[Term-I Final Examination, 2026]

Khulna University

School: SET **Discipline:** Statistics **Program:** Master of Science
Year: MS 1st **Term: I** **Examination Year:** 2026 **Session:** 2025-2026
Course Code: 0542 20 Stat 5101 **Course Title: Advanced Statistical Inference**
Total Marks: 60 **Time:** 03 Hours **Credit:** 03
Instructions:

- The figures in the margin indicate full marks. The questions are of equal value.
- Use separate sheet for each section.

Section-A

There are **FOUR** questions in this section. Answer any **THREE** questions.

1. a) Define the following terms: (i) Prior distribution (ii) Posterior distribution (iii) Risk function and (iv) Admissible estimator. [L1; CLO1] 04
   b) Define Bayes estimator. State the properties of Bayes estimator. [L1; CLO2] 03
   c) How can you derive the posterior distribution of a random variable? [L2, L6; CLO2] 03

2. a) What do you mean by Bayes risk, Native estimator and Refined estimator? Explain with suitable examples. [L1, L2; CLO4] 04
   b) Let, $X_1, X_2, \dots, X_n$ be a random sample from $b(1, \theta)$ using a s squared error loss function and a Beta prior of the parameter θ. Find the Bayes estimator, minimax estimator, and its risk. [L4; CLO3] 06

3. a) "If an estimator δ has a constant risk function $R(\theta, \delta) = c, \forall\, \theta$, and δ is admissible, then δ is the minimax estimator." Justify the statement. [L2; CLO2] 04
   b) Define two sample problems in Bootstrapping. Obtain the bootstrap sampling error of the problem. [L5; CLO2] 04
   c) "Bias can be estimated using Jackknife method." Do you agree or disagree with the statement? Give reason to support your answer. [L2; CLO4] 02

4. a) Define a kernel function. Discuss its applications in statistics. [L1, L2; CLO3] 03
   b) Define resampling with an example. Is there any difference between resampling and bootstrapping? Explain. [L1, L4; CLO3] 03
   c) Let $X_1, X_2, .., X_n$ be a random sample of size n from an exponential distribution with mean θ. Find Pitman estimator for the parameter θ. [L41, L5; CLO2] 04

Section-B

There are **FOUR** questions in this section. Answer any **THREE** questions.

5. a) Show that the best critical region is a function of sufficient statistics. [L4; CLO1] 03
   b) Construct a locally uniformly most powerful unbiased test from the generalized lemma. [L6; CLO2] 05
   c) Distinguish between 'Locally Uniformly Most Powerful Unbiased test' and 'Uniformly Most Powerful Unbiased test'. [L2; CLO1] 02

6. a) Show that, for unbiasedness, the power of the test should be greater than or equal to the size of the test. [L3; CLO1] 04

   b) A single observation is made of the density function $f(x,\theta) = \frac{1}{\theta}e^{-\frac{x}{\theta}}; x > 0$ *and* the critical region is defined by $x > 2$. Find the α, β and $\beta(w|H_1)$. [L3; CLO1] 03

   c) Find the relationship between the similar region test and the sufficient statistic. [L4; CLO2] 03

7. a) Distinguish between Sobel test and Armitage method for testing three hypotheses. [L4; CLO2] 03

   b) How Wald's equation can be utilized in getting ASN of SPRT? [L2; CLO1] 02

   c) Let X a sample of size 1. And the test rejects II, if :$4.2 \leq x \leq \infty$), where Ho: θ =3.2 vs. H_1: θ = 4.0. Find (i) the power function; (ii) the level of significance; (iii) draw a power curve for different values of θ. [L1, L3; CLO2] 05

8. a) Describe Wald's Sequential Probability Ratio Test. 03

   b) Explain how the sequential test procedure differs from the Neyman-Pearson test procedure. 03

   c) A random variable follows the normal distribution $N(\theta, \sigma^2)$. where σ^2 is known. Derive the SPRT for testing $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$. Obtain the approximate expression for the OC function. 04

The End

[Question of Term Final Examination, 2026]

Khulna University

School: SET **Discipline:** Statistics **Program:** Master of Science (MS)
Year: MS First **Term: I** **Examination Year:** 2026 **Session:** 2025-2026
Course Code: 0542 20 Stat 5103 **Course Title:** Applied Multivariate Analysis
Total Marks: 60 **Time:** 03 Hours **Credit:** 03
Instructions:

- The figures in the margin indicate full marks. The questions are of equal value.
- Use a separate sheet for each section.

Section-A

There are **FOUR** questions in this section. Answer any **THREE** questions.

1. a) Define a multivariate normal distribution with p random variables. Explain the derivation of the multivariate normal distribution. [L1, L4; CLO1,2] 05

   b) Mention at least two properties of the multivariate normal distribution. For X distributed as $N_3(\mu, \Sigma)$ find the distribution of $\begin{bmatrix} X_1 - X_2 \\ X_2 - X_3 \end{bmatrix}$. [L3, L4; CLO1] 05

2. a) How can the multivariate mean vector of p random variables be tested? Derive the appropriate test statistic for testing the multivariate mean. [L1, L4; CLO3] 03

   b) Evaluate the Hotelling T^2 value of the data matrix $X' = \begin{bmatrix} 6 & 10 & 8 \\ 9 & 6 & 3 \end{bmatrix}$ for testing null hypothesis $H_0: \mu_0' = [9 \quad 5]$. $(F_{2,1} = 198.5)$. [L5; CLO3] 04

   c) How is the likelihood ratio test related to Hotelling's T^2 statistic? Establish the relationship between them. [L2, L4; CLO3] 03

3. a) How does confidence region be defined for multivariate case. Differetiate it from simultaneous confidence intervals. [L1, L5; CLO2,3] 04

   b) Consider two samples of size n_1 and n_2 form multivariate normal population $N_p(\mu_1, \Sigma_1)$ and $N_p(\mu_2, \Sigma_2)$ respectively. Then test the null hypothesis $H_0: \mu_1 = \mu_2$, assuming that the population covariance matrices are equal. [L3, L4; CLO2,3] 06

4. a) Describe a repeated measures design. [L2; CLO3] 02

   b) How does it differ from a paired comparison design? [L4; CLO3] 02

   c) Explain the univariate and multivariate approaches to analyzing repeated measures data. [L4; CLO3] 04

   d) Explain a repeated measures design with two within-subject factors. [L4; CLO3] 02

Section-B

There are **FOUR** questions in this section. Answer any **THREE** questions.

5. a) Define multivariate regression analysis with examples. [L1, L2; CLO2,4] 02

   b) Construct a multivariate regression model and estimate its parameters using the maximum likelihood method. [L4, L5; CLO4] 04

*Levels of Bloom's cognitive domain: [L1=Remember; L2=Understand; L3=Apply; L4=Analyze; L5=Evaluate; L6= Create]

c) How does the maximum likelihood method differ or not from least square estimate? [L4, L5; CLO4] 04

6. a) Distinguish between Factorial MANOVA and MANCOVA. [L2; CLO4] 02

b) Explain how to assess the significance of MANOVA effects using the multivariate test statistic. [L2; CLO4] 04

c) A researcher collected performance scores on three skills for students from three different training programs. Outline the steps for conducting MANOVA on this dataset. [L2; CLO4] 04

7. a) Define correspondence analysis (CA) with an example. What are the main purposes of correspondence analysis (CA)? [L1, L5; CLO3] 03

b) Explain the key steps in performing Correspondence Analysis on a contingency table. [L2, L4; CLO2,4] 05

c) How does Multiple Correspondence Analysis (MCA) generalize simple Correspondence Analysis (CA) for the analysis of more than two categorical variables? [L3, L4; CLO5] 02

8. a) What is path analysis? Explain its main assumptions and how it differs from multiple regression. [L1, L5; CLO5] 03

b) Define direct effect, indirect effect, and total effect in path analysis. [L2, L3; CLO6] 03

c) Describe the main components of a path diagram. Explain the meaning of single-headed arrows, double-headed arrows, and error terms. [L4; CLO6] 04

The End

[Question of Term Final Examination, 2026]

Khulna University

School: SET **Discipline:** Statistics **Program:** Master of Science (MS)
Year: MS First **Term: I** **Examination Year:** 2026 **Session:** 2025-2026
Course Code: 0542 20 Stat 5115 **Course Title:** Bioinformatics
Total Marks: 60 **Time:** 03 Hours **Credit:** 03
Instructions:

- The figures in the margin indicate full marks. The questions are of equal value.
- Use a separate sheet for each section.

Section-A

There are **FOUR** questions in this section. Answer any **THREE** questions.

1. a) Define Bioinformatics. Discuss the importance and scope of Bioinformatics with its connectivity to Statistics. [L1, L2; CLO1] 05

   b) Explain the Central Dogma of molecular biology with a clear schematic description of transcription and translation. [L2; CLO1] 05

2. a) Define genomics and discuss the classification of genomics [L1, L2; CLO2] 04
   b) Describe genes, chromosomes, nucleic acids, and genotype. [L1, L2; CLO2] 04
   c) Differentiate between DNA and RNA. [L2; CLO1] 02

3. a) What is sequence alignment? Distinguish between global and local sequence alignment. [L2; CLO31] 04
   b) Explain the dynamic programming algorithm for sequence alignment. Mention its importance in Bioinformatics. [L2; CLO3] 06

4. a) Describe the heuristic working principles of BLAST and FASTA. How does BLAST utilize the E-value (Expectation value) to determine the statistical significance of a sequence match? [L1, L5; CLO3] 04
   b) Explain genotype and describe major genotyping technologies used in molecular genetics. [L2; CLO3] 02
   c) Given a scenario where SNP data is used for disease association studies, explain how bioinformatics contributes to identifying genetic variation. [L1, L3; CLO3] 04

Section-B

There are **FOUR** questions in this section. Answer any **THREE** questions.

5. a) Define phylogenetic analysis and discuss its motivation. [L1, L6; CLO2] 04
   b) Describe the Maximum Likelihood method in phylogenetic analysis. State its advantages and disadvantages. [L2; CLO3] 06

6. a) Define gene-expression data. Discuss the different types of microarray gene-expression data. [L2; CLO3] 05
   b) Why is Normalization a critical step in preprocessing microarray data? Briefly compare two methods used to remove systematic bias. [L2, L4; CLO3] 05

*Levels of Bloom's cognitive domain: [L1=Remember; L2=Understand; L3=Apply; L4=Analyze; L5=Evaluate; L6= Create]

7. a) What are differentially expressed (DE) genes? Explain how statistical tests are used to identify DE genes between two or more groups. [L1, L6; CLO4] 04

   b) What is the primary goal of Gene Set Enrichment Analysis (GSEA), and how does it differ from simply looking at a list of individual DE genes? [L1, L2; CLO4] 03

   c) Differentiate between Clustering and Classification in the context of gene expression profiles. Provide one biological example for each. [L2; CLO5] 03

8. a) Define Linkage Disequilibrium (LD) and explain its importance in "tagging" SNPs for GWAS. How does Haploview facilitate this analysis? [L1, L3; CLO5] 04

   b) Why are Linear Mixed Models (LMMs) preferred over simple linear regression in GWAS when dealing with population stratification? [L4; CLO5] 03

   c) Explain how Contingency Tables are utilized in SNP analysis to determine the association between a genotype and a phenotypic trait. [L2; CLO5] 03

The End

[Term-I Final Examination, 2026]

Khulna University

School: SET **Discipline:** Statistics **Program:** Master of Science (M.Sc)
Year: First **Term: I** **Examination Year:** 2026 **Session:** 2025-2026
Course Code: 0542 20 Stat 5121 **Course Title: Applied Multilevel Analysis**
Total Marks: 60 **Time:** 03 Hours **Credit:** 03
Instructions:

- The figures in the margin indicate full marks. The questions are of equal value.
- Use separate sheet for each section.

Section-A

There are **FOUR** questions in this section. Answer any **THREE** questions.

1. a) Define multilevel analysis. Why is multilevel modeling necessary for hierarchical data structures? [L1; CLO1] 03
   b) Why is multilevel analysis preferred over classical regression for hierarchical data? [L1; CLO1] 02
   c) Define Intra-class Correlation Coefficient (ICC). Explain its importance in multilevel analysis. [L3, L4; CLO1] 05

   Suppose the estimated variances from a null model are:
   $\sigma_{u0}^2=0.45$, $\sigma_e^2=1.05$. Calculate ICC and interpret the result regarding the necessity of MLM.

2. a) What is meant by fixed effects and random effects in MLM? [L1; CLO2] 03
   b) Define the variance component model and random intercept model. [L1; CLO2] 03
   c) What is meant by cross-level interaction in MLM? Explain the usefulness of cross-level interaction models. [L2; CLO2] 04

3. a) Define Log-likelihood and Deviance (-2 Log-likelihood). [L2; CLO2] 03
   b) What are AIC and BIC? Why are they used in model selection? [L1; CLO2] 04
   c) Describe the likelihood ratio (LR) test procedure for testing level-2 variance. [L5; CLO2] 03

4. a) Construct a three-level multilevel model for students within classrooms within schools. [L6; CLO3] 03
   b) Discuss the Maximum Likelihood (ML) and Restricted Maximum Likelihood (REML) estimation procedures in multilevel modeling. [L6; CLO4] 04
   c) Discuss the advantages and limitations of GLS estimation in MLM. [L2; CLO3] 03

Section-B

There are **FOUR** questions in this section. Answer any **THREE** questions.

5. a) Explain the concept of multilevel generalized linear models with examples. [L2; CLO1] 03
   b) List down the components of multilevel generalized linear models. [L1; CLO1] 03
   c) How to determine the sample size in multilevel analysis? Discuss with a proper example. [L1; CLO3] 04

6. a) Illustrate the mathematical form of a two-level binary logistic regression model. [L2; CLO3] 02
   b) Define Variance Partition Coefficients (VPCs) and Intra-Class Correlations (ICCs) for a binary response outcome. Are there any differences? [L2; CLO3] 04

*Levels of Bloom's cognitive domain: [L1=Remember; L2=Understand; L3=Apply; L4=Analyze; L5=Evaluate; L6= Create]

c) A study compares a random intercept model and a random slope model for a predictor variable. The log-likelihood values are: 04

- Model without random slope (M1) = –1945.20
- Model with random slope (M2) = –1940.10

(i) Calculate the LR statistic for testing the random slope effect.
(ii) Interpret the result if the critical value of χ^2 at 5% significance level with 2 degrees of freedom is 5.99. [L4; CLO3]

7. a) Calculate the value of explained variation (PCV) and ICC from the following table and comment on random part: [L5; CLO2] 05

Table: Multilevel model's output (random) identifying individual-level and community-level factors associated with improved water utilization in Bangladesh

Measures of Variation	**Model 0 (Null)**	**Model 1 (Individual-level variables)**	**Model 2 (Community-level variables)**	**Model 3 Both-level sig. variables)**
Variance (SE)	3.10 (0.30)	2.45 (0.26)	2.20 (0.24)	1.75 (0.22)
Explained Variation (%) (PCV)	Ref			
ICC = VPC				
Chi-square (χ^2)	820.45	610.32	545.10	498.76
P-value	(0.000)	(0.000)	(0.000)	(0.000)

SE=Standard Error

b) Explain the ever-changing latent scale in term of a multilevel regression model. [L2; CLO2] 02

c) Discuss the Poisson model for count data with examples. [L6; CLO3] 03

8. a) Discuss the significance and calculation process of the Intraclass Correlation Coefficient and Percentage of Proportional change in Variation applied in a multilevel generalized linear model. [L6; CLO2] 03

b) Distinguish between fixed and varying occasions of longitudinal data with proper examples. [L4; CLO5] 03

c) Mention some advantages and disadvantages of analyzing longitudinal data in MLM? [L1; CLO5] 02

d) Interpret the following output of testing randomness for under-five child mortality in Bangladesh: [L5; CLO4] 02

Individual-level Factors		**Model without random slope (M_1)**	**Model with random slope (M_2)**	**Likelihood Ratio (LR) Statistic $2(M_2 - M_1)$**	$\chi^2_{2,0.05}$
Age at First Birth	Below 18 vs. 18 and above	-11101.69	-11098.22	6.94	0.031
Mother's Education	No Education vs. At least Primary		-11091.32	20.74	<0.001
Working Status Of Mother	Employed vs. Unemployed		-11086.69	30.01	<0.001
Wealth Status	Poor vs. Non-poor		-11096.83	9.71	0.007
Media Exposure	No vs. Yes		-11090.25	22.89	<0.001
Total Children	At least 2 vs. 1		-11100.48	2.41	0.299

Note: M_1 and M_2 are log likelihood values,

The End

[Question of **Term Final Examination, 2026**]

Khulna University

School: SET **Discipline:** Statistics **Program:** Master of Science (MS)
Year: MS First **Term: I** **Examination Year:** 2026 **Session:** 2025-2026
Course Code: 0314 20 Soc 5151 **Course Title:** Population and Social Development
Total Marks: 60 **Time:** 03 Hours **Credit:** 03
Instructions:

- The figures in the margin indicate full marks. The questions are of equal value.
- Use a separate sheet for each section.

Section-A

There are **FOUR** questions in this section. Answer any **THREE** questions.

1. a) Write down the importance of studying population and social development. [L1; CLO1, 2] 03
   b) Define the Gender Based Domestic Violence (GBDV). Discuss the Impact of GBDV on reproductive health outcomes. [L2; CLO2] 03
   c) Distinguish between direct, structural, and cultural violence in Galtung's framework. [L3, L4; CLO1,4] 04

2. a) Analyze the structural transition from MDGs to SDGs, explaining how "MDG extensions" reshaped global development targets. [L2; CLO2] 04
   b) Identify the core SDG goals, targets, and tracking indicators that directly address maternal health, gender equality, and social protection. [L5; CLO2] 04
   c) Discuss the institutional and data-capacity challenges faced globally when tracking high-dimensional, cross-cutting SDG targets. [L2; CLO3] 02

3. a) Define population ageing. Explain the importance of studying population ageing in development planning. [L1, L5; CLO2,3] 03
   b) Establish the relationship between the Old Age Dependency Ratio (OADR) and the Potential Support Ratio (PSR), and explain its importance. [L3, L4; CLO4,5] 03
   c) In a population, the number of elderly (60+) is 800, and children (0–14) is 4000. 04

   (i) Compute the Ageing Index (AI) of the population.

   (ii) If the elderly population increases by 10%, holding the child population constant, evaluate the resulting change in the Ageing Index (AI) and discuss its demographic implications. [L4; CLO4]

4. a) What do you mean by gender preference? How do these preferences shape family size, gender dynamics, and policy responses in Bangladesh? [L1, L5; CLO3] 04
   b) How does child mortality affect the Parity Progression Ratio (PPR)? If the PPR from parity 2 to 3 is 0.6, interpret this in terms of fertility behavior. [L4, L5; CLO3,5] 03
   c) Explain the relationship between actual fertility and ideal fertility. [L1, L2; CLO3] 03

Section-B

There are **FOUR** questions in this section. Answer any **THREE** questions.

5. a) Define a megacity and evaluate the demographic projection methods and data challenges used to estimate its population growth. [L1; CLO4] 04

   b) Analyze the social, economic, and demographic implications of rapid urbanization on a nation's capacity to meet basic needs. [L4; CLO4] 04

   c) Explain how rapid rural-to-urban migration alters the age structure and dependency ratios of both sending and receiving areas. [L2; CLO4] 02

6. a) State the core assumptions of the Gompertz model and outline the step-by-step estimation process for its parameters. [L1; CLO5] 04

   b) Provide a formal mathematical derivation of the Gompertz model parameters from its fundamental differential equation. [L4; CLO5] 06

7. a) Define population momentum and explain the critical demographic factors policymakers must consider to reduce it successfully. [L1, L2; CLO4] 04

   b) Explain the tempo effect and the long-term demographic implications for a society if replacement-level fertility is never achieved. [L2; CLO4] 04

   c) Contrast the major advantages and disadvantages of using the Gompertz model over the Logistic growth model in demographic forecasting. [L4; CLO5] 02

8. a) Discuss the principles underlying DALY measurement. [L1, L5; CLO6] 03

   b) Discuss problems in constructing a disability life table. [L2, L3; CLO1] 04

   c) A public health analyst is evaluating the burden of a chronic disease. The disability weight is 0.25, the discount rate is 0.03, and the life expectancy at age 30 is 40 years. Compute the DALY for this group. [L4; CLO5] 03

The End