticity and tests of heteroscedasticity, autocorrelation, Durbin-Watson test, estimation under autocorrelated disturbances, errors in variable models, inconsistency of least squares method, instrumental variable method, seemingly unrelated regression equation (SURE) model and its estimation, simultaneous equations model, concept of structural and reduced forms, problem of identification, rank and order conditions of identifiability, indirect least squares, two stage least squares and limited information maximum likelihood estimation, idea of three stage least squares and full information maximum likelihood estimation.