

Department of Statistics
 Jahangirnagar University
 Part II B. Sc (Hons.) Examination-2017
 Course Code: Stat-201
 Course Title: Probability Distribution

Time: 2.5 Hours

Full Marks: 35

Answer Any Three of the Following Questions. All Questions Carry Equal Marks.

1. a) Find the first moments of hypergeometric distribution and hence find its mean and variance.

b) The Grameenphone Company has 14 employees who hold managerial position in an area. Of these 8 are female and 6 are male. The company is planning to send 3 managers to a conference randomly. Find the following probabilities using the probability rule and use the appropriate distribution to calculate the probabilities.

- Find the probability that all 3 of them are male.
- Find the probability that at most one of them is a female.

c) What is negative hypergeometric distribution? Explain the situations where this distribution is used. A bag contains 4000 white balls and 5000 red balls. A sample of 4 balls is taken such that all balls have got the same probability of being drawn.

- Find the probability that exactly 4 will be red.
- How many draws are necessary to get exactly 4 white balls?

2. a) Explain the concept of negative binomial distribution. Describe why it is called negative binomial distribution.

b) Show that negative binomial random variable is a sum of geometric random variable.

c) Find the mgf of negative binomial distribution and hence find its mean and variance.

d) Fit a negative binomial distribution to the following frequency distribution.

No of Calls Received	0	1	2	3	4	5
No of students	50	80	55	35	25	2

e) Find the probability that a person flipping a coin gets (i) the third head on the seventh flip; (ii) the first head on the fourth flip.

3. a) Define a gamma distribution with parameter α . Find the limiting form of gamma distribution as $\alpha \rightarrow \infty$.

b) If X and Y are two independent gamma variates with parameters α_1 and α_2 respectively, show that $Z_1 = X + Y$ and $Z_2 = \frac{X}{Y}$ are independently distributed.

c) What are the distributions of Z_1 and Z_2 ?

4. a) What is Cauchy distribution? Indicate its similarities and dissimilarities with the normal distribution.

b) Find the cumulative distribution function of Cauchy distribution. Show that the mgf of this distribution does not exist. Does its characteristic function exist?

c) The IQ of 6 years old is assumed to be a normal variable. If it is known that 15% of the children IQs under 90 and 2% exceed 135, what percentage of the children have IQs between 100 and 120?

5. a) State the general form of exponential distribution. What is its standard form? Write down one-parameter exponential distribution.

b) Find the mean, mode and median. State when mean is equal to variance.

c) In a certain city, the daily consumption of water (in millions of liters) follows approximately a gamma distribution with $\alpha = 2$ and $\beta = 3$. If the daily capacity of that city is 9 million liters of water, what is the probability that on any given day the water supply is inadequate?

- Find the mean and variance of the daily water consumption.
- According to Chebyshev's theorem, the probability of at least $3/4$ of the water consumption on any given day will fall within what interval?

Department of Statistics
Jahangirnagar University
Part II B. Sc. (Hon's) Examination - 2017
Course Name: Sampling Distribution
Course No.: Stat-202

Full Marks 70

Nihon

Time: 4 Hours

Answer any FIVE from the following Questions. Each question carries equal Marks.

Q1. a) Explain the concept of sampling distribution with example. How does it differ from the concept of frequency distribution and probability distribution?
b) What is the importance of sampling distribution in statistics? What are the uses of sampling distribution?
c) Obtain the sampling distribution of the linear correlation coefficient r when $\rho = 0$.

Q2. a) State and prove the Inversion Theorem.
b) How do you understand that the random variable is discrete or continuous.
c) Obtain the density function for which characteristic function is $\phi(t) = (q + pe^{it})^n$.

Q3. a) Explain convergence of a sequence of random variables. Discuss weak law of large numbers and strong law of large numbers.
b) What are the necessary and sufficient conditions for the existence of large numbers?
c) Differentiate between weak law of large numbers (WLLN) and strong law of large numbers (SLLN).
d) State and prove Chebyshev's weak law of large numbers.

Q4. a) State central limit theorem. Write down the important versions of central limit theorem.
b) In what situations central limit theorem can be used? What are the importance of central limit theorem. Mention its application.
c) State and prove De Movire-Laplace theorem.
d) Compare central limit theorem and law of large number.

Q5. a) What do you mean by Fisher's t -distribution? Also derive Fisher's t -distribution.
b) What are the assumptions for student's t test? Show that mode and median of student's t -distribution is equal to zero.
c) Prove that student's t distribution is symmetric. Also show that r is a sample correlation coefficient, then $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ follows t -distribution with $(n-2)$ d.f.

(Q6.)

- a) Explain the concept of central F -distribution with its properties.
- b) Suppose F follows F distribution with degrees of freedom n_1 and n_2 . Obtain the distribution of $\frac{n_1}{n_2} F$.
- c) If $n_1 = 1$ and $n_2 = n$ are substituted in the F distribution with degrees of freedom n_1 and n_2 then show that the distribution of $F = t^2$ converges to t distribution with n degrees of freedom.
- d) If $n_1 = 1$ and $n_2 = n$ are substituted in the F distribution with degrees of freedom n_1 and n_2 then show that the distribution of F converges to standard normal distribution.

(Q7.)

- a) What is meant by non-central t - distribution? Also derive the non-central t -distribution.
- b) What are the properties of non-central t distribution?
- c) Establish the relationship between central t distribution and non-central t distribution.
- d) Find the odd and even moments of non-central t distribution and try to obtain the shape characteristics of the non-central distribution.

(Q8.)

- a) What is non-central F distribution and derive its distribution.
- b) Establish the difference between central F distribution and non-central F distribution.
- c) State non-central χ^2 distribution and what are its properties?
- d) Using inversion theorem derives non-central χ^2 distribution.

-----Good Luck-----

Time: 4 Hours

Nihon

Full Marks: 70

Answer Any Five of the Following Questions. All Questions Carry Equal Marks.

1. a) What are the assumptions of classical linear regression model? Write down the applications of regression.
 b) Estimate the parameters of two-variable regression by method of maximum likelihood.
 c) Test the significance of the regression coefficients. Also, obtain a $100(1-\alpha)\%$ confidence interval for the slope coefficient of the two-variable regression.
2. a) What do you mean by correlation analysis? Write down the properties of simple correlation coefficient.
 b) Show that the coefficient of correlation between two variables is independent of origin and scale of measurement.
 c) Show that the coefficient of simple correlation lies between -1 and +1. Hence, show that the coefficient of variation lies between 0 and 1.
3. a) Explain the concept of coefficient of determination (R^2)? What is adjusted coefficient of determination (\bar{R}^2)? What is the justification of using it in a multiple regression model? Write down some important points with regard to \bar{R}^2 .
 b) Show that $\frac{\sum_{i=1}^n \hat{u}_i^2}{n-3}$ is an unbiased estimator of σ^2 in case of three-variable regression model.
 c) What is partial correlation coefficient? What do you mean by first order, second order and multiple order correlation coefficients? Explain $r_{12,3}$, $r_{13,2}$, $r_{23,1}$, $r_{12,34}$, $r_{12,345}$ and $r_{y,k(m)}$ with necessary formulas. Interpret the partial correlation coefficient, $r_{12,3}$. Establish a relationship among multiple coefficient of determination, simple coefficient of determination and partial coefficient of determination.
4. a) What is the purpose of Chow test? What are the underlying assumptions of Chow test? Write down the mechanics of Chow test with an example. What are the caveats about Chow test that must be kept in mind?
 b) How can you test one of the assumptions underlying the Chow test, namely, that the error variances in the two periods are the same? If the null hypothesis is rejected, then can you use Chow test?
 c) What are the advantages of dummy variable technique over the Chow test?
5. a) What do you mean by heteroscedasticity? Discuss some reasons of heteroscedasticity with examples.
 b) Discuss some informal and formal methods of detecting heteroscedasticity.
6. a) Define logistic regression. Explain the estimation procedure of the parameters in logistic regression.
 b) What is a dummy variable and dummy variable trap? Consider a model:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 D_{2i} + \varepsilon_i ; i = 1, 2, \dots, n$$
, where X_{1i} is the quantitative variable and D_{2i} is the dummy variable. Show the graphical representation of the model. If we include more dummy variables in the model then what will happen?

7. a) What do you mean by selecting the best regression equation? What methods are available for selecting the best regression equation? Discuss all possible regression procedure.

b) Explain the concept of ridge regression? Why is it intended in regression analysis? Briefly discuss the procedure of estimating the parameters of a ridge regression.

8. a) Define nonlinear regression model. Discuss how does it differ from linear regression? For the models shown below, determine whether they are linear, intrinsically linear or nonlinear?

i) $Y_i = \beta_1 X_{2i}^{\beta_2} X_{3i}^{\beta_3} e^{u_i}$

ii) $Y_i = \beta_1 + (0.75 - \beta_1) e^{-\beta_2(X_i-2)} + u_i$

iii) $Y_i = \theta_1 + \beta_1 X_{1i} + \beta_2 X_{2i}^{\beta_3} + \varepsilon_i$

iv) $\ln Y_i = \beta_1 - \beta_2 \left(\frac{1}{X_i} \right) + u_i$

v) $Y_i = \beta_1 + (0.75 - \beta_1) e^{-\beta_2(X_i-2)} + u_i$

b) Consider the following non-linear regression model:

$$Y_i = \beta_1 e^{\beta_2 X_i} + u_i$$

How can you estimate the parameters of the above model? Briefly discuss several approaches.

Department of Statistics
Jahangirnagar University
Part II B. Sc. (Hon's) Examination - 2017
Course Title: Demography
Course No.: Stat-204

Full Marks 35

Time: $2 \frac{1}{2}$ Hours

Niranjan

Answer any THREE from the following Questions. Each question carries equal Marks.

Q1. a) Discuss the scope and importance of demography.
b) What are the main sources of demographic data? Critically review these sources in terms of their scope and limitations.
c) What do you mean by age reporting error? Explain different methods for calculating age reporting error.

Q2. a) What do you mean by population growth? Explain geometric and exponential growth model.
b) What is meant by doubling time of a population? Derive an expression for computing the doubling time of a population.
c) Define CDR and CBR with advantages and disadvantages.

Q3. a) Distinguish between stable population and stationary population. Derive the Lotka's model developed for stable population theory.
b) Explain the calculation of mean length of generation and intrinsic death rate.
c) Explain the causes of high infant mortality in Bangladesh.

Q4. a) Differentiate between fertility and fecundity. Define general fertility rate and total fertility rate.
b) Distinguish between i) gross reproduction rate and net reproduction rate ii) mortality and morbidity iii) infant mortality and child mortality.
c) Describe the factors that affect fertility rates in Bangladesh.

Q5. a) Define life table. What are the different uses of life table? Distinguish between Cohort life table and current life table.
b) Explain different functions of life table and their interrelation.
c) What is probability of survival and expectation of life? Under usual notations, show that i) $p_x = \frac{e_x}{1+e_{x+1}}$ ii) $e_x^0 = \frac{t_x}{l_x}$

-----Good Luck-----

Answer Any Three of the Following Questions. All Questions Carry Equal Marks.

1. a) What is statistics quality control? Classify the four factors and processes of production which quality control mainly covers.
b) What are the differences between chance and assignable causes of variations?
c) In Statistical quality control, write down the six elements to a cybernetic or self-regulating system.

2. a) What do you mean by process and product control? What is control chart? Explain the basic principles underlying the control charts. Discuss the role of control charts in manufacturing processes.
b) Briefly write down the differences between the controls chars for variables and control charts for attributes.
c) What are the criterion for detecting the lack of control in \bar{x} and R charts?

3. a) What is acceptance sampling? Mention different types of acceptance sampling.
b) What is rectifying inspection plan? Let us consider a sample of size 30 in a double sampling plan with number of defectives 4 and acceptance number for 1st sample is 2 and 2nd sample is 8. What is your decision about the process?
c) Discuss why single sampling plan is binomial sampling? Give an example.

4. a) What do you mean by sequential probability ratio test (SPRT)? Discuss how sequential sampling plan differ from double sampling plan.
b) In a SPRT, state the situation when (i) the lot will be accepted (ii) the lot will be rejected and (iii) the production process will be continued.
c) Determine the OC and ASN curve for SPRT.

5. a) Explain briefly the concepts of reliability and validity of scores in educational and psychological experiments.
b) Define T-scores. Write down the procedure for obtaining T-scores for any frequency distribution.
c) What is official statistics? What are the sources of official statistics? Mention them with their at least three publications.

Department of Statistics
Jahangirnagar University
Part II B. Sc. (Hon's) Examination - 2017
Course Title: Macroeconomics and Economic Statistics
Course No.: Stat-206

Full Marks 35

Nihon

Time: $2\frac{1}{2}$ Hours

Answer any THREE from the following Questions. Each question carries equal Marks.

Q1. a) What is economics? Differentiate microeconomics and macroeconomics with examples. Why do you study economics as a student of Statistics discipline?
b) What are the goals of macro economics and what are the available instruments to achieve those goals? Explain with a table.

Q2. a) Define GDP (Gross Domestic Products). Distinguish between GNP and NNP.
b) Draw the circular flow diagram of GDP.
c) Which of the following transactions would be counted or would not be counted in GDP? Explain your own answer:
i) Bangladesh Government builds a new industrial plant.
ii) A drug dealer sells \$5000 worth of illegal drugs.

Q3. a) What is money? What are the official measures of money? Are all the measures really money?
b) Given a principle of \$1500 at 10 percent for 4 years, find the future value of money when the principle is compounded at i) annually ii) semiannually iii) quarterly and iv) continually.

Q4. a) Define aggregate demand (AD) curve and aggregate supply (AS) curve graphically. Why does the aggregate demand curve downward sloping? Explain any one reason.
b) Explain with appropriate figures what happen of price level and output level through aggregate demand (AD) curve or aggregate supply curve (AS):
i) If Government expenditures increases
ii) If technological innovation will occur in the economy
iii) If import level increase

Q5. a) Define poverty. What are the measures of calculating poverty and inequality?
b) Write down the differences between growth and development.
c) Discuss the economic effects of inflation in an economy.

Department of Statistics
 Jahangirnagar University
 Part II B. Sc (Hons.) Examination-2017

Course Code: Stat-207

Course Title: Mathematical Analysis and Differential Equations

Time: 4 Hours

Max

Full Marks: 70

Answer Any Five of the Following Questions. All Questions Carry Equal Marks.

1. Choose the correct option and justify your choice.
 - Determine the order of the given differential equation; also state whether the equation is linear or nonlinear: $\frac{d^2y}{dt^2} + \sin(t+y) = \sin t$
 - Second order, nonlinear
 - Second order, linear
 - First order, nonlinear
 - Which are the solutions of the differential equation: $y'' + 2y' - 3y = 0$
 - $y_1(t) = e^t, y_2(t) = 3t + t^2$
 - $y_1(t) = 3t + t^2, y_2(t) = e^{-3t}$
 - $y_1(t) = e^{-3t}, y_2(t) = e^t$
 - Classify the differential equation, $e^x \frac{dy}{dx} + 3y = x^2 y$.
 - Separable and not linear
 - Linear and not separable
 - Both separable and linear
 - The differential equation, $\frac{dy}{dx} = -\frac{ax+by}{bx+cy}$ is reducible to
 - homogeneous form
 - variable separable form
 - exact form.
 - Value of $L^{-1}\left(\frac{1}{s^n}\right)$ is
 - $\frac{t^{n-1}}{(n-1)!}$
 - $\frac{t^{n+1}}{n!}$
 - $\frac{t^n}{(n+1)!}$
 - Assume that $\sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} a_n x^n$ for all x , and determine what this implies about the coefficients a_n .
 - $a_{n+1} = \frac{a_n}{n+1}$
 - $a_{n+1} = \frac{(-1)^n a_n}{n-1}$
 - $a_{n+1} = \frac{a_{n-1}}{n+1}$
 - Identify whether the function, $f(x) = x^3 - 2x + 1$ is
 - even function
 - odd function
 - None.
2. a) What is differential equation? Using suitable examples, discuss different types of differential equation known to you.

 b) Consider the following initial value problem,

$$\frac{dy}{dt} = 3 - 2t - 0.5y, \quad y(0) = 1.$$
 Use Euler's method with step size $h = 0.2$ to find approximate values of the solution of the given equation at $t = 0.2, 0.4, 0.6, 0.8$, and 1 .

 c) Solve the following first order differential equation using suitable method
 - $y' - y = 2te^{2t}, y(0) = 1$
 - $\frac{dy}{dx} = \frac{x^2}{1+y^2}$
3. a) Define power series and convergent series. When a series is said to be convergent or divergent? Explain with example.

 b) Consider the differential equation: $y'' - xy' - y = 0$

(i) Find power series solutions of the form $y = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ of the given differential equation about the point $x_0 = 0$; find the recurrence relation for the coefficient a_n .

(ii) Find the first four terms in each of two solutions y_1 and y_2 .

(iii) By evaluating the Wronskian $W(y_1, y_2)(x_0)$, show that y_1 and y_2 form a fundamental set of solutions.

(iv) Find the general term in each solution.

4. a) Define Bessel's function. Find the general solution of Bessel's differential equation of order n .

b) If $J_\nu(x)$ denotes the Bessel function of order ν , then prove that

$$xJ'_\nu(x) = \nu J_\nu(x) - xJ_{\nu+1}(x)$$

c) What do you mean by Legendre polynomials? Write down the two Legendre's differential equations of order 3 and of order 4 and hence, find their particular solutions.

5. a) Describe Laplace transformation and inverse Laplace transformation using suitable example. How they can be used in solving differential equation?

b) Use the Laplace transform to solve the given initial value problem.

$y'' - y' - 6y = 0; y(0) = 1, y'(0) = -1$

c) Let $f(t) = \sin t$. Find the Taylor series for f about $t = 0$ and then assume that the Laplace transform of this function can be computed term by term to show that

$$(i) \sin t = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!}, \quad (ii) L\{\sin t\} = \frac{1}{s^2 + 1}, \quad s > 1.$$

6. a) Define complex number. Express $\frac{(1+5i)^2}{(2+i)^2}$ in the form of $a+ib$. Also find its modulus and argument.

b) Determine whether the function is even, odd or neither.

(i) $f(x) = x \cos x$ (ii) $f(x) = x^3 - 4x$ (iii) $f(x) = e^x - e^{-x}$

c) Find the general solution of the Bernoulli's differential equation $\frac{dy}{dx} + Py = Qy^n, n \neq 1$, where P and Q are only functions of x or constant.

7. a) Explain the concept of analytic function, harmonic function, and harmonic conjugate functions. Give Example.

b) Show that if a function $f(z) = u(x, y) + iv(x, y)$ is analytic in a domain D , then its component functions u and v are harmonic in D .

c) Show that the function $f(z)$ has a derivative everywhere in its domain of definition and also find $f'(z)$, if

(i) $f(z) = e^{-x} e^{-iy}$;

(ii) $f(z) = \sqrt{r} e^{i\theta/2}; (r > 0, \alpha < \theta < \alpha + 2\pi), \alpha$ is a fixed real number.

8. a) What do you mean by contour? When a contour is said to be simple closed contour?

b) Obtain the Taylor series representation of $f(z) = \cos z$.

c) Discuss singularities, Poles and Laurent series for complex numbers. Find the Pole and residue of $f(z) = \frac{e^z}{z^2(z-1)}$.

Department of Statistics
Jahangirnagar University
Part II B. Sc. (Hon's) Examination - 2017
Course Name: Statistical Simulation and Data Processing
Course No.: Stat-208

Full Marks 70

1/2 hour

Time: 4 Hours

Answer any FIVE from the following Questions. Each question carries equal Marks.

Q1. a) What do you mean by simulation? What are different types of simulation?
Mention with practical example.
b) What is random number? What are the different types of random numbers?
Explain briefly.
c) Write down the algorithm to generate pseudo random number using linear congruential generator.

Q2. a) What are the properties of a good random number generator? Describe an empirical testing technique to test random numbers.
b) Mention some variable reduction technique. Explain one of them.
c) Prove that the estimator of antithetic variables is more efficient than sample mean Monte Carlo method.

Q3. a) What do you mean by Monte Carlo simulation? What are the different methods of Monte Carlo simulation?
b) Discuss the method of Hit or Miss Monte Carlo simulation for finding the value of integral $I = \int_a^b f(x)dx$ with $0 \leq f(x) \leq c$. Also find the mean and variance of the estimator.
c) Discuss the advantage, disadvantage, and uses of Monte Carlo simulation.

Q4. a) What does SPSS stands for? When you start a session in SPSS, what did you see?
Discuss.
b) How can you read an i) excel file ii) text file in SPSS? Explain.
c) What are the general rules for naming a variable in SPSS?

Q5. a) Distinguish between data view and variable view. Mention with example different data types of SPSS. How data is encoded in SPSS?
b) Write down the steps to do the following task using SPSS syntax editor window pull down menu.
i) Regression analysis ii) Correlation iii) Box and Whisker plot
iv) Crosstabs.

c) Explain the following commands with suitable example:

- Recode (for string and numeric variables)
- Compute
- Frequency
- Add value label and
- Variable label.

Q6. a) What is SAS? What are the basic steps of SAS program? Discuss them briefly.
 What is SAS data set? What are the different types of data sets in SAS program?

b) What are the two INFILE options for comma delimited files in SAS? Write down the basic features of DLM and DSD options. When do you use MISSOVER and PAD options?

c) What do you mean by array in SAS? What are the advantages of it?

Q7. a) What is a code book and code plan? Write down the functions of code book in SPSS data analysis.

b) What are the different types of menus in SPSS? Discuss them.

c) Suppose you are given the filled in questionnaire of a survey that was conducted to examine possible linkage of tertiary students GPA with their gender, place of residence, university type, use of library, parents income and parents education.

- Develop a code plan to enter the survey data.
- Develop an analysis plan to serve the above purpose.

Q8. a) What are the basic characteristics of SAS statement? How can you create permanent SAS data set?

b) Given the following data, how much of the variance of SBP can be explained by the fact that there is variability in AGE (in years)? Write a SAS program.

Age (in yrs)	15	20	25	30	40	50	58	55
SBP	116	120	130	132	150	148	160	150

c) What are the naming conventions of SAS variables? How can you read text file located at the hard disk drive?

-----Good Luck-----

Time: 3.0 Hours

Full Marks: 21

Mirza
Answer the Following Questions. All Questions Carry Equal Marks.

Q1. To study the relationship between investment rate (investment expenditure as a ratio of the GDP) and savings rate (savings as a ratio of GDP), Martin Feldstain and Charles Horioka obtained data for a sample of 21 countries. (see Table 6.8). The investment rate for each country is the average rate for the period 1960-1974 and the savings rate is the average savings rate for the period 1960-1974. The variable Y represents the investment rate and the variable X represents the savings rate.

- 1) Plot the investment rate (Y) against the savings rate (X).
- 2) Plot the log of investment rate (Y) against the log of savings rate (X).
- 3) What is the difference between the two graphs?
- 4) From these two graphs, do you think that a log-linear model might provide a better fit to the data than the linear model?
- 5) Estimate both these models and obtain the usual statistics.
- 6) How would you interpret the slope coefficient in the linear model?
- 7) How would you interpret the slope coefficient in the log-linear model?
- 8) Is there any difference in the interpretation of these slope coefficients in the linear and log-linear models?
- 9) Would you compare the two coefficients of determinations? Why or why not? Given the results of the two regression models, which model would you prefer? Why?
- 10) How would you estimate the elasticity of the investment rate with respect to the savings rate for the linear model? What additional information, if any, do you need?
- 11) Is there any difference between the elasticity of the investment rate with respect to the savings rate estimated from the log-linear model and that estimated from the linear model? If so, which model would choose?

Q2. Table 8.8 shows the data of output, labor and capital of the Indian manufacturing sector, 2008-09.

- 1) See if the Cobb-Douglas production function fits the data given in the table and interpret the results.
- 2) Now, consider the regression model of the **labor productivity** (that is, **output labor ratio**) on the capital labor ratio. What is the economic significance of such a relationship, if any? Estimate the parameters of this model and interpret your results.

Q3. Consider the data given in Table 8.11. This table gives data on disposable personal income and gross domestic savings, in crores of rupees, for India for the period 1974-75 to 2004-2005. Suppose, we want to estimate a simple savings function that relates savings to disposable personal income. Suppose, we want to see if a structural change occurred within the data between two time periods. Split the data into three sections: 1974-75 to 1988-89, 1989-90 to 1995-96 and 1996-97 to 2004-05.

- 1) Determine if there is a significant break between the two time periods 1974-75 to 1988-89 and 1989-90 to 2004-05.
- 2) Determine if there is a significant break between the two time periods 1974-75 to 1995-96 and 1996-97 to 2004-05.

Q4. Table 11.3 gives the hypothetical data on consumption expenditure (Y) and income (X) of 30 families. Regress consumption expenditure (Y) on income (X) and comment on the nature of heteroscedasticity, if any, present in the data with the help of graphical method and Breusch-Pagan-Godfrey (BPG) test.

Department of Statistics
Jahangirnagar University

Part II B.Sc. (Hons.) Practical Examination 2017
Course No. STAT Lab. 209 (Group B) Statistical Data Analysis III
Time: 1 1/2 hours Marks: 14

Munmun

Answer the following questions. All questions carry equal marks.

1. Consider the following information on age group, number of women and number of births in a given year for a country. Calculate age-specific birth rates, Total fertility rate and Gross reproduction rate.

Age Group	No. of women	Total number of births
15-19	410,352	56,403
20-24	378,163	97,166
25-29	377,011	93,415
30-34	296,799	57,694
35-39	235,697	34,478
40-44	190,328	13,078
45-49	148,534	4,546

The sex ratio at birth is 103 male births to female births.

2. Calculate net reproduction rate, intrinsic rate of natural increase and mean length of generation by using the information in the following table:

Age of mothers	Annual birth of daughter per female	$\frac{5L_x^F}{l_0^F}$
15-19	0.0389	3.8973
20-24	0.1123	3.7869
25-29	0.1198	3.7593
30-34	0.0987	3.6943
35-39	0.0698	3.5321
40-44	0.0345	3.2399
45-49	0.0113	3.1986

3. Using the life table functions and their relationships complete the missing entries in the following table.

Age	q_x^n	l_x	d_x^n	L_x^n	T_x	e_x^0
0-1	.02593	100,000	-	97,815	6,989,030	-
1-2	-	-	165	-	-	70.75
2-3	.00104	-	-	97,192	-	-
3-4	-	97,141	78	-	-	-
4-5	.00067	-	-	97031	-	67.99
5-6	-	96,998	57	-	6,502,566	-
6-7	.00052	-	-	96,916	-	-
7-8	-	96,891	46	-	-	65.11
8-9	.00043	-	-	-	-	-
9-10	-	96,803	38	96,748	6,114,989	-

Good Luck

Department of Statistics
Jahangirnagar University
Part II B.Sc. (Hons.) Practical Examination 2017
Course No. STAT Lab. 210 (Group A) Statistical Data Analysis IV
Time: 1 hour Marks: 7
Nihon

Answer the following questions. All questions carry equal marks.

1. The following data give the number of defectives in 10 independent samples of varying size from a production process.

Sample No.	Sample Size	No. of defectives
1	2000	425
2	1500	430
3	1400	216
4	1350	341
5	1250	225
6	1760	322
7	1875	280
8	1955	306
9	3125	337
10	1575	305

Draw the control chart for fraction defective and comment on it (use the method I only).

2. (a) A test is administered on 400 pupils. It gave mean 60 and standard deviation 12. Complete the following table of equivalent raw scores.

Raw score:	84	78	72	66	60	54	48	72	36
σ score:	-	-	1	-	0	-	-	-	-
Standard score	-	-	-	-	-	40	-	-	-

(b) Convert the scores 1, 2, ..., 10 into standard scores with mean 50 and standard deviation 10.

Department of Statistics
Jahangirnagar University, Savar, Dhaka
Part II B.Sc (Hons.) Practical Examination-2017
Course Title: Statistical Data Analysis III
Course No. : Stat-lab-210 (Group B)

Full Marks: 28

Time: 3 hrs

Nihen
[Answer all the following questions]

Q1. a) Generate 1000 random numbers from $U(0,1)$ distribution using Linear congruential generator for $X_0 = (110 + \text{last three digits of your exam roll number})$ $a = 7^3, c = 11$ and $m = 111111$. (Using Excel/SAS/SPSS) [4]

b) Find the period of this generator. Also find the descriptive statistics for the above generated random numbers using SPSS. [2]

Q2. Suppose that the discrete random variable X has probability function $p(x = -2.5) = 0.35, p(x = 1.0) = 0.25, p(x = 8.5) = 0.40$ and $p(x) = 0$, otherwise. Set up a table to generate realizations from this distribution. Also draw a random sample of size 200 from this distribution generating PRNs U_i . Obtain the sample mean, sample variance and sample standard deviation of random variable X . (using Excel/SAS/SPSS) [4]

Q3. Consider the inverse transformation method to generate 20 standard normal $(0,1)$ random variable. i) Demonstrate your technique using the PRNs U_i , where, $i = 1, 2, \dots, 200$. ii) Using answer in (i) generate an $N(1,9)$ random variable. iii) Check whether the generated data in (ii) follows normal distribution (use any graphical method). [5]

Q4. Suppose that X_1, X_2, \dots, X_n are IID exponential random variables with rate $\alpha = 2$. Then $Y = \sum_{i=1}^n X_i$ is said to have an Earlong distribution with parameters n and α . Generate an Earlong realization with parameter $n = 10$ and $\alpha = 2$. [4]

Q5. Generate 1000 random values from the following distributions using SAS and SPSS:
i) exponential distribution with parameter 0.8, ii) uniform distribution over (2,5), iii) binomial distribution with parameter 10 and 0.7. Also find mean, variance and standard deviation of generated data series. [4]

Q6. A SAS data set prob_1 (prob_1.txt) in your desktop contains three variables X, Y and Z.

	PROB_1	
X	Y	Z
15	55	90
31	25	88
45	35	65
50	21	24

Perform the following tasks using SAS statements: [5]

- Select all the variables where Z is greater than 50.
- Get a contingency table of Z vs. Y. Carry out a chi-square test of independence for these two variables.
- Find the correlation coefficients for X, Y, with Z. Here, test the hypothesis of no correlation. Also fit the regression equation for estimating Z using X and Y.



Random variable is
a variable whose outcome
depends on random experiment

Department of Statistics

Jahangirnagar University

Part II B. Sc (Hons.) Final Examination-2018

Course #: STAT-201

Course Title: Probability Distributions

(Answer any THREE (03) questions. All questions carry equal marks)

Time: 2.5 hours

Full Marks: 35

1. (a) Define probability distribution of a random variable. How does it differ from a frequency distribution? Distinguish between probability density function (pdf) and probability mass function (pmf).
(b) Derive the pmf of binomial distribution. Find mean and variance of binomial distribution. Show that mean is greater than variance of binomial distribution.
(c) If x_1, x_2 are two independent binomial variates with parameters (n_1, p) and (n_2, p) then $x = x_1 + x_2$ is a binomial variate with parameters $(n_1 + n_2, p)$. Hence show that binomial distribution does not possess the additive property, except $p_1 = p_2 = p$.
2. (a) What is a Poisson variate? Give some typical examples of a Poisson variate.
(b) Derive the Poisson distribution as a limiting case of a binomial distribution with parameters n and p .
(c) Find the moment generating function of a Poisson distribution and hence find its β_1 and β_2 . Also comments on the shape of the distribution.
3. (a) Define uniform distribution with examples. Show that for rectangular distribution $f(x) = 1/2a$, $-a < x < a$, the m.g.f. about the origin is $\frac{1}{at} \sinh(at)$. If $X \sim U[0,1]$, find the distribution of $-2 \log x$.
(b) What is an exponential distribution? When exponential distribution is called standard exponential distribution? If X is an exponential variate with pdf $f(x; \alpha) = \alpha e^{-\alpha x}$; $x > 0, \alpha > 0$ then show that the m.g.f. of x is $M_x(t) = \left(1 - \frac{t}{\alpha}\right)^{-1}$.
(c) Let X be an exponential variate with pdf $f(x; \alpha) = \alpha e^{-\alpha x}$; $x > 0, \alpha > 0$. Show that the median of the distribution is $\frac{1}{\alpha} \log 2$, and the mean deviation about mean is $\frac{2}{\alpha e}$.
4. (a) What do you mean by Normal distribution and standard Normal distribution? Describe the characteristics of the Normal distribution.
(b) If X and Y are independent log-normal variates, then prove that, XY and X/Y are also log-normal variates.
(c) Let X be a standard normal random variable. Find the probability that X will,
(i) less than 1.5, (ii) greater than 2.4, (iii) between 1.5 and 2.14, and (iv) less than 1.32.
5. (a) Define negative binomial distribution and show that negative binomial random variable is a sum of geometric random variable.
(b) If X and Y are independent Gamma variate with parameter l and m , then show that, $U = X + Y$ and $V = \frac{X}{Y}$ are independent and U is a Gamma variate with parameter $l+m$ and V is a Beta variate of the second kind with parameter l and m .
(c) Define bivariate normal distribution. If X and Y have a bivariate normal distribution with mean and variance $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ respectively and ρ is the correlation coefficient of X and Y . Obtain the moment generating function of bivariate normal distribution.



Department of Statistics
Jahangirnagar University

Part II B. Sc (Hons.) Final Examination-2018

Course #: STAT-202

Course Title: Sampling Distributions

(Answer any FIVE (05) questions. All questions carry equal marks)

Time: 4 hours

Full Marks: 70

1. (a) Explain the concept of sampling distribution with an example. Discuss the purpose of sampling distribution.
(b) How do you obtain the distribution of a random variable? Discuss the case of single random variable and several random variables.
(c) Elaborate the analytical method to find the distribution of a function of random variables.

2. (a) Write the definition of characteristic function (c.f.). What are the necessary and sufficient conditions for $\phi(t)$ to be a c.f? Mention its properties.
(b) State and prove Inversion theorem.
(c) Find the density function whose c.f is $\phi(t) = (q + pe^t)^n$.

3. (a) What do you mean by variance stabilization transformation? Explain with a suitable example.
(b) Discuss Fisher's Z-transformation. Show that the distribution of Z-statistic is approximately normally distributed with mean $E(Z) = \frac{1}{2} \log_e \frac{1+\rho}{1-\rho}$ and variance $V(Z) = \frac{1}{n-3}$.
(c) The variance of the Binomial and the Poisson random variables are some functions of their means. Choose appropriate variance stabilizing procedure and obtain the constant variance for both the random variables.

4. (a) Explain the concept of Laplace transformation. State the important properties, uses and importance of Laplace transformation.
(b) Let χ^2 is a chi-square variable with n d.f and $f(\chi^2) = \frac{1}{2^{n/2} \Gamma(n/2)} e^{-\frac{\chi^2}{2}} (\chi^2)^{\frac{n}{2}-1}; \chi^2 > 0$. Obtain the Laplace transformation and hence find mean, variance, β_1 and β_2 . Also, make comments.
(c) State central limit theorem. Also, state and prove, De Moivre-Laplace theorem.

5. (a) Discuss chi-square distribution and show that, $\int_0^\infty f(\chi^2) d\chi^2 = 1$.

7. (a) What do you mean by the non-centrality parameter? Explain with example. How does a non-central distribution differ from a central distribution?
(b) Find the characteristic function of a non-central chi-square distribution. Hence or otherwise find the mean and variance of the distribution.
(c) If X_k is distributed as normal with mean μ and variance-covariance matrix D , where D is the diagonal matrix, then show that $X^T D^{-1} D$ is distributed as a non-central chi-square with k degrees of freedom and non-centrality parameter $\mu^T D^{-1} \mu = \lambda$.

8. (a) Define order statistics with an example. How would you obtain the joint distribution of the n order statistics? Also, obtain the marginal distribution of the n^{th} order statistics.
(b) How would you find the exact moments of order statistics?
(c) Obtain the large sample approximations to the mean and variance of the r^{th} order statistics.



Department of Statistics
Jahangirnagar University
Part II B. Sc (Hons.) Final Examination-2018
Course #: STAT-203
Course Title: Regression Analysis
(Answer any **FIVE (05)** questions. All questions carry equal marks)

Time: 4 hours

Full Marks: 70

1. (a) What do you mean by regression analysis? Distinguish between population regression function and sample regression function. What are the underlying assumptions of classical linear regression model?
(b) Estimate the parameters of a simple linear regression model by the ordinary least squares method and show that they are the best linear unbiased estimators.
2. (a) How can you test the overall significance of a multiple regression model? Discuss the testing procedure of equality of two regression coefficients in case of a three-variable regression model.
(b) Establish a relationship between R^2 and F . How can you test the incremental or marginal contribution of an explanatory variable in a regression model?
(c) How can you choose an appropriate regression model between a linear regression model and a log-linear regression model? Discuss with the help of the MacKinnon, White and Davidson (MWD) test.
3. (a) Define partial correlation coefficient. How are they related to total regression and total correlation coefficient?
(b) Define the coefficient of partial determination. Show that, $r_{12,3}^2 = \frac{(R_{123}^2 - r_{13}^2)}{1 - r_{13}^2}$ and interpret the result.
(c) If $R_{123} = 1$, then prove that $R_{2,13} = 1$ and $R_{3,12} = 1$ and if $r_{12} = r_{23} = r_{13} = r \neq 1$, then prove that $R_{1,23} = R_{2,31} = R_{3,12} = \frac{r\sqrt{2}}{\sqrt{1+r}}$.
4. (a) What do you mean by auto-correlation? What are the consequences of using OLS method in case of auto-correlation?
(b) What do you mean by multicollinearity? What assumption concerning multicollinearity? What are the main reasons to arise multicollinearity?
(c) If the inter-correlation between the explanatory variables is perfect ($r_{x_i x_j} = 1$) then prove that (i) the estimates of the coefficients are indeterminate (ii) the standard errors of these estimates become infinitely large.
5. (a) Explain the concept of dummy variables with examples. What are the cautions in the use of dummy variables? What are the advantages of dummy variable technique over the Chow test?
(b) What do you mean by ANOVA and ANCOVA models in case of dummy variables? Consider the following model:
$$Y_i = \beta_1 + \beta_2 D_{2i} + \beta_3 D_{3i} + u_i$$

$$Y_i = \text{Average consumption expenditure (Tk.) per person per 30 days in state } i \text{ in Bangladesh}$$

$$D_{2i} = 1, \text{ if the state is in the Eastern region of Bangladesh}$$

$$= 0, \text{ otherwise}$$

$$D_{3i} = 1, \text{ if the state is in the North-west-central region of Bangladesh}$$

$$= 0, \text{ otherwise}$$

How can you interpret the parameters of this model? Explain.
(c) Consider the following model:

$$Y_i = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 (D_{2i} D_{3i}) + \beta X_i + u_i$$

where, Y_i = Literacy rate (percent)

X_i = Per capita net state domestic product

$D_{2i} = 1$, if female

= 0, otherwise

$D_{3i} = 1$, urban

= 0, otherwise

How can you interpret the parameters of this model? Explain.

6. (a) What do you mean by binary response regression model? What are the available approaches for developing a probability model for a binary response variable?
 (b) What is linear probability model? Discuss the several problems associated with it. What are the alternatives to linear probability model?
 (c) What do you mean by Logit model? What are the features of this model? How can you estimate the parameters of this model in case of grouped or replicated data?

7. (a) What do you mean by selecting the best regression equation? What methods are available for selecting the best regression equation? Discuss the stepwise regression procedure to select the best regression equation.
 (b) Explain the concept of ridge regression? Why is it intended in regression analysis? Briefly discuss the procedure of estimating the parameters of a ridge regression.

8. The following table gives the data on child mortality (Y), that is the number of deaths of children under age 5 in a year per 1000 live births, per capita GNP in 1980 (X_2), female literacy rate in percent (X_3) and total fertility rate (X_4), 1980-1985, the average number of children born to woman, using age specific fertility rate for a given year.

Obs#	Y	X2	X3	X4
1	128	1870	37	6.66
2	204	130	22	6.15
3	202	310	16	7
4	197	570	65	6.25
5	96	2050	76	3.81
6	209	200	26	6.44
7	170	670	45	6.19
8	240	300	29	5.89
9	241	120	11	5.89
10	55	290	55	2.36
11	75	1180	87	3.93
12	129	900	55	5.99
13	24	1730	93	3.5
14	165	1150	31	7.41
15	115	1130	102	5.03
Average	150	840	50	5.5

$$\sum x_{2i}^2 = 6025400, \sum x_{2i} x_{3i} = 165140$$

$$\sum x_{2i} x_{4i} = -3621, \sum x_{3i}^2 = 11906$$

$$\sum x_{3i} x_{4i} = -384.45, \sum x_{4i}^2 = 29.5342$$

$$\sum y_i^2 = 63908, \sum x_{2i} y_i = -414020$$

$$\sum x_{3i} y_i = -21526, \sum x_{4i} y_i = 1056.8$$

Here,

$$x_{2i} = (X_{2i} - \bar{X}_2), y_i = (Y_i - \bar{Y})$$

$$x_{2i} x_{3i} = (X_{2i} - \bar{X}_2)(X_{3i} - \bar{X}_3)$$

$$x_{2i} y_i = (X_{2i} - \bar{X}_2)(Y_i - \bar{Y})$$

(a) Estimate the parameters of the regression model of child mortality on per capita GNP, female literacy rate and total fertility rate. That is, find the following fitted model:
 $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \hat{\beta}_4 X_{4i}$. Also, interpret the parameters of the model.

(b) Calculate SST, SSR and SSE. Also, construct the ANOVA table.

(c) Calculate $\text{var}(\hat{\beta}_2)$, $\text{var}(\hat{\beta}_3)$ and $\text{var}(\hat{\beta}_4)$. Hence, calculate $\text{var}(\hat{\beta}_1)$. Also, calculate their standard errors.

(d) Test the significance of the model and individual parameters of the model. Calculate the coefficient of determination of the model and comment.



**Department of Statistics
Jahangirnagar University
Part II B. Sc (Hons.) Final Examination-2018**

Course #: STAT-204

Course Title: Demography

(Answer any THREE (03) questions. All questions carry equal marks)

Time: 2.5 hours

Full Marks: 35

1. (a) Define demography. Discuss in brief the scope and importance of demography.
(b) What are the major sources of demographic data? Critically review these sources in terms of their scope and limitations.
(c) Explain the importance of age and sex composition of population data. Also, explain population pyramid.

2. (a) What do you mean by mortality? What are the advantages and disadvantages of mortality? What is growth rate? What are the different methods that are used to compute growth rates?
(b) What do you mean by standardization of death rates? Write down the necessity of standardization.
(c) Explain in detail the direct and indirect method of standardization.

3. (a) What do you mean by population growth? Explain exponential growth model.
(b) What is meant by doubling time of a population? Derive an expression for computing the doubling time of a population.
(c) Differentiate between fertility and fecundity. Define general fertility rate and total fertility rate.

4. (a) Explain the terms: coverage error, content error, under enumeration, age heaping, and dependency ratio.
(b) What do you mean by age reporting error? Explain Whipple's index and Myer's index for calculating age reporting error.
(c) Define crude birth rate (CBR) and crude death rate (CDR) with advantages and disadvantages.

5. (a) What is life table? Why life table is so important in population policy and life insurance? Distinguish between cohort life table and current life table.
(b) What are the underlying assumptions in the construction of life table? Explain different functions of life table and their relationships.
(c) What do you mean by central mortality rate? Under usual notations, show that
$$(i) m_x = \frac{d_x}{l_x - \frac{1}{2}d_x}$$

$$(ii) q_x = \frac{2m_x}{2 + m_x}$$



Department of Statistics
Jahangirnagar University

Part II B. Sc (Hons.) Final Examination-2018

Course #: STAT-205

Course Title: Quality Control and Educational Statistics
(Answer any THREE (03) questions. All questions carry equal marks)

Time: 2.5 hours

Full Marks: 35

1. (a) Define quality, quality control and statistical quality control. Discuss its need and utility in industry.
(b) What do you mean by causes of variation? The quality of materials, manpower, machines, and management are the essential part in any production process, which one do you think the most important and why?
(c) What do you mean by productivity enhancing and regulatory technique (PERT)?
2. (a) Discuss how control charts for \bar{X} and R are constructed. Explain briefly the different criteria for detecting lack of control in \bar{X} and R charts. Discuss different situations where R chart is in a state of control but \bar{X} chart is not. Interpret the points falling outside the control limits on these charts.
(b) Explain p -charts for fraction defective and d -chart for number of defectives.
(c) Write short notes of the following terms (i) Process Average Fraction Defective (ii) Acceptance Quality Level (iii) Average Sample Number (iv) Average Amount Total Inspection.
3. (a) Briefly explain the acceptance sampling by attributes and rectifying inspection plan.
(b) Describe single sampling plan. Obtain O. C. and A. O. Q curve for this plan. Distinguish clearly between: (i) Producer's and Consumer's risk, (ii) Average Sample Number (ASN) and Average Amount of Total Inspection (ATI).
(c) What is double sampling plan? What are the differences between single sampling plan and double sampling plan?
4. (a) What are the importances of studying educational statistics? Define z-score. How does it differ from the standard score? Also, explain the T-score.
(b) State different methods of testing reliability. Explain the Spilt-half method with its relative advantages and disadvantages.
(c) Convert the scores 1, 2, 3, ..., 15 into standard scores with mean 50 and standard deviation 10.
5. (a) What do you mean by official statistics? Also define semi-official statistics and non-official statistics. Who are the main compiler and publisher of official statistics in Bangladesh?
(b) What are the limitations of official Statistics in Bangladesh? Discuss them. What steps should be taken to improve the quality of official statistics?



Department of Statistics
Jahangirnagar University
Part II B. Sc (Hons.) Final Examination-2018
Course # STAT-206
Course Title: Macroeconomics and Economic Statistics
(Answer any Three (03) questions. All questions carry equal marks)

Time: 2.5 hours

Full Marks: 35

1. (a) Suppose in a hypothetical economy, three (3) goods are produced: Mango, Banana and Chicken. The quantities of output and prices of Mango, Banana and Chicken are given in the following table:

	Output			Price (Tk./unit)		
	2014	2015	2016	2014	2015	2016
Mango	125	135	140	50	55	65
Banana	500	525	540	06	07	09
Chicken	200	215	225	110	115	125

Define nominal and real GDP (Gross Domestic Product). Calculate the nominal GDP for the year 2014, 2015 and 2016. Suppose that 2014 is the base year, and then calculate real GDP for the year 2014, 2015 and 2016. Define growth rate. Calculate the growth rate of real GDP for the year 2015 and 2016.

(b) Explain GDP deflator. What does it measure? According to GDP deflator, how much prices have increased for the year 2015 and 2016?

2. (a) What is aggregate demand (AD)? Why does the AD curve slopes downward to the right?
(b) Explain with the diagram what will happen to the short run equilibrium of the economy if the government adopts expansionary fiscal policy.
(c) Distinguish between the movement along the AD curve and shift of the AD curve.

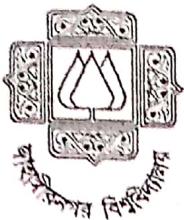
3. A consumer's basket of goods for a hypothetical economy includes 10 Pizzas, 15 Cokes, and 12 Hamburgers. Use the table below to answer the following questions (2012 is the base year)

Year	Prices of Pizza (Tk.)	Prices of Coke (Tk.)	Prices of Hamburger (Tk.)
2011	100	20	80
2012	110	25	90
2013	120	28	95

(a) Calculate the cost of the consumer's basket for each of the years given
(b) Calculate the CPI for each of the years
(c) From the CPI, calculate the inflation rate for the year 2012 and 2013
(d) Which one is the better measure of inflation, CPI or GDP deflator? Explain

4. (a) Why there is structural unemployment? Explain.
(b) In 2011, Bangladesh Bureau of Statistics (BBS) reported that the total population was 156 million of which 70% were adult. The numbers of adults who were not in the labor forces and unemployment were 20 million and 18 million respectively. What is the labor force participation rate? What is the unemployment rate?

5. (a) Define money. Briefly explain the major functions of money.
(b) Define monetary policy. Explain the tools to control money supply



Time: 4 hours

Department of Statistics
Jahangirnagar University

Part II B. Sc (Hons.) Final Examination-2018

Course #: STAT-207

Course Title: Mathematical Analysis and Differential Equations
(Answer any FIVE (05) questions. All questions carry equal marks)

Full Marks: 70

Q1. a) Write down the order and degree of the differential equation

$$x^5(y'')^3 + y(y')^4 + y^4 = 0$$

Is it linear? How many constants does the general solution of the differential equation must contain?

b) Solve the following initial-value problem

$$\frac{dy}{dx} + P(x)y = 4x, \quad y(0) = 3, \quad \text{where } P(x) = \begin{cases} 2 & ; 0 \leq x \leq 1 \\ -\frac{2}{x} & ; x > 1 \end{cases}$$

c) Reduce the equation $\frac{dy}{dx} = (-2x + y)^2 - 7$ to separable variables and solve it.

d) Express the solution of the initial-value problem $y' - 2xy = 2, \quad y(0) = 1$, in terms of $\text{erf}(x)$.

e) The population of a community is known to increase at a rate proportional to the number of people present at time. If an initial population P_0 has doubled in 5 years, how long will it take to triple?

Q2. a) Verify that the given differential equation is not exact. Multiply the given differential equation by an appropriate integrating factor to make it exact and then solve it.

$$xy \, dx + (2x^2 + 3y^2 - 20) \, dy = 0.$$

b) Classify each differential equation as linear, homogenous or Bernoulli. Some equations may be more than one kind. Solve them using appropriate transformation:

(i) $x \frac{dy}{dx} + y = \frac{1}{y^2}$ (ii) $(y^2 + yx) \, dx - x^2 \, dy = 0$

(c) Consider the initial value problem $y' = x + y^2, \quad y(0) = 0$. Use Euler's method to obtain an approximation of $y(0.1)$, use step size $h = 0.05$.

Q3. a) Find the interval and radius of convergence of the power series $\sum_{n=1}^{\infty} \frac{(x-3)^n}{2^n n}$

b) Consider the differential equation: $3xy' + y' - y = 0$. Determine whether $x = 0$ is a regular singular point of the given equation.

c) Find the recurrence formula and indicial equation for the equation given in Q2. (b). Also obtain the general solution of that equation.

d) Find the general solution near $x = -1$ of $y'' + xy' + (2x-1)y = 0$

Q4. (a) Write down four properties of Laplace transformation.

Find i) $L(e^{-2x} \cos 4x)$ (ii) $L\{f(x)\}$ where $f(x) = \begin{cases} 1 & 0 < x \leq 1 \\ -1 & 1 < x \leq 2 \end{cases}$ (iii) $L^{-1}\left(\frac{1}{s-4} e^{-2x}\right)$

b) Solve $y'' - 6y' + 9y = x^2 e^{3x}$, $y(0) = 2$, $y'(0) = 17$ using Laplace transformation method

c) Let $f(x) = \begin{cases} 0 & 0 \leq x < 1 \\ 5 & x \geq 1 \end{cases}$. Write the function in terms of step function. Find the Laplace transformation of the given function.

d) Solve the equation: $y'' + y = f(x)$, $y(0) = 0$, where $f(x)$ is given in Q4 (c).

Q5. a) Find the general solution of the Bessel equation: $x^2 y' + xy' + \left(x^2 - \frac{1}{4}\right)y = 0$.

b) Mention some properties of Bessel's function of order m , $m = 0, 1, 2, \dots$

c) Show that $J_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \sin x$ where $J_{\frac{1}{2}}(x)$ Bessel's function of first kind of order $\frac{1}{2}$.

c) Derive the formula $x J_{\nu}'(x) = \nu J_{\nu}(x) - x J_{\nu+1}(x)$, $J_{\nu}(x)$ is the Bessel's function of order ν .

Q6. a) Perform the indicated operation both analytically and graphically

$$(-3+5i) + (4+2i) + (5-3i) + (-4-6i)$$

b) Express the following complex number in polar form: (i) $2+2\sqrt{3}i$ (ii) $-3i$

c) Define harmonic function and write Cauchy-Riemann Equation. Prove that the function $u(x, y) = e^{(-x)}(x \sin y - y \cos y)$ is harmonic. Find conjugate harmonic function $v(x, y)$, and express $f(z)$ in terms of z .

c) Evaluate $\oint_C \frac{dz}{z-a}$ where C is any simple closed curve and $z = a$ (i) outside C (ii) inside C

Q7. a) Find Laurent series about the indicated singularity for each of the following functions:

Give the name of the singularity in each case. (i) $\frac{e^{2z}}{(z-1)^3}$; $z=1$ (ii) $\frac{z-\sin z}{z^3}$; $z=0$.

b) What is the Laurent series of the function $f(z) = \frac{z}{(z-1)(2-z)}$ around $a = 0$ in the region (i) $1 < |z| < 2$ (ii) $|z| > 2$.

c) Use Cauchy's integration formula to evaluate: $\oint_C \frac{e^{3z}}{z-\pi i} dz$, where C is a circle of radius 4, centered at 1.

Q8. a) State Residue theorem.

b) Find $\oint_C \frac{z^2 dz}{(z-1)^3}$ where C is the circle $|z| = 2$.

c) Evaluate $\int_0^{2\pi} \frac{\cos 3\theta}{5-4\cos\theta} d\theta$.

d) $\int_{-\infty}^{\infty} \frac{dx}{(x^2+4x+5)^2}$.



Department of Statistics
Jahangirnagar University
Part II B. Sc (Hons.) Final Examination-2018
Course #: STAT-208

Course Title: Statistical Simulation and Data Processing
(Answer any FIVE (05) questions. All questions carry equal marks)

Time: 4 hours

Full Marks: 70

1. (a) What is SAS? What are the different types of SAS window? What are the rules for the variable names in SAS?
- (b) You have completed an experiment and recorded a subject ID, and values for variables A, B, and C. You want to compute means for A, B, and C but, unfortunately, your lab technician, who didn't know SAS programming, arranged the data like this:

ID	TYPE	SCORE
1	A	44
1	B	9
1	C	203
2	A	50
2	B	7
2	C	188
3	A	39
3	B	9
3	C	234

Write a program to read these data and produce means.

- (c) What do you mean by array in SAS? What are the different advantages of it?
2. (a) Explain the terms "PROC MEANS" and "PROC UNIVARIATE" to creating summary of Data Sets. Which statistic(s) are found by PROC UNIVARIATE for a data set?
- (b) A survey is conducted and data are collected and coded. The data layout is shown below:

Variable	Description	Columns	Coding Values
ID	Subject identifier	1-3	
GENDER		4	M=Male, F=Female
PARTY	Political party	5	1=Republican, 2=Democrat, 3=Not registered
VOTE	Did you vote in the last election?	6	0=No 1=Yes
FOREIGN	Do you agree with the government's foreign policy?	7	0=No 1=Yes
SPEND	Should we increase domestic spending?	8	0=No 1=Yes

007M1110
013F2101
137F1001
117M1111
428M3110
017F3101
037M2101

- i) Create a SAS data set, complete with variable labels and value labels or formats for this questionnaire.
- ii) Generate frequency counts for the variables GENDER, PARTY, VOTE, FOREIGN, and SPEND.
- (c) How can you read a *.csv data file in SAS? Explain with example.

3. (a) What's wrong with this program?

```

DATA REGESSN;
INPUT ID DOSAGE EXERCISE LOSS;
DATALINES;
1 100 0 -4
2 100 0 0
3 100 5 -7
(more data lines)
PROC REG DATA=REGESSN;
TITLE 'Weight Loss Experiment- Regression Example';
MODEL LOSS = DOSAGE EXERCISE/P R;
RUN;
QUIT

```

(b) You have two data files, one from the year 1996 and the other from the year 1997, as follows:

File for 1996 (DATA96)

ID	Height	Weight
2	68	155
1	63	102
4	61	111

File for 1997 (DATA97)

ID	Height	Weight
7	72	202
5	78	220
3	66	105

Create a SAS data set from each file. Combine the data from each data set into a single file (call it BOTH).

(c) How can create a permanent data set in SAS? Explain.

4. (a) What are SPSS and SPSS Syntax? Explain the different types of window in SPSS? Write down the use of dialog box in SPSS.

(b) What are the rules for defining variable names in SPSS? Write down the Rules for writing command Syntax in SPSS.

(c) Write an SPSS program to calculate the correlation co-efficient of X and Y and also to calculate regression co-efficient of Y on X .

5. (a) How you open an existing Excel data file in SPSS? How you merge data files in SPSS? Write down the checking procedure for categorical and continuous variables?

(b) Write down the procedure for creating Histogram, Bar graphs, Pie chart, Matrix Scatter plots and Box plots. Also, discuss the procedure for assessing Normality using Explore.

(c) How can you use Journal file to create new program or syntax in SPSS? Discuss the following command file of SPSS and print the results

```

DATA LIST ID 1-2 SEX 4(A) AGE 6-7 OPTION1 TO OPTION4 9-15.
BEGIN DATA
01 M 20 3 2 1 3
02 M 25 1 3 2 1
03 F 32 2 2 1 1
04 M 42 3 1 3 3
05 F 55 2 2 2 1
END DATA.
FREQUENCY ALL.

```

6. (a) What do you mean by Monte Carlo simulation? Write down the difference between Monte Carlo simulation and Stochastic simulation. How is simulation process similar to statistical inference process?

(b) Write the steps of Hit or Miss Monte Carlo Method Algorithm. What is the aim of correlated sampling? In which situations correlated sampling can be employed?

(c) Find the mean and variance of the estimator of the integral $I = \int g(x)dx$ with $0 \leq g(x) \leq c$ using Hit or Miss Monte Carlo method.

7. (a) Describe the algorithm to generate random variates from Geometric distribution and Negative Binomial distribution. Generate 10 random numbers from $Geom(0.3)$.

(b) Let $X \sim \text{Cauchy}(\mu = 10, \sigma = 15)$, using inverse transformation method generate 20 random numbers and find mean and variance.

(c) Give description of Convolution, Acceptance-Rejection general approaches for generating random variates.

8. (a) Describe the Composite Generators or Feedback Shift Register Generators to generate uniform random variates.

(b) Describe a test procedure of uniform random numbers to examine the uniformity.

(c) What is mid-square method to generate uniform random number? What's the limitation of it? Generate 10 random numbers using this method, given $Z_0 = 7182$.

Time: 3.0 Hours

Full Marks: 21

Answer the Following Questions. All Questions Carry Equal Marks.

Q1. Table 6.4 gives the data on child mortality (CM), that is the number of deaths of children under age 5 in a year per 1000 live births, per capita GNP in 1980 (PGNP), female literacy rate in percent (FLR) and total fertility rate (TFR), 1980-1985, the average number of children born to woman, using age specific fertility rate for a given year.

Regress child mortality (CM) on per capita GNP (PGNP). Suppose, you have decided to add female literacy rate (FLR) to the model. Is the incremental contribution of FLR statistically significant?

Q2. Table 7.6 shows the demand for chicken in the states of India for 1992-1993. The data for the following variables are given:

Y = Consumption of chicken in kg

X_2 = Price of chicken in Rs per kg

X_3 = Price of fish in Rs per kg

X_4 = Per capita income in thousands of rupees at 1993-1994 prices

Consider the demand for chicken as a function of the prices of chicken and fish, leaving out the income variable for the time being. Consider the following models:

Linear model: $Y_i = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + u_i$

Log-linear model: $\ln Y_i = \beta_1 + \beta_2 \ln X_{2i} + \beta_3 \ln X_{3i} + u_i$

Which model would you choose between the above two? Justify.

Q3. Table 15.4 gives the data on several families grouped or replicated (repeat observations) according to income level and the number of families owing a house at each income level. Here,

X = Income (thousands of dollar).

N = Number of families at income level X .

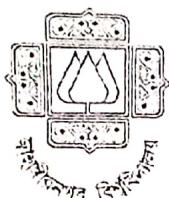
n = Number of families owning a house at income level X .

- 1) Develop a Logit model from the above data. Can you use ordinary least squares (OLS) method for estimating the parameters of this model? Justify your answer. If OLS cannot be used, what method should be used for estimating the parameters of this model?
- 2) Estimate the parameters of this model using appropriate method. Interpret the parameters of this model.
- 3) Compute the probabilities of owning a house and rate of change of probability of owning a house at each income level. Also, interpret them at a particular level of income (say, \$20000).

Q4. Consider the data given in Table 8.11. This table gives data on disposable personal income and gross domestic savings, in crores of rupees, for India for the period 1974-75 to 2004-2005. Suppose, we want to estimate a simple savings function that relates savings to disposable personal income. Suppose, we want to see if a structural change occurred within the data between two time periods. Split the data into three sections: 1974-75 to 1988-89, 1989-90 to 1995-96 and 1996-97 to 2004-05.

- 1) Determine if there is a significant break between the two time periods 1974-75 to 1988-89 and 1989-90 to 2004-05.
- 2) Determine if there is a significant break between the two time periods 1974-75 to 1995-96 and 1996-97 to 2004-05.

Q5. Table 25.1 gives the data of heat evolved in calories per gram of cement (Y) on various amount of tricalcium aluminates (X_2), amount of tricalcium silicates (X_3), amount of tetra calcium alumino ferrite (X_4) and amount of dicalcium silicates (X_5), where X_2, X_3, X_4 and X_5 are measured as percent of the weight of the clinkers from which the cement was made. Fit the ridge regression solution for the above data for a general value of θ and for the straight line model: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + u_i$, and show that the ridge solution fit is: $\hat{Y}_i = 83.41 + 1.3X_{2i} + 0.3X_{3i} - 0.14X_{4i} - 0.35X_{5i}$.



Department of Statistics
Jahangirnagar University
Part II B. Sc (Hons.) Practical Examination-2018
Course #: STAT-LAB-209
Course Title: Statistical Data Analysis III (Group-B)
(Answer the following questions.)

Time: 1.5 hours

Full Marks: 14

Q1. Calculate GFR, ASFR, TFR and GRR for the following data:

Age Group	Female Population	Annual Number of Births
15 – 19	11195	691
20 – 24	10662	1674
25 – 29	8904	1398
30 – 34	7963	843
35 – 39	7817	370
40 – 44	7977	95
45 – 49	6895	9
Total	61413	5080

Number of female live birth is 2464.

Q2. (a) Complete the following section of the life table using given information:

x	q_x	l_x	d_x	L_x	T_x	e_x
0	.0321	100000				
1	.0038					
2	.0008					
3	.0008					
4	.0012					
5	.0045				6513761	

(b) Consider the following information on a complete life table. Fill up the missing entries:

Age interval	nq_x	l_x	d_x	L_x	T_x	e_x^0
0-1	.02593	100,000	--	97,815	6,989,030	--
1-2	--	97,407	165	--	--	70.75
2-3	.00104	--	--	97,192	--	--
3-4	--	97,141	78	--	6,696,699	68.94
4-5	.00667	--	65	97,031	6,599,597	--

Q3. (a) The enumerated population of Bangladesh for the last 5 censuses were as follows:

Census year	Enumerated population (in thousands)
1974	71479
1981	87120
1991	106315
2001	124355
2011	142319

Estimate the intercensal growth rates using exponential formula.

(b) Suppose the intercensal rate of growth for population of 2001 and 2011 is 1.35 percent. How many years would it take the population of 2011 to double itself if this growth rate persists over the doubling time? (Use exponential formula)



Department of Statistics
Jahangirnagar University
Part II B. Sc (Hons.) Practical Examination-2018
Course #: STAT-LAB-202
Course Title: Statistical Data Analysis IV (Group-A)
(Answer any FOUR (04) questions. All questions carry equal marks)

Time: 2 hours

Full Marks: 08⁷ 7

1. The following are the figures of defectives in 20 lots each containing 2000 rubber belts:

425	430	216	341	225	322	280	306	337	305
402	216	264	126	409	193	326	280	389	451

Draw control chart for fraction defective and comment on the state of control of the process.

2. Construct a control chart for mean and the range for the following data on the basis of fuses, samples of 5 being taken every hour (each set of 5 has been arranged in ascending order of magnitude). Comment on whether the production seems to be under control, assuming that these are the first data.

42	19	36	42	51	60	18	15	69
45	25	54	51	74	60	20	30	109
68	80	69	57	75	72	27	39	113
72	81	77	59	78	95	42	62	118
90	81	84	78	132	138	60	84	153

3. From a lot consisting of 2,200 items, a sample of size 225 is taken. If it contains 14 or less defectives, the lot is accepted otherwise rejected. Plot the OC, ATI and AOQ curves. Also obtain the value of A.O.Q.L.

4. Given a test question solved by 10% of a large un-selected group, a second question solved by 20% of the same group and third question solved by 30%, determine the relative difficulty of the questions assuming the capacity measured by the test questions to be distributed normally. Given that with

$$f(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-x^2/2} dx, \quad f(0.52) = 0.7, \quad f(0.84) = 0.8 \text{ and } f(1.28) = 0.9.$$

5. (a) A test is administered on 400 pupils. It gave mean 60 and standard deviation 12. Complete the following table of equivalent raw scores.

Raw score:	84	78	72	66	60	54	48	42	36
σ score:	-	-	1	-	0	-	-	-	-
Standard score:	-	-	-	-	-	45	-	-	-

(b) Convert the ten scores 1, 2, ..., 10 into standard scores with mean 50 and standard deviation 10.



Time: 3 Hours

Department of Statistics
Jahangirnagar University
Part II B. Sc (Honors) Examination - 2018
Course Name: Statistical Data Analysis IV
Course No.: Stat LAB-210

Marks: 28

Answer the following Questions.

Q1. Generate 1000 uniform random numbers using $X_n = (7X_{n-1} + 3) \bmod 1111$, where $X_0 = (5 + \text{last digit of your Exam roll})$ and evaluate the value of the following integrals

$$(i) I = \int_{-2}^2 e^{x+x^2} dx \quad (ii) I = \int_{-\infty}^{\infty} e^{-x^2} dx.$$

Compare your estimate with the exact answer.

Q2. Check the uniformity of the generated uniform random numbers in Question 1.

Q3. Generate $n=10000$ from Cauchy($\mu=3, \sigma=2$), using inverse transformation method and hence find mean and variance. Repeat the process 100 times and store the mean and standard deviation for each time. Plot the 100 means against time. Also, Calculate the average of the mean(s) and standard deviation(s). Finally comment on your results.

Q4. Generate 100 random numbers from $\text{negbin}(s=7, p=0.3)$. From the generated data calculate mean and standard deviation.

Q5. By using Box-Muller transformations draw 1000 random numbers from $N(\mu=17, \sigma=20)$. Also draw 1000 random numbers from $LN(\mu=17, \sigma=20)$. Draw histogram and density curve in each case. Comment on your results. Use SPSS syntax to solve the problem.

Q6. We have recorded the following data from an experiment:

SUBJECT	DOSE	CATEGORY	RECAT	LIVER WT	SPLEEN
1	1	Y	5.4	10.2	8.9
2	1	Y	5.9	9.8	7.3
3	1	N	4.8	12.2	9.1
4	1	Y	6.9	11.8	8.8
5	1	N	15.8	10.9	9.0
6	2	N	4.9	13.8	6.6
7	2	Y	5.0	12.0	7.9
8	2	N	6.7	10.5	8.0
9	2	Y	18.2	11.9	6.9
10	2	Y	55	9.9	9.1

Department of Statistics
Jahangirnagar University
Part II BSc (Honors) Final Examination 2019
STAT 201: Probability Distribution

Time: 2.5 Hours

Full Marks: 35

Answer any THREE questions. All questions carry equal marks.

1. a) Define Poisson distribution with example. Find moment generating function of Poisson distribution and hence or otherwise find the first four moments. Also, find the coefficient of skewness and kurtosis and comment on your result.
b) State the practical applications of the negative binomial distribution. Show that, negative binomial distribution tends to Poisson distribution when $r \rightarrow \infty$, $q \rightarrow 0$, such that $rq \rightarrow \lambda$ (a finite constant).
c) Compare the binomial distribution with negative binomial distribution. If x_1, x_2 are two independent negative binomial variates with parameters (n_1, p) and (n_2, p) then show that $x = x_1 + x_2$ is also a negative binomial variate with parameters (n_1+n_2, p) .
2. a) What is geometric distribution? Is it following lack memory property? Justify your answer.
b) A fast food chain puts a winning game piece on every fifth package of French fries. Find the probability that you will win the prize,
 - (i) with your third purchase of French fries,
 - (ii) with your third or fourth purchase of French fries.
c) Define hypergeometric distribution. Explain how do you use hypergeometric distribution in order to estimate the number of fishes in a pond.
3. a) What is an exponential distribution? Find the moment generating function of exponential distribution. And hence find the mean and variance of exponential distribution.
b) If X_1, X_2, \dots, X_n are independent random variables, X_i having an exponential distribution with parameter θ_i , $i = 1, 2, \dots, n$ then $Z = \min(X_1, X_2, \dots, X_n)$ has exponential distribution with parameter $\sum_{i=1}^n \theta_i$.
c) Define Gamma distribution. Show that it is related to exponential distribution. Find the moment generating function of Gamma distribution $\Gamma(n, p)$.
$$f(x, n) = \frac{x^{n-1} e^{-x}}{\Gamma(n)} \quad (x > 0)$$
 Hence find the probability density function of the associated random variable $Z = \min(X_1, X_2, \dots, X_n)$.

1. Define a gamma distribution. If X_1, X_2, \dots, X_n are independent and identically distributed $\text{Exp}(\lambda)$ variates, then find the distribution of $\sum X_i$.

2. Define a beta distribution. If X_1, X_2, \dots, X_n are independent $\text{Exp}(\lambda)$ variates, then find the distribution of $\sum X_i$.

3. Define a lognormal distribution. When do you use this distribution? Find the median and mode of this distribution.

4. a) Define Beta distribution of first kind and Beta distribution of second kind.
 b) If X and Y are independent Gamma variates with parameters l and m respectively, show that the variates

$$U = X + Y, \text{ and } V = \frac{X}{Y}$$

are independent and that U is a $\gamma(l+m)$ variate and V is a $\beta_2(l, m)$ variate. Also show that $\frac{X}{X+Y}$ is $\beta_1(l, m)$ variate.

c) Define bivariate normal distribution. If X and Y have a bivariate normal distribution with mean and variance $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ respectively and ρ is the correlation coefficient of X and Y . Obtain the moment generating function of bivariate normal distribution.

Good Luck

Department of Statistics
Jahangirnagar University
Part II BSc (Honors) Final Examination 2019
STAT 202: Sampling Distribution

Time: 4 Hours

Full Marks: 70

Answer any FIVE questions. All questions carry equal marks.

1. a) Distinguish between probability distribution and sampling distribution. Discuss the distribution of random variables in case of single random variables and explain the role of sampling distribution in statistics.
b) Explain the concept of Laplace transformation and state the important properties of Laplace transformation
c) If $f(x) = \lambda e^{-\lambda x}$; $x > 0$, find the Laplace transformation, mean, variance, β_1 , β_2 and shape characteristics.
2. a) Explain the concept of variate transformation. Give detail description of (i) Fisher's z-transformation and (ii) orthogonal linear transformation.
b) Write the uses of variate transformation.
c) State and derive the distribution of sample mean.
3. a) State and prove inversion theorem.
b) Find the density function for which c.f. is
$$\Phi(t) = e^{-t^2 \sigma^2 / 2}$$

c) State the central limit theorem. Illustrate the importance of the central limit theorem with an example.
4. a) Explain χ^2 distribution, χ distribution, and non-central χ^2 distribution. Write the properties of χ^2 , χ and χ^2 (non-central) distributions.
b) Derive the χ^2 distribution algebraically.
c) Derive mode m_0 of the χ variate by setting $\frac{d}{dx} \log f(x) = 0$ and $\frac{d^2}{dx^2} \log f(x) < 0$.
5. a) Explain student's t -distribution. Also, mention properties of student's t -distribution. Derive the student's t -distribution geometrically.
b) Show that mode of student's t -distribution is equal to zero and obtain the mean deviation of student's t -distribution with 10 degrees of freedom.
c) Establish the relationship between the central and the non-central t distributions.
6. a) Define F distribution with its properties. Establish the relationship between t , and F distribution.
b) Show that F distribution with degrees of freedom n_1 and n_2 converges to standard normal distribution when $n_1=1$ and $n_2=n$.
c) Discuss the properties of F -distribution. Show that mean of F -distribution is always greater than 1.

Time: 4 Hours Answer any FIVE questions. All questions carry equal marks Full Marks: 70

1. a) What do you mean by regression analysis? How does it differ from the correlation analysis?
 b) Observations Y_1, \dots, Y_n are made according to the model: $Y_i = \alpha + \beta x_i + \varepsilon_i$, where x_1, \dots, x_n are fixed constants and $E(\varepsilon_i) = 0$, $\text{Var}(\varepsilon_i) = \sigma^2$ with $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$ if $i \neq j$. What are the method of moment (MoM) estimators of α and β , denoted by $\hat{\alpha}$ and $\hat{\beta}$ respectively.
 c) Suppose we assume

$$(E(\varepsilon_i), \text{Var}(\varepsilon_i)) = \begin{cases} (0, \sigma^2) & ; \text{ with probability } 1 - \delta \\ (\mu, \tau^2) & ; \text{ with probability } \delta \end{cases}$$
 (i) the MoM estimator of β would be an unbiased estimator of β ?
 (ii) in this case, what is the variance of $\hat{\beta}$?
2. Suppose that we have the summary statistics of the monthly marketing cost (in Euro) (x) and the corresponding profit (in Euro) (Y) of baker: $\sum_{i=1}^{10} x_i = 1880$, $\sum_{i=1}^{10} x_i^2 = 371200$, $\sum_{i=1}^{10} Y_i = 19170$, $\sum_{i=1}^{10} Y_i^2 = 37387300$, and $\sum_{i=1}^{10} x_i Y_i = 3696800$.
 - a) Find the least squares estimates of α and β for a simple linear regression model.
 - b) Estimate the error variance σ^2 .
 - c) Develop 95% confidence interval for $\hat{\alpha}$ and $\hat{\beta}$.
 - d) The statement is that the monthly marketing cost has no influence to the profits. Can you reject that statement at the 5% level?
3. a) Discuss the procedure of estimating the parameters of a three-variable linear regression model by the method of maximum likelihood.
 b) How can you test the normality assumption of the error term in a classical linear regression model with the help of residuals, normal probability plot and Jarque-Bera (JB) test?
 c) Discuss the concept of mean prediction and individual prediction in case of a simple regression model.
4. Consider a bivariate random vector (X, Y) , and iid observations $(X_1, Y_1), \dots, (X_n, Y_n)$ from (X, Y) . The general interest is in a linear regression model

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon$$
 where β is the vector of regression coefficients, say of dimension $p \times 1$, \mathbf{X} is the design matrix, \mathbf{Y} is the column vector of all Y -observations, and ε is the column vector of the error terms $\varepsilon_1, \dots, \varepsilon_n$ for which we assume that they are independent of each other and have the same distribution as the random variable ε . One uses the least squares method to estimate the regression coefficient β .

a) We make the following assumption regarding the random variable ε :

$$E(\varepsilon|X=x) = 0 \text{ and } \text{Var}(\varepsilon|X=x) = \sigma^2 \quad (1)$$

with $0 < \sigma^2 < \infty$, the constant error variance. Derive the bias and variance of the least squares regression estimator $\hat{\beta}$. Discuss which assumption you would further impose on the error variable ε to derive distributional properties of the least square regression estimator $\hat{\beta}$.

b) Assume now that the error random variable is such that

$$E(\varepsilon|X=x) = 0 \text{ and } \text{Var}(\varepsilon|X=x) = \sigma^2(x) \quad (2)$$

with $\sigma^2(\cdot)$ a positive function. Derive the conditional as well as unconditional bias and variance of the least squares regression estimators (by conditionally we mean conditionally upon X_1, \dots, X_n).

c) Consider now the special case of the linear simple regression model: $Y = \beta X + \varepsilon$ i.e., a linear regression model with only a slope parameter (and no intercept parameter). Provide the expression for the variance of the least squares slope parameter $\hat{\beta}$ under both model assumptions (1) and (2). Show that a lower bound for the variance of the least squares slope parameter β under (2) is given by

$$E\left[\left(\sum_{i=1}^n \frac{X_i^2}{\sigma^2(X_i)}\right)^{-1}\right]$$

5. a) Discuss the nature of dummy variables with examples. What are the cautions in the use of dummy variables? Write down the advantages of dummy variable technique over the Chow test. What do you mean by ANOVA and ANCOVA models in case of dummy variables?

b) Consider the model: $Y_t = \alpha_1 + \alpha_2 D_t + \beta_1 X_t + \beta_2 (D_t X_t) + u_t$

where Y = Savings, X = Income, t = Time, $D = 1$, for observations in 1989-90 to 1995-96
 $= 0$, otherwise (Observations in 1974-75 to 1988-1989).

Discuss the procedure of testing the structural stability between the two time periods 1974-75 to 1988-1989 and 1989-90 to 1995-96.

c) Consider a hypothetical data of sales of refrigerator in thousands of units in the first, second, third and fourth quarter. With the help of dummy variable, how can you test the seasonal pattern in this data associated with various quarters? How do you obtain the deseasonalized time series of refrigerator sales in this case? Discuss.

6. a) Explain the concept of linear probability model. Discuss the consequences of this model. What are the alternatives to linear probability model?

b) What do you mean by Logit model? What are the features of this model? How can you estimate the parameters of this model in case of grouped? Under usual notations, suppose that the estimated weighted log-odds is given by:

$$\hat{L}_i = -1.59\sqrt{w_i} + 0.078X_i \quad \left(\begin{array}{l} \text{where, } L_i = L_i \sqrt{w_i} = \ln\left(\frac{p_i}{1-p_i}\right) \sqrt{w_i} \\ \text{and } \sqrt{w_i} = 4.1825 \end{array} \right)$$

Discuss the Logit interpretation, odds interpretation, computing probabilities and computing the rate of change of probabilities in this case.

7. a) What do you mean by residual analysis? How can you check the homoscedasticity of variance of the error term with the help of the plot of residuals against the values of independent variable and against the predicted values of the dependent variable?

b) Distinguish between autocorrelation and serial correlation with examples. Discuss some graphical and mathematical procedure of testing autocorrelation.

8. a) What are the different statistical procedures for selecting the best regression equation? Briefly discuss all possible regression method using three criteria: R^2 , s^2 and the Mallows C_k and stepwise regression and backward elimination procedure for selecting the best regression equation.

b) Suppose we have a data set with dependent variable (Y) and explanatory variables X_1 , X_2 , X_3 and X_4 given by Hald (1952, p.647) with the following information:

k = Number of parameters	R^2 values	Residual mean squares	Average $s^2(k)$	C_k values
2	0.534(1), 0.666(2), 0.286(3), 0.675(4)	115.06(1), 82.39(2), 176.31(3), 80.35(4)	113.53	202.5(1), 142.5(2), 315.2(3), 138.7(4)
3	0.979(12), 0.548(13), 0.972(14), 0.874(23), 0.680(24), 0.935(34)	5.79(12), 122.71(13), 7.48(14), 41.54(23), 86.89(24), 17.57(34)	47.0	2.7(12), 198.1(13), 5.5(14), 62.4(23), 138.2(24), 22.4(34)
4	0.98228(123), 0.98234(124) 0.98128(134), 0.97282(234)	5.35(123), 5.33(124), 5.65(134), 8.20(234)	6.13	3.0(123), 3.0(134), 3.5(134), 7.3(234)
5	0.98237(1234)	5.98(1234)	5.98	5.0(1234)

The values in the parentheses indicate the explanatory variables. (1) means variable X_1 , (2) means variable X_2 , (12) means variables X_1 and X_2 , (13) means variables X_1 and X_3 and so on.

From the above information, select the best regression model on the basis of all possible regression method using three criteria: R^2 , s^2 and the Mallows C_k .

Good Luck

Department of Statistics
Jahangirnagar University
Part II BSc (Honors) Final Examination 2019
STAT 204: Demography

Time: 2.5 Hours

Full Marks: 35

Answer any THREE questions. All questions carry equal marks.

1. a) Discuss the scope and importance of demography.
b) Describe the primary sources of demographic data.
c) Explain the importance of age and sex composition of population data.
d) Explain the terms: coverage error, under enumeration, age heaping, and dependency ratio.

2. a) Define crude birth rate with advantages and disadvantages.
b) Differentiate between fertility and secundity.
c) Define general fertility rate and total fertility rate.
d) Discuss the factors affecting fertility rates in Bangladesh.

3. a) What are the meaning of mortality and morbidity? Define neonatal, postneonatal, infant, and child mortality.
b) What are the major causes of infant mortality? What are the factors that lead to reduce the infant mortality dramatically in Bangladesh?
c) What do you mean by standardization of death rates?
d) Explain in detail the direct and indirect method of standardization.

4. a) What is sex ratio? Write down the importance of sex ratio.
b) Distinguish between stable population and stationary population.
c) Derive the Lotka's model developed for stable population theory.
d) What is mid-year population? How is it calculated? Define mean length of generation.

5. a) Distinguish between gross reproduction rate and net reproduction rate.
b) What is life table? How did the concept of life table originate? What purpose does a life table serve? How does a complete life table differ from an abridged life table?
c) What are the underlying assumptions in the construction of a life table for representing the actual experience of an observed population? Elaborate them.
d) Explain different functions of life table and their interrelation.

*** Good Luck***

Department of Statistics
Jahangirnagar University
Part II BSc (Honors) Final Examination 2019
STAT 205: Quality Control and Applied Statistics

Time: 2.5 Hours

Full Marks: 35

Answer any THREE questions. All questions carry equal marks.

1. a) What is quality and statistical quality control? Write down the major factors and processes of production which quality control mainly covers.
b) Define causes of variation and write down the differences between chance and assignable causes of variations.
c) Write down the six elements to a cybernetic or self-regulating system in statistical quality control.

2. a) What is a control chart? Give example. Explain the basic ideology underlying the control charts. Discuss the function of control charts in manufacturing processes.
b) Discuss construction procedure for mean control charts and range control charts. Explain briefly the different criteria for detecting lack of control in \bar{X} and R charts. Discuss different situations where R chart is in a state of control but \bar{X} chart is not. Interpret the points falling outside the control limits on these charts.
c) Distinguish between (i) quality characteristics and quality engineering (ii) defective and nonconforming (iii) target value and nominal value.

3. a) What is single sampling plan? Describe briefly. What are the differences between single sampling plan and double sampling plan?
b) Distinguish clearly between: (i) Producer's and Consumer's risk, (ii) A. Q. L. and L. T. P. D.
c) What do you mean by acceptance sampling plan? Discuss its utility with examples. Define control chart for fraction defective (p-chart) and number of defects per unit (c-chart).

4. a) Briefly discuss the importance of studying educational statistics. What is z-score and T-score? How does z-score differ from the standard score?
b) State different methods of testing reliability? Explain the Split-half method with its relative advantages and disadvantages.
c) Convert the scores 2, 4, 6, ..., $2n$ into standard scores with mean 50 and standard deviation 10 where $9 < n < 16$.

5. a) What is official statistics? What are the sources of official statistics in Bangladesh? Mention them with their at least three publications.
b) Define the fundamental principles of official statistics and also write down the system of collection of official statistics.
c) What are the factors that affect the quality of official statistics? What steps should be taken to improve it?

*** Good Luck***

Department of Statistics
 Jahangirnagar University
 Part II BSc (Honors) Final Examination 2019
 STAT 206: Microeconomics and Economic Statistics

Time: 2.5 Hours

Full Marks: 35

Answer any THREE questions. All questions carry equal marks.

1. a) What is aggregate demand (AD)? What are the components of AD? Why does the AD curve slopes download to the right?
 b) Examine with the diagram what will happen to the short run equilibrium of the economy if the government adopts expansionary fiscal policy.
2. A consumer's basket of goods for a hypothetical economy includes 7 Pizzas, 9 Pepsis and 10 Hamburgers. Use the Table below to answer the following questions. Take 2016 as the base year.

Year	Price of Pizza (Tk.)	Price of Pepsi(Tk.)	Price of Hamburger(Tk.)
2015	130	70	110
2016	135	85	120
2017	140	95	140

- a) Calculate the cost of consumer's basket for each of the years given.
 b) Calculate the CPI for each of the years given.
 c) From the CPI, calculate the inflation rate for the year 2016 and 2017.
 d) Which one is a better measure of inflation, CPI or GDP deflator? Explain.

3. Suppose in a hypothetical economy, three goods are produced: Mango, Banana and Chicken. The quantity of output and prices of mango, banana and chicken are given in the following table:

	Output			Price, Tk./Unit		
	2014	2015	2016	2014	2015	2016
Mango	50	55	65	35	40	45
Banana	60	65	75	6	7	8
Chicken	31	35	40	110	125	150

- a) Define nominal and real GDP. Calculate the nominal GDP for the year 2014, 2015 and 2016. Suppose, 2014 is the base year, and then calculate the real GDP for the year 2014, 2015 and 2016. Define growth rate. Calculate the growth rate of real GDP for the year 2015 and 2016.
- b) Define GDP deflator. What does it measure? According to the GDP deflator, how much prices have increased for the year 2015 and 2016.
4. a) What is unemployment? Briefly explain the different types of unemployment.
 b) The Bangladesh Bureau of Statistics in 2016 reported that the total population was 160 millions of which 80% were adult. The adults Not in labor force were 25 millions and 23 millions were unemployed. What was the labor force participation rate? What was the unemployment rate?
5. a) What is money? Explain the types of money? Explain the tools of expansionary monetary policy.
 b) Define inflation. Explain demand-pull and cost-push inflation in brief.

Department of Statistics
Jahangirnagar University
Part II BSc (Honors) Final Examination 2019
STAT 207: Mathematical Analysis and Differential Equations

Time: 4 Hours

Full Marks: 70

Answer any FIVE questions. All questions carry equal marks.

1. a) Why are differential equations useful? Explain briefly.
b) Classify each differential equation as separable, linear, homogenous or Bernoulli. Some equations may be one kind. Solve them using appropriate transformation:
 - (i) $(x^2 + 1) \frac{dy}{dx} + 4xy = x$ (ii) $(x^2 + 3y^2)dx + 2xydy = 0$ (iii) $\frac{dy}{dx} = 2 + \sqrt{y - 2x + 3}$
 - c) Is the equation $(3y + 4xy^3)dx + (2x + 3x^2y)dy = 0$ exact? If it is exact, solve it; otherwise solve it by finding appropriate integrating factor.
 - d) Consider the initial value problem $y' = 2x + y$; $y(0) = 1$. Use Euler's method to obtain an approximation of $y(0.6)$ use step size $h = 0.05$.
2. a) Determine the order and degree of the following D.E.
(i) $\frac{d^2y}{dx^2} + 3 \frac{dy}{dx} - 2y = 0$, (ii) $\frac{d^2y}{dx^2} + 3\left(\frac{dy}{dx}\right)^2 + y = x$; which one of the above is linear?
Give reasons in support of your answer.
b) Show that the differential equation of the family of circles $x^2 + y^2 + 2gx + 2fy + c = 0$ is $(1 + y_1^2)y_1 - 3y_1y_2 = 0$.
c) What is homogeneous differential equation? Solve the following D.E.
$$x \cos\left(\frac{y}{x}\right)(ydx + xdy) = y \sin\left(\frac{y}{x}\right)(xdy - ydx).$$

d) Solve the equation:
$$3e^x \tan y dx + (1 - e^x) \sec^2 y dy = 0.$$
3. a) Find (i) $L(xe^{4x})$ (ii) $L^{-1}\left\{5s - 3\frac{e^{-3s}}{s} - 2\frac{e^{-2s}}{s}\right\}$ (iii) $L^{-1}\left\{\frac{1}{s^2 + 7}\right\}$
b) Solve $y'' - 3y' + 2y = e^{-x}$; $y(0) = 1$, $y'(0) = 5$ using Laplace transformation method.
c) Let $f(x) = \begin{cases} 2 & 0 \leq x < 3 \\ -2 & x \geq 3 \end{cases}$. Solve the equation: $y'' + y = f(x)$.
4. a) Find the general solution of the Bessel equation: $x^2 y'' + xy' + (x^2 - 1)y = 0$.
b) Mention some properties of Bessel's function of order m , $m = 0, 1, 2, \dots$
c) Show that $J_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \cos x$ and $Y_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \sin x$
where $J_{\frac{1}{2}}(x)$ Bessel's function of first kind of order $-\frac{1}{2}$ and $Y_{\frac{1}{2}}(x)$ Bessel's function of second kind of order $-\frac{1}{2}$ respectively.
d) What do you mean by Legendre polynomials?

5. a) Explain the graphical representation of complex numbers. Write the vector interpretation of the complex numbers.

b) Let z be a complex number, then represent iz geometrically.

c) Express the following complex number in polar form: (i) $-5 + 5i$ (ii) $-2i$.

d) Define harmonic function and write the Cauchy-Riemann equation. Prove that $u = 2x - x^3 + 3xy^2$ is harmonic. Find v such that $f(z) = u + iv$ is analytic.

6. a) What do you mean by simply and multiply connected domains?

b) Give two Laurent series expansions in the power of z for the function

$$f(z) = \frac{1}{z^2(1-z)}$$

and specify the regions in which those expressions are valid.

c) Find Laurent series about the indicated singularity for each of the following functions:

$$(i) \frac{e^{2z}}{(z-1)^3}; z=1 \quad (ii) \frac{z-\sin z}{z^3}; z=0$$

Give the name of the singularity in each case.

d) State Cauchy integral formula. Show that $\int \frac{z dz}{(9-z^2)(z+i)} = \frac{\pi}{5}$ where C be the positively oriented circle $|z| = 2$.

7. a) Evaluate $\oint \frac{\sin \pi z^2 + \cos \pi z^2}{(z-1)(z-2)} dz$, where C is the circle $|z| = 3$.

b) State the Residue theorem. Find the residue of the following function at all its poles in the finite plane (i) $f(z) = \frac{z^2 - 2z}{(z+1)^2(z^2+4)}$ (ii) $f(z) = \frac{e^z}{\sin^2 z}$.

8. a) State why the function $f(z) = xy^2 + i(2x - y)$ is entire.

b) The population of a town grows at a rate proportional to the population present at time t . The initial population of 500 increases by 15% in 10 years. What will be the population in that town in 30 years?

c) Solve $\frac{dy}{dx} + y = f(x)$, $y(0) = 0$ where $f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases}$.

d) Consider the function $f(z) = (z+1)^2$ and the closed triangular region R with vertices at $z = 0, z = 2, z = i$. Find the points in R where $|f(z)|$ has its maximum and minimum values.

Good Luck

Department of Statistics
Jahangirnagar University
Part II BSc (Honors) Final Examination 2019
STAT 208: Statistical Simulation and Data Processing

Time: 4 Hours

Full Marks: 70

Answer any FIVE questions. All questions carry equal marks.

1. a) Explain the term simulation and different types of simulation with example.
b) Explain the system and different types of system. Also explain with example different component of system.
c) Generate 15 uniform random numbers using $X_n = (11X_{n-1} + 3) \bmod 31759$, where $X_0 = 13$ and evaluate the value of $I = \int_0^2 e^{3x^2-2x+3} dx$. Compare your estimate with the exact answer.
2. a) What is random number? What are the different methods to generate random numbers? What are the different tests for a good random number?
b) What is cycle of length? Verify that the following generators are full cycle or not:
(i) $X_{i+1} = 5X_i \bmod 5$; (ii) $X_{i+1} = 11X_i \bmod 32$; (iii) $X_{i+1} = (7X_i + 2) \bmod 24$
c) What is spectral test? Describe about serial correlation of Linear Congruential Generators.
3. a) Explain the inverse transformation approaches to generating random variates.
b) Describe the algorithm to generate random variates from Poisson distribution and Binomial distribution. Generate 15 random numbers from *Discrete Uniform* ($a=3, b=9$).
c) Derive the equations using the Polar method for generating normal variables.
4. a) What do you mean by Monte Carlo simulation? What are the difference and similarity between Hybrid simulation and Monte Carlo simulation?
b) What are the advantages, disadvantages, and pitfalls of simulation?
c) Write down the algorithm to generate random variate from Geometric distribution and Negative Binomial distribution.
5. a) What is meant by SAS? What are the different types of window used in SAS? Write down different rules for the variable names in SAS?
b) Discuss the terms "PROC FREQ", "PROC MEANS", and "PROC UNIVARIATE" to creating summary of Data Sets in SAS. Which statistic(s) are found by PROC UNIVARIATE for a data set?
c) What's wrong with the following program? There may be missing values for X, Y, and Z.

```
1 DATA MANY-RRR;
2  INPUT X Y Z;
3  IF X LE 0 THEN X=1;
4  IF Y LE 0 THEN Y=1;
5  IF Z LE 0 THEN Z=1;
6  LOGX=LOG(X);
7  LOGY=LOG(Y);
8  LOGZ=LOG(Z);
9 DATALINES;
1  2  3
.  7  8
4  .  10
7  8  11
;
10 PROC CORR DATA=MANY-RRR / PEARSON SPEARMAN;
11      VAR X ~ LOGZ;
RUN;
```

6. a) What is the application of ARRAY in SAS? How can you add value label and variable label in a SAS data set? Explain with example.

b) How can you read a *.txt data file in SAS? Explain with example.

c) Consider two SAS data sets one containing the employee ID with name and salary and another containing employee ID with employee ID and department.

Data set 1			Data set 2	
ID	NAME	SALARY	ID	DEPT
1	RICK	623.3	1	IT
2	DAN	515.2	2	OPS
3	MIKE	611.5	3	IT
4	RYAN	729.1	4	HR
5	GARY	843.25	5	FIN

i) Write a SAS code to enter the above two datasets.

ii) Merge these two datasets by ID to get the complete information for each employee and calculate the average salary.

iii) Write down the SAS output.

7. a) What do you mean by SPSS and SPSS Syntax? What are the different types of window in SPSS? Write down the uses of dialog box in SPSS.

b) Suppose a data set 'DATA.xls' is provided where Column 1 represents ID (Identity), Column 2 represents Year (I =Part I, II =Part II, III =Part III, IV =Part IV), Column 3, Column 4, and Column 6 represent Height, Weight, and Age respectively, and Column 5 represents Gender (0=Male, 1=Female). Also given Column 1 and Column 5 are nominal, Column 2 is ordinal, and rests of the Columns are continuous. The data file is in following –

1	I	63.8	61	1	22.2
2	II	66	69	0	21.6
3	III	68.8	69.6	1	22.5
4	IV	60.1	64.1	0	21.1
⋮	⋮	⋮	⋮	⋮	⋮
17	I	67.1	68.2	0	23.5
18	IV	69.7	64.7	0	20.4
19	IV	63	67.5	1	22.4
⋮	⋮	⋮	⋮	⋮	⋮
n	IV	63.5	67.8	1	22.7

Write down the SPSS syntax for the following operations-

- Read the Excel Data file from Drive D.
- Rename the Variables to ID, YEAR, HEIGHT, WEIGHT, GENDER, and AGE.
- Add variable labels and value labels.
- Change variables alignment to Center.
- Set the measurement level for all variables.
- To find the frequency for YEAR and GENDER with Bar Charts
- To find Descriptive Statistics for HEIGHT, WEIGHT, and AGE.
- To find Cross tabulation between YEAR and GENDER.
- Save the data file to Drive E.

8. a) What is SPSS portable file? How to load an existing Excel file in SPSS? How can you determine if the values of the summary statistics for a variable differ for subgroups of cases?

b) How to define variables in SPSS? Write down the general rules of establishing variable names in SPSS.

c) Write a SPSS command syntax file that contains both commands and inline data. Add variable label of id, gender and age as "id of the student", "gender of the student", and "age of the student" respectively. Add value label of gender 0 "female" and 1 "male".

id	gender	age	op1	op2
1	0	29	2	1
2	1	56	1	1
3	0	34	2	2
329	1	23	2	1

*** Best of Luck***

Subject Code:STAT-201



DEPARTMENT OF STATISTICS

Jahangirnagar University

BSc Final Examination - 2020

Subject: Probability Distribution

Course No. 201

Time: 45 mins.

Full Marks: 10

Date: 02.01.2022

Answer any four questions from the following. Each question carries equal marks.

1. (a) The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that (a) at least 10 survive, (b) from 3 to 8 survive, and (c) exactly 5 survive?
(b) Explain the meaning and execution procedure of each line of the following PYTHON Code.

```
from scipy.stats import binom
import seaborn as sb
binom.rvs(size=10, n=20, p=0.8)
data_binom = binom.rvs(n=20, p=0.8, loc=0, size=1000)
ex = sb.distplot(data_binom,
kde=True,
color='blue',
hist_kws={"linewidth": 25, 'alpha':1})
ex.set(xlabel='Binomial', ylabel='Frequency')
prob1 = binom.cdf(k=8, n=20, p=0.8)
prob2 = binom.pmf(k=10, n=20, p=0.8)
print(prob2)
```

2. (a) What is a Poisson variate? Give two examples of Poisson variate.
(b) Verify that

$$f(x, \mu) = \frac{e^{-\mu} \mu^x}{x!} ; \quad x = 0, 1, 2, \dots, \infty$$

is a probability function.

3. (a) What is the concept of negative binomial distribution? From which distribution negative binomial distribution is originated?
(b) A drug is known to be effective in 60 percent of the cases in curing a patient from a severe pain. Find the probability that the fifth patient to experience relief from pain is the seventh patient to receive such drug during a given period of time.

4. (a) The finished inside diameter of a piston ring is normally distributed with a mean of 10 centimeters and a standard deviation of 0.03 centimeter. What is the probability that a piston ring will have an inside diameter between 9.97 and 10.03 centimeters?
(b) Suppose that telephone calls arriving at a switchboard follow a Poisson process with an average of 5 calls coming per minute. What is the probability that up to a minute will elapse until 2 calls have come in to the switchboard?

5. (a) The length of time for one individual to be served at a cafeteria is a random variable having an exponential distribution with a mean of 4 minutes. What is the probability that a person is served in less than 3 minutes on at least 4 of the next 6 days?
(b) Define a bivariate exponential distribution. Mention important properties of bivariate exponential distribution.

Subject Code:STAT-202



DEPARTMENT OF STATISTICS

Jahangirnagar University

BSc Final Examination - 2020

Subject: Sampling Distribution

Course No. 202

Time: 45 mins.

Full Marks: 10

Date: 06.01.2022

Answer any four questions from the following. Each question carries equal marks.

1. (a). Differentiate between sampling distribution and probability distribution with an example.

(b). Define characteristic function (c.f.). Find the density function whose c.f. is

$$\varphi(t) = (q + pe^{it})^n,$$

under usual notations.

2. (a). Define order statistics with an example. How would you obtain the joint distribution of the n order statistics?

(b). What are the properties of Laplace transformation. Find the Laplace transformation of $f(x)$ where $f(x) = \sin(ax)$, $x > 0$.

3. (a). Differentiate between weak law of large number and strong law of large number.

(b). A random variable X has the following probability distribution:

Values of X : x	1	2	3	\dots	N
$\Pr(X = x)$	$\frac{1}{N}$	$\frac{1}{N}$	$\frac{1}{N}$	\dots	$\frac{1}{N}$

Let X_1, X_2, \dots, X_n be a random sample of size n from the above probability distribution and $S_n = X_1 + X_2 + \dots + X_n$. Check whether $\bar{X} = \frac{S_n}{n}$ follows laws of large number.

4. (a). Define χ^2 distribution and mention its important applications.

(b). If t follows Student t distribution with n degrees of freedom, then show that $y = t^2/n$ follows beta distribution of second kind.

5. (a). What is non-central F distribution? If F' follows non-central F distribution, then find the distribution of $Z = \frac{n_1}{n_2} F'$.

(b). If $\chi'^2_1, \chi'^2_2, \dots, \chi'^2_k$ follows non-central χ^2 distribution with df n_1, n_2, \dots, n_k and non-centrality parameter $\lambda_1, \lambda_2, \dots, \lambda_k$, then find the sampling distribution of $\sum_{i=1}^k \chi'^2_i$.

Subject Code:STAT-203



DEPARTMENT OF STATISTICS

Jahangirnagar University

Part II B.Sc.(Hons.) Examination - 2020

Subject: Regression Analysis

Course No. 203
Time: 45 mins.
Full Marks: 10
Date: 13.01.2022

Answer any four questions from the following. Each question carries equal marks.

1. What do you mean by regression analysis? What are the underlying assumptions of classical linear regression model? What do you mean by correlation analysis? Write down the properties of simple correlation coefficient.
2. Write down the normality assumption of the error term in a classical linear regression model? Discuss some graphical and mathematical procedures of testing the normality assumption of the error term in a classical linear regression model.
3. What do you mean by dummy variables? What are the cautions in the use of dummy variables? What are the advantages of dummy variable technique over the Chow test?
4. What do you mean by binary response regression model? What are the available approaches for developing a probability model for a binary response variable? What is linear probability model? Discuss the several problems associated with it.
5. How can you test the incremental or marginal contribution of an explanatory variable in a regression model? How can you choose an appropriate regression model between a linear regression model and a log-linear regression model?

Subject Code:STAT-204

DEPARTMENT OF STATISTICS

Jahangirnagar University

Part II B.Sc.(Hons.) Examination - 2020

Course Name: Demography

Course No. 204

Time: 45 mins.

Full Marks: 10

Date: 19.01.2022

Answer any four questions from the following. Each question carries equal marks.

1. (a) Differentiate population studies from demography.
(b) Discuss the importance of age and sex composition in policies.
2. (a) What do you mean by assessment of quality of demographic data?
(b) What is coverage error? Write down the causes of coverage error.
3. (a) Write down the concept of De facto and De Jure population with examples.
(b) In which situations erroneous enumeration occurs? Explain shortly. Define net over count enumeration and net under count enumeration.
4. (a) What is the basic difference between fertility rate and birth rate? Distinguish between mean age at child bearing and mean length of generation.
(b) Show relationship among TFR, GRR and NRR with examples
5. (a) Distinguish between cohort life table and period life table.
(b) Under usual notations show that

$$\mu_x = \frac{d_{x-1} + d_x}{2l_x}$$

W6m² t or 2-71
Mortality count One number
or? life birth 15-49
mortality absent

End of exam questions

Subject Code:STAT-205

DEPARTMENT OF STATISTICS



Jahangirnagar University

Part II B.Sc.(Hons.) Examination - 2020

**Course Name: Quality Control and
Applied Statistics**

Course No. 205

Time: 45 mins.

Full Marks: 10

Date: 26.01.2022

Answer any four questions from the following. Each question carries equal marks.

1. Briefly differentiate between quality and statistical quality control. What are the important elements to a cybernetic or self-regulating system in statistical quality control? What are the causes of variation?
2. Define control charts along with its basic thoughts. Discuss the function of control charts in industrial processes. In \bar{x} and R charts, write down the criterion for detecting the lack of control.
3. Discuss how the Lorenz curve and Gini concentration ratio are applied for measuring income inequality.
4. Define operating characteristics (OC) curve for a single sampling plan. If the probability of acceptance (P_a) = 0.892 and fraction defective (P) = 0.005. What is your comment on these results?
5. Considering the educational and psychological experiments define reliability and validity scores. For any frequency distribution write down the procedure for obtaining T-scores. What is psychological scale?

Subject Code:STAT-206

DEPARTMENT OF STATISTICS



Jahangirnagar University

Part II B.Sc.(Hons.) Examination - 2020

Course Name: Macroeconomics and
Economic Statistics

Course No. 206
Time: 45 mins.
Full Marks: 10
Date: 01.02.2022

Answer any four questions from the following. Each question carries equal marks.

1. Below are some data on two goods produced in the economy. Compute nominal GDP, real GDP and growth rate each year, using 2018 as the base year.

Year	Price of rice (/kg)(Tk)	Quantity of rice (kg)	Price of cloth (/m)(Tk)	Quantity of cloth (/m)
2018	40	200	140	150
2019	45	350	150	200
2020	50	300	200	300

2. A typical consumer of Dhaka city buys 5 hotdogs and 3 hamburgers. Use the table below to answer the following questions. Take 2017 as the base year.

Year	Price of Hotdogs (Tk)	Price of Hamburgers (Tk)
2017	50	80
2018	55	85
2019	58	90

Calculate CPI for the years 2017, 2018 and 2019. Also calculate inflation rate for the year 2018 and 2019.

3. Define aggregate demand(AD) and aggregate supply (AS) curve. Why the aggregate demand is downward sloping? Explain any one reason. Why the short-run aggregate supply (SRAS) curve is upward sloping? Explain any one reason.
4. The BBS in 2018 reported that the total population was 180 millions of which 75% were adult. The adults Not in labor force were 30 millions and 25 millions were unemployed. What was the labor force participation rate? What was the unemployment rate?
5. Explain the tools to control money supply

Subject Code:STAT-207

DEPARTMENT OF STATISTICS



Jahangirnagar University

Part II B.Sc.(Hons.) Examination - 2020

Course Name: Mathematical Analysis

and Differential Equations

Course No. 207

Time: 45 mins.

Full Marks: 10

Date: 07.02.2022

Answer any four questions from the following. Each question carries equal marks.

1. (a) Classify each differential equation as separable, linear, homogeneous or Bernoulli:

$$(i) (x^2 + 1) \frac{dy}{dx} + 4xy = x \quad (ii) \frac{dy}{dx} = 2 + \sqrt{y - 2x + 3}$$

(b) Is the equation $xydx + (2x^2 + 3y^2 - 20)dy = 0$ exact? If it is exact, solve it.

2. Suppose $\int_{(0,3)}^{(2,4)} (2y + x^2) dx + (3x - y) dy$ and calculate the line integral of straight line from $(0,3)$ to $(2,4)$.

3. Solve the differential equation : $\frac{d^2y}{dx^2} - 2k \frac{dy}{dx} + k^2y = e^x$.

4. Evaluate: (i) $L(1 + 5t)$ (ii) $L^{-1}\left\{\frac{1}{s^2+7}\right\}$.

5. Evaluate the integral: $\oint_C \frac{e^{-iz}}{(z+3)(z-i)^2} dz$;
 $c = \{z : z = 1 + 2e^{i\theta}; 0 \leq \theta \leq 2\pi\}$.

Subject Code:STAT-208



DEPARTMENT OF STATISTICS
Jahangirnagar University
Part II B.Sc.(Hons.) Examination - 2020
Course Name: Statistical Simulation
and Data Processing

Course No. 208
Time: 45 mins.
Full Marks: 10
Date: 13.02.2022

Answer any four questions from the following. Each question carries equal marks.

1. (a) What is meant by the term simulation? Discuss different types of simulation briefly with example.

(b) Generate 15 uniform random numbers using $X_n = (9X_{n-1} + 5) \bmod 746515$, where $X_0 = \text{Last five digits of your Examination Roll}$ and evaluate the value of $I = \int_{-1.5}^{3.0} \frac{2}{5} e^{-\frac{3x^2}{6}} dx$. Compare your estimate with the exact answer.
2. (a) What do you mean by cycle length of Linear Congruential Generators (LCG)? Test the uniformity of the following random numbers: 0.44, 0.81, 0.14, 0.05, 0.93 at 5% level of significance by using Kolmogorov-Smirnov test.

(b) Describe the algorithm to generate random numbers from Discrete Uniform distribution and Geometric distribution.
3. (a) What is meant by SAS? What are the different types of SAS window? What are the rules for the variable names in SAS? How can you import a CSV File into SAS with header and without header?

(b) What is wrong with the following problem?

```
DATA TEST;  
INPUT AGE X Y;  
CARDS;  
10 3 5  
15 3 9  
20 4 8  
22 4 8  
25 8 3  
;  
RUN;  
PROC CORR DATA = TES;  
VAR X y;  
PARTIAL age,  
RUN.  
PROC REG DATA=TEST,  
MODEL Y = AGE;  
RUN.
```
4. (a) Show the derivation and derive the equations using the Polar method for generating normal variables.

(b) Suppose you are comparing the time to relief of three headache medicines: brand 1, 2 and 3. The time to relief data is reported in minutes. For this experiment, 15 subjects were randomly placed on one of the three medicines. The data for this example are as follows:

Brand 1	Brand 2	Brand 3
24.5	28.4	26.1
23.5	34.2	28.3
26.4	29.5	24.3
27.1	32.2	26.2
29.9	30.1	27.8

Write a SAS program to enter these data, perform one-way ANoVA, produce means for brands and plot the data. Notice that SAS expects the data to be entered as two variables, a group and an observation.

5. (a) Discuss the different rules defining variable names and the rules for writing command syntax in SPSS. Mention the SPSS syntax to solve the system of linear equations.

$$3x + 5y + 2z = 2$$

$$2x + 3y + 3z = 5$$

$$x + y + z = 1$$

(b) Suppose a data set 'DATA.xlsx' is provided where column 1 represents ID (Identity), column 2 represents Year (I =Part I, II =Part II, III =Part III, IV =Part IV), column 3, column 4 and column 6 represent Height, Weight and Age respectively. And column 5 represents Gender (0=Male, 1=Female). Also given column 1 and column 5 are nominal, column 2 is ordinal and rests of the columns are continuous. The data file is as following

1	I	67.1	68.2	0	23.5
2	II	69.7	64.7	0	20.4
.
.
.
19	IV	65	67.5	1	22.4
.
.
.
n	IV	83.5	67.8	1	82.7

Write down the SPSS syntax for the following operations.

- Read the Excel data file from Drive D. Rename the variables to ID, YEAR, HEIGHT, WEIGHT, GENDER and AGE.
- Add variable labels and value labels.
- Change variables alignment to center.
- Set the measurement level for all variables.
- To find the frequency for YEAR and GENDER with bar charts.
- To find descriptive statistics for HEIGHT, WEIGHT, and AGE.
- To find cross tabulation between YEAR and GENDER.
- To save the data file to Drive E.

Subject Code:STAT LAB-209(Group-B)

Problem 1: Using the following information compute the general fertility rate (GFR), age-specific fertility rate (ASFR), total fertility rate (TFR) and gross reproduction rate (GRR). The sex ratio at birth is 105 male birth to 100 female birth. Comment on your results.

Table: Calculation of Gross Reproduction Rate (GRR).

(5)

Age Group	Number of Women	Total Number of reported birth B_i
15-19	10427	699
20-24	10702	1616
25-29	9195	1289
30-34	8313	817
35-39	7430	321
40-44	7847	80
45-49	7630	10
Total		

Problem 2: The following Table shows the enumerated population of 2011 census of Bangladesh in single years by sex. (8)

Age	Sex		Age	Sex	
	Male	Female		Male	Female
23	878	1073	43	334	309
24	945	1081	44	282	260
25	2193	2979	45	1983	1907
26	1055	1170	46	303	241
27	931	966	47	271	213
28	1521	1580	48	576	455
29	525	560	49	231	200
30	2647	2996	50	1957	1845
31	430	433	51	213	166
32	1092	1056	52	397	286

Scanned with CamScanner

33	511	527	53	209	165
34	340	409	54	176	137
35	2463	269	55	1173	1067
36	535	527	56	203	136
37	424	402	57	164	105
38	908	880	58	255	175
39	367	364	59	128	93
40	2596	2515	60	1518	1423
41	338	294	61	127	98
42	731	602	62	211	160

Compute the Whipple's Index by taking 5 years age interval for both male and female and comment on your results

Problem 3: The Health and Demographic Surveillance System (HDSS) , Matlab In 2009 reported the age-specific fertility rates as shown in the accompanying Table along with the distribution of the female population by age as recorded in the 2007 Bangladesh Demographic and Health Survey (BDHS). Calculate the standardized CBR and GFR from the given data using the BDHS 2007 female age distribution as standard (Total population given in BDHS=48749). (3)

Age Group	Female Population W_i	Age-specific fertility rate f_i
15-19	11195	0.0617
20-24	10662	0.1570
25-29	8904	0.1570
30-34	7963	0.1059
35-39	7817	0.0473
40-44	7977	0.0119
45-49	6895	0.0013
Total		

Scanned with CamScanner

Scanned with CamScanner

Department of Statistics
Jahangirnagar University
Part II B. Sc. (Hons.) Examination 2020
Course Title: Statistical Data Analysis IV
Course No: STAT LAB 210 (Group A)

Time: 1 hour

Full Marks: 07

(Answer all questions. All questions carry equal marks)

1. The net weight (in oz.) of a dry bleach product is to be monitored by \bar{x} and R control charts using a sample size of $n=5$. Data for 20 preliminary samples are as follows:

Sample number	Mean	Range	Sample number	Mean	Range
1	16.18	0.7	11	16.24	0.5
2	16.14	0.5	12	16.38	0.8
3	16.3	0.4	13	16.32	0.5
4	16.2	0.5	14	16.34	0.3
5	16.22	0.5	15	16.24	0.3
6	16.32	0.9	16	16.2	0.3
7	16.3	0.4	17	16.3	0.2
8	16.18	0.2	18	16.24	0.5
9	16.34	0.3	19	16.3	0.4
10	16.38	0.5	20	16.22	0.7

(a) Set up \bar{x} and R control charts using these data. Does the process exhibit statistical control?
 (b) Estimate the process mean and standard deviation.
 (c) If the specifications are at 16.2 ± 0.5 , what conclusion would you draw about process capability?
 (d) What fraction of containers produced by this process is likely to be below the lower specification limit of 15.7 oz.?
 2. From a lot consisting of 100 items, a sample of size 20 is taken. If it contains 2 or less defectives, the lot is accepted, otherwise rejected. Calculate ATI of the process using $p=0.01, 0.05, 0.25$ and 0.8.
 3. In a distribution with mean 86 and $\sigma = 15$, X's score is 91 and Y's score is 83. Express these scores in standard scores in a distribution with mean 500 and $\sigma = 100$.