Westminster International University in Tashkent

School of Business and Economics

Module Code: 4ECON006C_n
Introduction to Statistics and Data Science Examination

Thursday, 14th December 2023 — 10:00 – 12:15 or 15:00 - 17:15

Venue: WIUT (Sports Hall or IB 301) — **Time allowed:** 2 hours and 15 minutes

---

The following items are provided:

One Examination Question Paper (6 printed pages in total inclusive of cover page) and one Answer Book (8 printed pages in total inclusive of cover page)

Candidates are permitted to bring into the examination room:

- Any type of calculator
- Black or blue pen
- Ruler

Instructions:

Candidates must answer all questions.

- The paper consists of Section A (15 questions, 30 marks) and Section B (5 tasks, 70 marks).
- All answers must be written in the answer book(s) provided.
- All steps of solutions for Section B must be written in the answer book(s) provided.
- A line should be drawn through any rough work to indicate to the examiner that it is not part of the work to be graded.
- Black or Blue pen only must be used for written answers and for all drawings and sketches. The use of pencils or erasable pens is not permitted.
- Formula sheet and tables are attached at the back of this exam booklet (starting page 5).

Ensure you have the correct number of pages on your examination question paper. **Note to Candidates: Please check the Module Code and Module Title to ensure that you are attempting the correct examination. If in doubt, please contact the Invigilator.**

DO NOT TURN OVER THIS PAGE UNTIL INSTRUCTED TO DO SO BY THE INVIGILATOR

Topics covered

Week 1: Introduction to Statistics & Descriptive Statistics
Week 2: Probability topics (I)
Week 3: Probability topics (II)
Week 4: Discrete Random Variables: Binomial Distribution
Week 5: Discrete Random Variables: Poisson Distribution
Week 6: Continuous Probability Distribution: Uniform Distribution
Week 7: Normal Probability Distribution
Week 8: Sampling Methods and CLT
Week 9: Parameter estimation
Week 10: Hypothesis testing for a single parameter
Week 11: Hypothesis testing for two parameters

Tables: z and F tables will be provided if there are any relevant questions.
If Binomial or Poisson tables are not provided, then the students should know how to solve using the formulas.

For Section A, you just need to provide the correct answer.
E.g. 1 A, 2 B, 3 C, etc.

For Section B, you must provide all steps of calculations in detailed form. Round your answers to at least two decimals (e.g. 2.13 or 2.125).

No need to copy the questions to your answer sheet! Just indicating the number is sufficient!

Formula Sheet

$$\sigma^2 = \frac{\sum(x_i-\mu)^2}{N} \qquad s^2 = \frac{\sum(x_i-\bar{x})^2}{n-1}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad P(A \cap B) = P(A) * P(B \mid A)$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)*P(B|A)}{P(B)} \qquad P(A|B) = \frac{P(A)*P(B|A)}{[P(A)*P(B|A)] + [P(A^C)*P(B|A^C)]}$$

	With repetition	Without repetition
Combinations	${}_{n+r-1}C_r = \frac{(n+r-1)!}{r!*(n-1)!}$	${}_nC_r = \frac{n!}{r!*(n-r)!}$
Permutations	n^r	${}_nP_r = \frac{n!}{(n-r)!}$

$$E(x) = \sum x * p(x) \qquad E(x) = \int x * p(x)dx$$

$$E(x^2) = \sum x^2 * p(x) \qquad E(x^2) = \int x^2 * p(x)dx$$

$$Var(x) = E(x^2) - [E(x)]^2 \qquad Var(x) = \sum(x-\mu)^2 * p(x) \qquad Var(x) = \int(x-\mu)^2 * p(x)dx$$

$$P(x) = {}_nC_{x^*}p^{x}{}_*(1-p)^{n-x} \qquad \mu = np \qquad \sigma^2 = npq$$

$$P(x = k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad P(x = k) = \frac{(\lambda t)^k e^{-(\lambda t)}}{k!} \qquad e \approx 2.7182$$

$$P(x) = \frac{1}{b-a} \qquad \mu = \frac{a+b}{2} \qquad \sigma^2 = \frac{(b-a)^2}{12}$$

$$z = \frac{x-\mu}{\sigma} \qquad z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \qquad z \cong \frac{p-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$z = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \qquad z = \frac{p_1-p_2}{\sqrt{P(1-P)(\frac{1}{n_1}+\frac{1}{n_2})}} \qquad P = \frac{x_1+x_2}{n_1+n_2}$$

$$\bar{x} \pm z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}} \qquad p \pm z_{\frac{\alpha}{2}} * \sqrt{\frac{p(1-p)}{n}}$$

$$\bar{x}_1 - \bar{x}_2 \pm z_{\frac{\alpha}{2}} * \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad p_1 - p_2 \pm z_{\frac{\alpha}{2}} * \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

$$F = \frac{s_1^2}{s_2^2} \qquad df_1 = n_1\text{-}1 \qquad df_2 = n_2\text{-}1$$

Sample revision questions

1. The following tables show the scores of two groups of students in a test question.

i.

Mark	0	1	2	3	4
Frequency	10	5	8	x	3

Determine the value of x if the median of these marks is 1.5.

ii.

Mark	0	1	2	3	4
Frequency	10	5	8	y	3

Determine the value of y if the mean of these marks is 1.5.

2. A sales promotion offers a special price for a scanner with a printer. The probabilities that the scanner and the printer function satisfactorily for 3 years are 0.68 and 0.72, respectively. Assume that the scanner and the printer function independently.
What is the probability that at the end of three years:

i. both will function satisfactorily
ii. only the printer will function satisfactorily
iii. at least one will function satisfactorily?

3. A survey is conducted to compare customer satisfaction in two branches of a bank. Various customers visiting the two branches were selected randomly, and asked if they were satisfied with the services provided by the branch. The results of this survey are shown in the following table.

	Number satisfied	Number dissatisfied
Branch A	124	82
Branch B	118	65

a. You are asked to consider an appropriate hypothesis test to determine whether the proportion of satisfied customers is lower in Branch A compared to Branch B.
b. Compute the p-value of the test.
c. Compute a 97% confidence interval for the difference in proportion of satisfied customers visiting branches A and B.

4. A manufacturer of detergent claims that the contents of boxes sold weigh on average at least 16 ounces. The distribution of weight is known to be normal. A random sample of 16 boxes yielded a sample mean weight of 15.84 ounces and a sample standard deviation of 0.4 ounce.
a. Test at 5% significance level the null hypothesis that the population mean weight is at least 16 ounces.
b. What is the p-value of test?
c. What type of error is possible with your decision made in part a. Explain.

5. On average, there are 1.2 defects in a sheet of rolled steel. Assume that the number of defects

follows a Poisson distribution. What is the probability of having exactly two defects in a roll?

6. The probability that a person catches a cold during the cold and flu season is 0.4. Assume that 10 people are chosen at random. Find the probability that at most 1 of them will catch a cold.

7. The amount of time you have to wait at a particular stoplight is uniformly distributed between zero and three minutes. Eighty percent of the time, the stoplight will change before you have to wait X seconds, what is the value of X?

8. A recent survey of Fortune 500 firms found that on average, they contribute $330 per month for each salaried employee's health insurance. If you are told that almost 93% of salaried employees at Fortune 500 firms receive a health insurance contribution between $290 and $370, and assuming a bell-shaped distribution, what must the approximate standard deviation for this data be?

(5 marks) Two fair dice are thrown.

Given that a total score of at least 10 is obtained, what is the probability of at least one six occurring?