Applied Statistics Syllabus

Session 2020-2021

Collected by,
Md Nayeem Chowdhury
ISRT, DU

Basic Statistics

❖ Introduction to Statistics:

- Meaning of statistics
- Scopes and limitations
- Concepts of descriptive and inferential statistics

❖ Basic Concepts:

- Data
- Sources of data (primary and secondary)
- Population, sample, parameter, statistic
- Variables and types of variables (qualitative, quantitative discrete, and continuous)
- Scales of measurements
- Classification of variables by scales of measurements

❖ Producing Data:

- Approaches to producing data
- Concept of experimental study and non-experimental study
- Introduction to sample survey and questionnaire
- Concept of electronically recorded data and hospital recorded data
- Concept of data cleaning and checking before statistical analysis

❖ Organization and Presentation of Data:

- Graphical presentation for qualitative and quantitative data
- Sorting data
- Grouping qualitative and quantitative data
- Construction of frequency distribution and relative frequency distribution
- Graphical presentation of frequency distribution (histogram, frequency polygon, ogive)

❖ Concept of Distribution:

- Location, scale (spread), and shape
- Illustration with stem-and-leaf diagram
- Descriptive measures of data
- Measures of location
- Measures of dispersion
- Moments and their interrelationship
- Measures of skewness and kurtosis
- Three- and five-number summary
- Box-plot and modified box-plot

❖ Description of Bivariate Data:

- Bivariate frequency distribution
- Graphical presentation of bivariate data
- Contingency table
- Concept of association between two variables
- Percentage table and interpretation of cell frequencies
- Measures of association for nominal and ordinal variables
- Measures of association for interval or ratio variables
- Correlation
- Relationship between two variables: simple linear regression
- Basic issues in inferential statistics

Probability and Random Variables

❖ Combinatorial Analysis:

- Basic principles of counting
- Permutations
- Combinations

❖ Axioms of Probability:

- Sample space and events
- Axioms of probability
- Sample spaces having equally likely outcomes
- Probability as a measure of belief

❖ Conditional Probability and Independence:

- Conditional probabilities
- Bayes formula
- Independent events

❖ Random Variables:

- Introduction
- Discrete random variables
- Expectation
- Expectation of a function of a random variable
- Variance
- Bernoulli and binomial random variables
- Poisson random variable
- Other discrete random variables (geometric, negative binomial, hypergeometric)
- The expected value of a sum of random variables
- Properties of the cumulative distribution function
- Continuous random variables
- Expectation and variance of the continuous random variable
- Normal random variable
- Normal approximation to the binomial distribution
- Exponential random variables

❖ Jointly Distributed Random Variables:

- Joint distribution functions
- Independent random variables
- Sums of independent random variables
- Conditional distributions (discrete and continuous cases)

- Properties of expectation
- Expectation of sums of random variables
- Covariance
- Variance of sums
- Correlations
- Conditional expectation
- Moment generating functions
- Probability generating function

❖ Generating Function Techniques:

- Moment generating function
- Cumulant generating function
- Probability generating function
- Characteristic function
- Finding distributions of functions of random variables
- Change of variable technique
- Distribution function technique
- Moment-generating function technique
- Probability integral transformation
- Statistic and sampling distribution
- Law of large numbers
- Central limit theorem
- The exact distribution of the sample mean
- Chi-square distribution and its properties
- F-distribution and its properties
- T-distribution and its properties
- Non-central chi-square, F, and t distributions
- Concept of order statistics
- Distributions of single-order statistics
- Joint distribution of two or more order statistics
- A brief review of some discrete distributions
- Continuous probability distributions and their properties (uniform, normal, exponential, gamma, beta, log-normal, Cauchy)
- Definition of the truncated distribution
- Definition of compound and mixture distribution
- Family of distributions: Pearsonian distribution

❖ Simulation and Random Number Generations:

- Concepts of simulation and its uses in statistics
- Random number generators (congruential generators, seeding)
- Random variate generations (inversion method, rejection method)
- Simulating discrete random variables

- Simulating normal random variables (rejection with exponential envelope, Box-Muller algorithm)
- Monte-Carlo integration (hit-and-miss method, improved Monte-Carlo integration)
- Variance reduction techniques (antithetic sampling, importance sampling, control variates)

Design of Experiments

❖ Introduction to Design of Experiments:

- Strategy of experimentation
- Some typical examples of experimental design
- Basic principles
- Guidelines for Designing Experiments

❖ Experiments with a Single Factor:

- The analysis of variance
- Analysis of fixed effects model
- Estimation of model parameters
- Unbalanced data
- Model adequacy checking
- Regression model
- Comparisons among treatment means
- Graphical comparisons of means
- Contrasts
- Orthogonal contrasts
- Multiple testing
- Scheffe's method
- Comparing pairs of treatments means
- Comparing treatment means with a control
- Determining sample size
- Operating characteristic curve
- Specifying standard deviation increase
- Confidence interval estimation method
- Discovering dispersion effects
- Regression approach to analysis of variance
- Least squares estimation of model parameters
- General regression significance test

❖ Randomized Blocks, Latin Squares, and Related Designs:

- The randomized complete block designs (RCBD)
- Statistical analysis of RCBD
- Model adequacy checking
- Estimating model parameters
- Latin square design
- Graeco-Latin square design
- Balanced incomplete block design (BIBD)
- Statistical analysis of BIBD

- Least squares estimation of BIBD
- Recovery of intra-block information in the BIBD

❖ Introduction to Factorial Designs:

- Basic definition and principles
- Advantage of factorials
- Two-factor factorial design
- Statistical analysis of fixed effects model
- Model adequacy checking
- Estimating model parameters
- Choice of sample size
- Assumption of no interaction in a two-factor model
- General factorial design
- Fitting response curves and surfaces
- Blocking in a factorial design

❖ 2^k Factorial Design:

- Introduction
- 2^2 design
- 2^3 design
- General 2^k design
- A single replicate in 2^k factorial design
- Blocking in a 2^k factorial design
- Confounding in 2^k factorial design
- Confounding in 2^k factorial design in two blocks
- Confounding in 2^k factorial design in four blocks
- Confounding in 2^k factorial design in 2^p blocks
- Partial confounding

❖ Two-Level Fractional Factorial Designs:

- One-half fraction of the 2^k design
- One-quarter fraction of 2^k design
- General 2^{k-p} fractional factorial design
- Resolution III designs
- Resolution IV and V designs

❖ Three-Level and Mixed-Level Factorial and Fractional Factorial Designs:

- 3^k factorial design
- Confounding in 3^k factorial design

- Fractional replication of 3^k factorial design
- Factorials with mixed levels

❖ Response Surface Methods:

- Introduction to Response Surface Methodology
- Method of steepest ascent
- Analysis of second-order response surface
- Experimental designs for fitting response surfaces
- Mixture experiments
- Robust designs

❖ Experiments with Random Factors:

- Random effects model
- Two-factor factorial with random factors
- Two-factor mixed model
- Sample size determination with random effects
- Rules for expected mean squares
- Approximate F tests
- Approximate confidence intervals on variance components
- Modified large-sample method
- Maximum likelihood estimation of variance components

❖ Nested and Split-Plot Designs:

- Two-stage nested designs
- Statistical analysis, diagnostic checking, variance components
- General m-staged nested design
- Designs with both nested and factorial factors
- Split-plot design
- Split-plot designs with more than two factors
- Split-split-plot design
- Strip-split-plot design

Sampling Theory and Design

❖ Introduction:

- Concept of sampling and related terms
- Role of sampling theory
- Requirements of a good sampling design
- Steps in a sample survey
- Probability and nonprobability sampling
- Selection (draw-to-draw) and inclusion probability
- Sampling weight
- With and without replacement sampling
- Characteristics of estimates: bias, mean square error, and variance (precision)
- Errors in sample survey and census
- Sample size determination: basics and complex scenarios

❖ Simple Random Sampling (SRS):

- Sample selection
- Estimation: mean, total, proportion, ratio of two quantities
- Unbiasedness and variances/standard errors (SEs) of the estimators
- Estimators of the SEs
- Confidence interval (normal approximation)
- Finite population correction
- Estimation over subpopulation
- Computation: inclusion probabilities and sampling weights

❖ Systematic Sampling:

- Motivation, use, and challenges
- Sample selection
- Different estimators and their unbiasedness and variances
- The estimator of the variances
- Comparison with SRS
- Sampling from a population with linear trend or periodic variation

❖ Stratified Random Sampling:

- Concept, reasoning, and needs in a heterogeneous population
- Number and formation of strata
- Sample selection
- Estimators (total, mean, proportion)
- Variances of the estimators
- Estimators for the variances

- Different allocation techniques
- Comparison with SRS
- Design effect and its uses
- Poststratification
- Quota sampling

❖ Auxiliary Information in Estimation:

- Ratio estimators (total, mean)
- Properties: unbiasedness, variance (approximate), estimated variance, confidence interval
- Comparison with mean per unit estimates
- Conditions for best linear unbiased ratio estimator
- Application in stratified sampling
- Unbiased ratio-type estimates
- Product estimator
- Regression estimator: linear regression estimate and its properties under preassigned b and estimated b
- Comparison with mean per unit estimate
- Application in stratified sampling
- Relative merits and demerits

❖ Cluster Sampling:

- Motivation and reasoning
- Formation and size of clusters
- Cluster sampling with equal-sized clusters
- Estimators and their various properties
- Comparison with SRS and systematic sampling
- Optimum cluster size
- Stratification in cluster sampling: estimation and comparison with simpler sampling designs

❖ Special Sampling Designs:

- Capture-recapture sampling: implementation, Peterson and Chapman estimators for population size and their variances, Hypergeometric and Multinomial models for estimating population abundance
- Ranked set sampling: sample selection and estimation

❖ Probability Proportional to Size (PPS) Sampling:

- Motivating examples
- With replacement (WR) sampling: cumulative measure of size method and Lahiri's method, Hansen-Hurwitz (H-H) estimator
- Comparison with SRS

- Optimum measure of size
- Without replacement (WOR) sampling: challenges and solutions, initial probabilities, normalizing probabilities, inclusion probabilities and their relation with the sample size, Horvitz-Thompson (H-T) estimator
- Different methods of PPSWOR
- Multinomial distribution for PPSWR sampling
- H-T estimator in case of PPSWR sampling

❖ Sub-sampling and Multi-stage Sampling:

- Sub-sampling of unequal-sized clusters: different estimators and their variances
- Two-stage sampling: design, estimators (total, mean), variances, and their unbiased estimators
- Three-stage sampling: design, estimators (total, mean), variances and their estimators
- General framework (two-stage and three-stage) for estimating population total
- Different sampling designs at different stages
- Determination of sample sizes in two and three-stage sampling
- Optimum sampling and sub-sampling fractions
- Use of information from a pilot survey

❖ Double Sampling and Repeated Sampling:

- Concept of double sampling and its necessity
- Application in stratified sampling, and Ratio and Regression estimators
- Repeated sampling from the same population: sampling on two and more than two occasions

❖ Complex Survey:

- Definition and challenges involved in complex surveys
- Approaches of variance estimation (VE)
- Replication methods for VE: random group method, balanced repeated replication (balanced half-sample replication) method, Jackknife method, and Bootstrap method
- Implementation of replication methods in complex sampling designs
- Post-stratification

❖ Non-sampling Errors:

- Sources of errors
- Effects of nonresponse
- Inference on population proportion in the presence of nonresponse
- Types of nonresponse
- Call-backs and their effects
- Hansen and Hurwitz's approach to nonresponse
- Politz-Simmons adjustment for bias reduction

- Mathematical model for errors of measurement
- Mechanism of nonresponse
- Imputation and its different techniques

❖ Special Sampling Designs:

- Multiplicity
- Network sampling: design and estimation (multiplicity and Horvitz-Thompson estimators for population total, and their different properties)
- Adaptive sampling: adaptive cluster sampling (ACS) and related concepts used in ACS, Hansen-Hurwitz and Horvitz-Thompson estimators for population total, and their different properties

Statistical Inference

❖ Basic Concepts:

- Fundamental ideas of statistical inference
- Parametric and non-parametric inference
- Estimators, statistics, parameters
- Sampling distributions and their uses in inference
- Point estimation, interval estimation, and test of hypotheses
- Theory and reality

❖ Point Estimation of Parameters and Fitting of Probability Distributions:

- Descriptive statistics
- Exploratory data analysis
- Least squares estimation
- Moments based estimation
- Maximum likelihood estimation
- Uses of graphical tools for assessing goodness of fit
- Asymptotic distributions of maximum likelihood estimators

❖ Interval Estimation:

- Methods for constructing confidence intervals
- Pivotal quantity method
- Wald-type method
- Likelihood ratio-based method
- Confidence intervals for means, the difference of two means, proportions
- Interpretation of confidence intervals

❖ Testing Hypotheses and Assessing Goodness of Fit:

- Heuristics of hypothesis testing
- Errors in hypothesis testing
- Statistical significance and power
- Exact tests and approximate tests
- Tests about one population mean, equality of two population means, more than two population means, proportions
- Likelihood ratio test
- Statistical tests applied to categorical data: Fisher’s exact test, chi-square test of homogeneity and independence, chi-square goodness of fit tests

❖ Testing Hypotheses:

- Approaches to hypothesis testing (Neyman-Pearson, Fisher, Jeffreys)
- Error probabilities and the power function
- Best test concept
- The most powerful test via the Neyman-Pearson lemma
- Uniformly most powerful (UMP) test via the Neyman-Pearson Lemma
- Likelihood ratio property
- UMP test via maximum likelihood ratio property
- Unbiased and UMP unbiased tests

❖ Principle of Data Reduction:

- Sufficiency
- Conditional distribution approach
- Neyman factorization theorem
- Minimal sufficiency
- Lehmann-Scheffe approach
- Information in one-parameter and multi-parameter situations
- Ancillary
- Completeness
- Complete sufficient statistics
- Basu's theorem

❖ Likelihood-Based Inference in Exponential Families:

- Formulation
- Estimation in one-parameter and multi-parameter cases
- Approximate normality of MLEs
- Wald tests and confidence intervals
- Likelihood ratio test and confidence interval
- Inference about $g(\theta)$ using the delta method applied to MLEs

❖ Criteria to Compare Estimators:

- Unbiasedness, variance, mean squared error
- Best unbiased and linear unbiased estimators
- Improved unbiased estimator via sufficiency
- Rao-Blackwell theorem
- Uniformly minimum variance unbiased estimator (UMVUE)
- Cramer-Rao inequality and UMVUE
- Lehmann-Scheffe theorems and UMVUE
- Generalization of the Cramer-Rao inequality
- Evaluation of conditional expectations
- Unbiased estimation under incompleteness

- Consistent estimators
- Comparison of estimators using the decision-theoretic approach (loss function, risk function)

❖ Statistical Inference:

- Parametric, nonparametric, and semiparametric inference

❖ Approximate and Computationally Intensive Methods:

- The general problem of inference
- Likelihood functions
- Maximum likelihood estimation
- Optimization techniques (Newton-type methods)
- EM algorithm (simple form, properties, uses in analyzing missing data, fitting mixture models and latent variable model)
- Restricted maximum likelihood (REML) method of estimation
- Multi-stage maximization
- Efficient maximization via profile likelihood
- Confidence interval and testing hypothesis in complex cases
- Bayesian method of inference (prior and posterior distribution, different types of prior, credible intervals, testing hypothesis)
- Analytical approximations (asymptotic theory, Laplace approximation)
- Numerical integral methods (Newton-Cotes type methods)
- Monte Carlo methods
- Simulation methods (Markov chain Monte Carlo)

❖ Resampling Techniques:

- Bootstrap (confidence intervals, tests, parametric bootstrap)
- Jackknife (confidence interval, test, and permutation test)

❖ Nonparametric Inference and Robustness:

- Introduction
- Inference concerning cumulative distribution function (CDF)
- Quantiles and statistical functionals
- Empirical CDF
- Estimating statistical functionals
- Influence functions
- Testing statistical hypothesis in one sample and two or more sample settings
- Tolerance limit
- Empirical density estimation (histograms, kernel, kernel density estimation)

Multivariate Analysis

❖ Preliminaries of Multivariate Analysis:

- Applications of multivariate techniques
- Organization of data
- Data display and pictorial representations
- Distance

❖ Random Vectors and Random Sampling:

- Basics of matrix and vector algebra
- Positive-definite matrices
- Square-root matrix
- Random vectors and matrices
- Mean vectors and covariance matrices
- Matrix inequalities and maximization
- Geometry of the sample
- Random sample and expected values of sample means and covariance matrix
- Generalized variance
- Sample mean, covariance, and correlation as matrix operations
- Sample values of linear combinations of variables

❖ The Multivariate Normal Distribution:

- Multivariate normal density and its properties
- Sampling from a multivariate normal distribution and maximum likelihood estimation
- Sampling distribution and large sample behavior of sample mean vector and sample variance-covariance matrix
- Assessing the assumption of normality
- Detecting outliers and data cleaning
- Transformation to near normality

❖ Inferences about a Mean Vector:

- Plausibility of mean vector as a value for a normal population mean
- Hotelling T^2 and likelihood ratio tests
- Confidence regions and simultaneous comparisons of component means
- Large sample inference about a population mean vector
- Inferences about mean vectors when some observations are missing
- Time dependence in multivariate data

❖ Comparisons of Several Multivariate Means:

- Paired comparisons and a repeated measures design

- Comparing mean vectors from two populations
- Comparison of several multivariate population means (one-way MANOVA)
- Simultaneous confidence intervals for treatment effects
- Two-way multivariate analysis of variance
- Profiles analysis
- Repeated measures designs and growth curves

❖ Multivariate Linear Regression Models:

- The classical linear regression model
- Least squares estimation
- Inferences about the regression model
- Inferences from the estimated regression function
- Multivariate multiple regression
- Comparing two formulations of the regression model
- Multiple regression model with time-dependent errors

❖ Principal Components:

- Population principal components
- Summarizing sample variations by principal components
- Graphing the principal components
- Large sample inference

❖ Factor Analysis:

- Orthogonal factor models
- Methods of estimation (maximum likelihood estimates and principal factor analysis)
- Selection of loadings and factor
- Factor rotation
- Varimax rotation
- Quartimax rotation
- Oblimin rotations
- Factor scores
- Structural equations models

❖ Canonical Correlation Analysis:

- Canonical variates and canonical correlations
- Sample canonical variates and sample canonical correlations
- Large sample inference

❖ Discrimination and Classification:

- Separation and classification of two populations
- Classification of two multivariate normal populations

- Evaluating classification functions
- Fisher's discriminant function
- Classification with several populations
- Fisher's method for discriminating several populations

❖ Clustering:

- Similarity measures
- Hierarchical clustering methods
- Non-hierarchical clustering methods
- Fuzzy clustering
- Determination of the number of clusters
- Gap statistics and its modifications
- Cluster validity indices
- Cluster's homogeneity test
- Multidimensional scaling

Lifetime Data Analysis

❖ Basic Concepts and Models:

- Lifetime distributions
- Continuous models
- Discrete models
- General formulation
- Important models
- Exponential, Weibull, log-normal, log-logistic, gamma distributions
- Log-location-scale models
- Inverse Gaussian distribution models
- Mixture models
- Regression models

❖ Observation Schemes, Censoring, and Likelihood:

- Right censoring and maximum likelihood
- Other forms of incomplete data
- Truncation and selection effects
- Information and design issues

❖ Nonparametric and Graphical Procedures:

- Nonparametric estimation of survivor function and quantiles
- Descriptive and diagnostic plots
- Estimation of hazard or density functions
- Methods of truncated and interval-censored data
- Life tables

❖ Inference Procedures for Parametric Models:

- Inference procedures for exponential distributions
- Inference procedures for gamma distributions
- Inference procedures for inverse Gaussian distributions
- Inference for grouped, interval-censored, or truncated data
- Mixture models
- Threshold parameters
- Prediction intervals

❖ Inference Procedure for Log-Location-Scale Distributions:

- Inference for location-scale distributions
- Weibull and extreme-value distributions
- Log-normal and log-logistic distributions

- Comparison of distributions
- Models with additional shape parameters
- Planning experiments for life tests

❖ Parametric Regression Models:

- Introduction to log-location-scale regression models
- Proportional hazards regression models
- Graphical methods and model assessment
- Inference for log-location-scale models
- Extensions of log-location-scale models
- Hazard-based models

❖ Brief Introduction to Cox's Proportional Hazards Model:

- Partial likelihood function
- Estimation and interpretation of model parameters

Analysis of Time Series

❖ Introduction and Examples of Time Series:

- Time series plots
- Trend
- Seasonal effects
- Sample autocorrelation
- Correlogram
- Filtering

❖ Probability Models:

- Stochastic processes
- Stationarity
- Second-order stationarity
- White noise model
- Random walks
- Moving average (MA) processes
- Autoregressive (AR) processes
- ARMA processes
- Seasonal ARMA processes
- General linear process
- Properties, estimation, and model building
- Diagnostic checking

❖ Forecasting:

- Naive procedures
- Exponential smoothing
- Holt-Winters
- Box-Jenkins forecasting
- Linear prediction
- Forecasting from probability models

❖ Non-Stationary Time Series:

- Non-stationarity in variance
- Logarithmic and power transformations
- Non-stationarity in mean
- Deterministic trends
- Integrated time series
- ARIMA and seasonal ARIMA models
- Modeling seasonality and trend with ARIMA models

- ❖ Stationary Processes in the Frequency Domain:
  - Spectral density function
  - The periodogram
  - Spectral analysis
- ❖ Concept of State-Space Models:
  - Dynamic linear models
  - The Kalman filter

Statistical Modeling and Analysis

❖ Measures of Association for Quantitative Data:

- Correlation and inference concerning correlation
- Regression and model building
- Motivating examples
- Uses of regression

❖ Simple Linear Regression Model:

- Model for E(Y |x)
- Least squares estimation
- Assumptions related to errors
- Maximum likelihood estimation (MLE) of the model
- Sampling distribution of MLEs of the model parameters
- Inferences concerning the model parameters (confidence intervals and t-test)
- Confidence interval estimate of E(Y |x) (confidence band)
- Model accuracy and diagnostics
- Goodness of fit test (F-test, coefficient of determination, R^2)
- Prediction and prediction interval for a new Y at a specific x
- Residual analysis
- Detection and treatment of outliers
- Concept of lack of fit and pure error
- Test for lack of fit
- Transformations as a solution to problems with the model
- Weighted least squares
- Matrix representation of simple linear regression model, inference, and prediction

❖ Multiple Linear Regression Models:

- Formulation of multiple regression models
- Estimation of the model parameters
- Least squares estimation
- Maximum likelihood estimation
- Sampling distributions of the MLEs
- Confidence interval and hypothesis testing concerning model parameters
- Model accuracy and diagnostics
- Goodness of fit test (F-test, R^2, adjusted R^2)
- Prediction of a new observation
- The extra sum of squares principles and its application in testing the general linear hypothesis
- Checking all assumptions concerning the model and use of remedy measures when assumptions are not valid

- Detection and treatment of outliers and influential observations

❖ Polynomial Regression Model:

- Introduction
- Polynomial models in one variable
- Basic principles
- Piecewise polynomial fitting
- Polynomial models in two or more variables
- Orthogonal polynomials

❖ Indicator Variables:

- The general concept of an indicator variable
- Use of indicator variables in linear regression
- Models with only indicator variables
- Regression models with an indicator response variable

❖ Variable Selection and Model Building:

- Model building problem
- Consequences of model misspecification
- Criteria for evaluating subset regression models
- Computational techniques for variable selection

❖ Validation of Regression Models:

- Concept of cross-validation

❖ Generalized Linear Model:

- Exponential family of distributions
- Estimation methods
- Method of maximum likelihood
- Method of least squares
- Inference
- Sampling distribution for scores
- Sampling distribution for maximum likelihood estimators
- Confidence intervals for model parameters
- Adequacy of a model
- Sampling distribution for log-likelihood statistic
- Log-likelihood ratio statistic (deviance)
- Assessing goodness of fit
- Hypothesis testing
- Multiple regression
- Maximum likelihood estimation

- Log-likelihood ratio statistic

❖ Models for Binary Responses:

- Probability distributions
- Generalized linear models
- Dose-response models
- General logistic regression
- Maximum likelihood estimation and log-likelihood ratio statistic
- Other criteria for goodness of fit
- Least square methods
- Multinomial distributions
- Nominal logistic regression models
- Ordinal logistic regression models

❖ Models for Count Data, Poisson Regression, and Log-linear Models:

- Probability distributions
- Maximum likelihood estimation
- Hypothesis testing and goodness of fit

Demography

❖ Basic Concepts of Demography:

- Role and importance of demographic/population studies
- Sources of demographic data: census, vital registration system, sample surveys, population registers, and other sources, especially in Bangladesh

❖ Errors in Demographic Data:

- Types of errors
- Methods of testing the accuracy of demographic data
- Quality checking and adjustment of population data
- Post Enumeration Check (PEC) and detection of errors and deficiencies in data and the needed adjustments and corrections

❖ Fertility:

- Basic measures of fertility
- Crude birth rate
- Age-specific fertility rates (ASFR)
- General fertility rate (GFR)
- Total fertility rate (TFR)
- Gross reproduction rate (GRR) and net reproduction rate (NRR)
- Child-woman ratio
- Concept of fecundity and its relationship with fertility

❖ Demographic Theory:

- Transition theory and the present situation in Bangladesh
- Malthus' theory and its criticism

❖ Mortality:

- Basic measures of mortality
- Crude death rate (CDR)
- Age-specific death rates (ASDR)
- Infant mortality rate
- Child mortality rate
- Neonatal mortality rate
- Standardized death rate, its need, and use
- Direct and indirect standardization of rates
- Commonly used ratios: sex ratio, child-woman ratio, dependency ratio, density of population

❖ Fertility and Mortality in Bangladesh Since 1951:

- Reduction in fertility and mortality in Bangladesh in recent years
- Role of socio-economic Development on Fertility and mortality

❖ Nuptiality:

- Marriage
- Types of marriage
- Age of marriage
- Age at marriage and its effect on fertility
- Celibacy, widowhood, divorce, and separation, their effect on fertility and population growth

❖ Migration:

- Definition, internal and international migration
- Sources of migration data
- Factors affecting both internal and international migration
- Laws of migration
- Impact of migration on origin and destination
- Effect on population growth, age and sex structure, labor supply, employment and unemployment, wage levels, and other socio-economic effects
- Migration of Bangladeshis abroad and its impact on the overall economic development of the country

❖ Graduation of Data:

- Meaning and its need
- Techniques of graduation
- Graduation of age distribution
- Life table: concept, structure, and calculation
- Complete life table (life table by single year of age) and abridged life table
- Multiple decrement life tables
- Working life table
- Different life table functions and inter-relationships among them
- Use of life table
- Model life tables
- Coale and Demeny regional model life tables

❖ Force of Mortality:

- Idea and definition
- Calculation of life table with the help of the force of mortality

❖ Population Growth:

- Techniques to measure it
- Doubling time concept in demography
- Population estimates and projections
- Different techniques of population projection
- Need for population projections
- Use of Lee-Carter model in population projections

❖ Stable and Stationary Population:

- Characteristics and uses
- Lotka’s characteristics equation
- Intrinsic birth and death rates
- Effect of a uniform drop in the force of mortality on the growth rate
- Effects of changes in fertility and mortality on the age distribution of the population

❖ Population in Bangladesh:

- History of the growth of the population in Bangladesh
- Implications of the growth of population in Bangladesh
- Population policy in Bangladesh
- Level, trends, and determinants in fertility, mortality, and migration in Bangladesh
- Interrelationship between population and development
- Prospects of population and population control in Bangladesh
- Aged and aging of the population in Bangladesh

Epidemiological Statistics

❖ Introduction:

- Disease processes
- Statistical approaches to epidemiological data
- Study design
- Binary outcome data
- Causality

❖ Measures of Disease Occurrence:

- Prevalence and incidence
- Disease rates
- Hazard function
- Review of simple random samples
- Probability, conditional probabilities, and independence of two events

❖ Measures of Disease-Exposure Association:

- Relative risk
- Odds ratio
- Relative hazard
- Risk attributable risk

❖ Study Designs:

- Population-based studies
- Cohort studies
- Case-control studies
- Case-cohort studies
- Assessing the significance of 2 × 2 tables obtained from cohort designs, case-control designs

❖ Estimation and Inference for Measures of Association:

- Odds ratio
- Sampling distribution
- Confidence interval for odds ratio
- Relative risk
- Excess risk
- Attributable risk

❖ Confounding and Interaction:

- Causal inference
- Counterfactuals
- Confounding variables
- Control of confounding variables by stratification
- Causal graphs
- Controlling confounding in causal graphs
- Cochran-Mantel-Haenszel test
- Summary estimates and confidence intervals for odds ratio and relative risk after adjusting for confounding factors

❖ Interaction:

- Multiplicative and additive interaction
- Interaction and counterfactuals
- Test of consistency of association across strata
- Overall test of association
- Test for the trend in risk

❖ Introduction to Matching:

- Types of matching
- Analysis of matched studies

Biostatistics and Statistical Methods

❖ Definition, Scope, and Importance:

- Definition and scope of statistics and biostatistics
- Uses and importance of statistics and biostatistics
- Limitations of using statistics and biostatistics
- Population, sample, parameter, statistic, estimator, and estimate

❖ Data Presentation:

- Meaning of data
- Level of measurement
- Variables (discrete and continuous)
- Summarizing and presenting data
- Tabular, graphical, and diagrammatic representation of data

❖ Descriptive Statistics:

- Measures of central tendency (Arithmetic mean, geometric mean, harmonic mean, median, and mode)
- Measures of dispersion (Range, mean deviation, variance, coefficient of variation, standard deviation)
- Moments, skewness, and kurtosis

❖ Probability Distribution:

- The normal, binomial, and Poisson distribution
- Derivation, means, and variances
- Point estimation
- The Mean and its Standard Error
- The Central Limit Theorem
- Confidence Intervals (Mean of a sample, Binomial proportions)

❖ Hypothesis Testing:

- General concepts
- The basic idea of significance tests
- Simple significance tests based on the normal distribution
- Comparison with a known standard
- Comparison of means of two large samples

❖ t-Tests:

- The use of 't' tests for small samples

- Importance of small sample comparison
- Comparison of sample means with a standard
- Comparison of means of two small samples (unknown variances-assumed equal, not assumed equal)
- Confidence limits
- One-sample t-test, Paired & unpaired t-tests
- Nonparametric analogs

❖ Contingency Tables:

- Comparing binomial proportions
- X^2-tests of goodness of fit and homogeneity
- Introduction to the general idea
- Testing the fit of a whole frequency distribution to data
- Tests of homogeneity
- Variance ratio test

❖ Correlation Coefficients:

- Properties of ρ (rho)
- Pearson's correlation coefficient
- Spearman's ρ (rho)
- Kendall's τ (tau)
- Types of Correlation (Bivariate Correlations, Partial Correlations, Distances)

❖ Linear Regression:

- Introduction to Linear Regression
- Purposes and assumptions for using the linear regression model
- Least Squares Method
- R^2 - Variance explained

❖ Analysis of Variance (ANOVA):

- One-way ANOVA
- Two-way ANOVA

❖ Power, Sample Size, and Effect Size:

- Practical vs. statistical significance.