The University of Melbourne

Semester 1 Mid-Semester Exam — Sample Paper

School of Mathematics and Statistics

MAST90105 Methods of Mathematical Statistics

Exam Duration: 3 Hours

Reading Time: 15 Minutes

This paper has 6 pages

Authorised materials:
The calculator authorised at the University of Melbourne is the CASIO FX82 and this is permitted.
Two A4 double-sided handwritten sheets of notes are permitted.
A formula sheet is attached at the end of this paper.

Instructions to Invigilators:
Two 16-page script books will be supplied to each student.

Instructions to Students:
This paper has **11** questions.
Attempt as many questions, or parts of questions, as you can.
Questions carry marks as shown in the brackets after the question statement.
The total number of marks available for this examination is **100**.
Working and/or reasoning must be given to obtain full credit.

1. You ask your neighbour to water 3 sickly plants in your front yard while you are on vacation. Each of the plants will be alive with probability 0.85 if watered; it will be alive with probability 0.2 if not watered. You are 90 percent certain that your neighbour will remember to water the plants. [7]

   (a) What is the probability that you will find at least two of the three plants alive when you return home?

   (b) What is the expected number and variance of the number of plants that will be alive when you return home?

2. A certain cancer is found in one person in 5000. If a person does have the disease, in 92% of the cases the diagnostic procedure will show that he or she actually has it. If a person does not have the disease, the diagnostic procedure in one out of 500 cases gives a false positive result. [6]

   (a) Determine the probability that a person with a positive test result has the cancer.

   (b) Four people are found to have positive test results. Let X be the number of people in these four that actually have the cancer. Determine the probability $P(X = 1)$.

   (c) What is $E(X)$ and $Var(X)$

3. Let X be a discrete random variable with pmf (probability mass function)

$$f(x) = \frac{k}{2^x}, \quad x = 0, 1, 2$$

[7]

   (a) Find the value of k

   (b) Find the mgf (moment generating function) of X and hence find $E(X)$ and $E(X^2)$

   (c) Find $Var(X)$

4. In an annual charity drive, 30% of a population of 1500 make contributions. Suppose, after the drive, 15 people are selected at random and without replacement from the population to be given a statistical survey. Let X be the number of people in the survey who have made contributions. [6]

   (a) What is the name and the associated parameter values of the exact distribution of X? Also determine the probability $P(X \geq 2)$.

   (b) Since the survey sample size is much smaller than the population size, X can be approximated by a binomial distribution. Use this approximation to give an approximate value of $P(X \geq 2)$.

5. Let X be a continuous random variable with pdf (probability density function)

$$f(x) = \frac{1}{30}(x + 6), \quad -4 < x < 2$$

[12]

(a) Find $E(X)$ and $Var(X)$

(b) Find the cdf (cumulative distribution function) of X.

(c) Find the 25-th percentile of X.

(d) Consider the transformation $Y = 3X + 2$ of X.

i. Derive the cdf of Y.

ii. Find the pdf of Y.

6. A certain rare blood type can be found in only 0.05% of people. Let X be the number of people, in a population of 3000, that have this blood type. Note that X would follow a binomial distribution. [6]

(a) Find $P(X \leq 1)$.

(b) A binomial distribution $b(n, p)$ can be approximated by a Poisson($\lambda = np$) distribution if p is small and n large. Use this result to approximate the probability in part (a) by a Poisson probability.

(c) The probability in part (a) may also be approximated by a normal probability based on the central limit theorem. Give a normal approximation (using continuity correction) to $P(X \leq 1)$.

7. Police are to conduct random breath testing on drivers on a busy road one Friday evening. Suppose 2% of the drivers drink and drive at the time. Let X be the number of drivers that police need to test to capture the first case of drinking and driving. Let Y be the number of drivers tested to find 3 such cases. [11]

(a) Name the probability distribution and specify the value of any parameter(s) for each of the two random variables X and Y.

(b) What is the probability that at least 10 drivers are to be tested to capture the first drinking and driving case?

(c) What is the probability that exactly 30 drivers are to be tested to capture 3 drinking and driving cases?

(d) On average, how many drivers do police need to test to find 3 cases of drinking and driving?

(e) Find $P(Y > 50)$.

8. An insurance company employs four telephone operators who receive customers' calls from different regions, so the calls occur independently of each other. The number of calls received by each operator during follows a Poisson process with rate of occurrence parameter $\lambda = 2$ per minute. Let X_1, X_2, X_3 and X_4 be respectively the number of calls received by each of the four operators during a 1-minute period. [21]

(a) Calculate the probability $P(X_1 = 0)$.

(b) Let $Y = X_1 + X_2 + X_3 + X_4$ be the total number of calls received by the four operators during the 1-minute period. Name the probability distribution of Y and specify its parameter(s) value(s).

(c) Continuing (b), find the conditional probability $P(X_1 = 1|Y = 4)$.

(d) Let Z be the number of operators who do not receive any calls during the given 1-minute period. What probability distribution does Z have? Calculate the probability $P(Z = 2)$.

(e) If T_1, T_2, T_3, T_4 are the random variables giving the time in minutes till the first call of each of the operators arrives, derive the cdf (cumulative distribution function) of T_1.

(f) Derive the pdf (probability density function) of T_1

(g) Derive the mgf (moment generating function) of T_1 and hence $E(T_1)$

(h) Derive the mgf of $T_1 + T_2 + T_3 + T_4$.

(i) What is the probability that $T_1 + T_2 + T_3 + T_4 > 2$?

9. The error when using a surveying instrument to measure a distance is normally distributed with mean 0 and variance 1 mm. Measurements of the height, width and depth of the building are taken with this measuring device. [10]

(a) What is the probability that the error in (say) height exceeds an acceptable level of 2 mm?

(b) Derive the mgf (moment generating function) of the square on the error of one of these measurements? What is the name of this distribution?

(c) Derive the mgf of the sum of the squares of the three errors. What is the name of this distribution?

(d) The total error is defined as the square root of sum of the squares of the three errors. What is the probability that this exceeds 6 mm, that is three times the acceptable level in one measurement? Comment.

10. Let X_1, X_2, X_3 be independent Binomial$(n = 4, p = \frac{1}{2})$ random variables. Define $Y_1 = X_1 + X_3$ and $Y_2 = X_2 + X_3$. [10]

(a) Find the value of $\mathrm{Cov}(Y_1, Y_2)$.

(b) Define

$$Z_1 = \begin{cases} 1 & \text{if } Y_1 = 0, \\ 0 & \text{otherwise;} \end{cases} \quad \text{and} \quad Z_2 = \begin{cases} 1 & \text{if } Y_2 = 0, \\ 0 & \text{otherwise.} \end{cases}$$

i. Find the marginal pmf (probability mass function) of Z_1 and Z_2 respectively.

ii. Find the joint pmf (joint probability mass function) of (Z_1, Z_2).

iii. Find the correlation coefficient between Z_1 and Z_2.

11. Suppose X and Y are continuous random variables with the joint pdf (probability density function)

$$f(x, y) = \begin{cases} \frac{3}{2} & \text{if } 0 \le x \le 1 \text{ and } x^2 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

[4]

(a) Find the marginal pdf of Y.

(b) Find the conditional pdf of X given $Y = y$, $0 \leq y \leq 1$.

(c) Are X and Y independent? Explain.

`Total marks = 100`

End of the exam questions.
Formulas are on the next page.

Table XII: Discrete Distributions

Probability Distribution and Parameter Values	Probability Mass Function	Moment-Generating Function	Mean $E(X)$	Variance $\mathrm{Var}(X)$	Examples
Bernoulli $0 < p < 1$ $q = 1 - p$	$p^x q^{1-x}$, $x = 0, 1$	$q + pe^t$	p	pq	Experiment with two possible outcomes, say success and failure, $p = P(\text{success})$
Binomial $n = 1, 2, 3, \ldots$ $0 < p < 1$	$\binom{n}{x} p^x q^{n-x}$, $x = 0, 1, \ldots, n$	$(q + pe^t)^n$	np	npq	Number of successes in a sequence of n Bernoulli trials, $p = P(\text{success})$
Geometric $0 < p < 1$ $q = 1 - p$	$q^{x-1}p$, $x = 1, 2, \ldots$	$\dfrac{pe^t}{1 - qe^t}$	$\dfrac{1}{p}$	$\dfrac{q}{p^2}$	The number of trials to obtain the first success in a sequence of Bernoulli trials
Hypergeometric $x \le n, x \le N_1$ $n - x \le N_2$ $N = N_1 + N_2$ $N_1 > 0,\ N_2 > 0$	$\dfrac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}$		$n\left(\dfrac{N_1}{N}\right)$	$n\left(\dfrac{N_1}{N}\right)\left(\dfrac{N_2}{N}\right)\left(\dfrac{N-n}{N-1}\right)$	Selecting r objects at random without replacement from a set composed of two types of objects
Negative Binomial $r = 1, 2, 3, \ldots$ $0 < p < 1$	$\binom{x-1}{r-1} p^r q^{x-r}$, $x = r, r+1, \ldots$	$\dfrac{(pe^t)^r}{(1 - qe^t)^r}$	$\dfrac{r}{p}$	$\dfrac{rq}{p^2}$	The number of trials to obtain the rth success in a sequence of Bernoulli trials
Poisson $0 < \lambda$	$\dfrac{\lambda^x e^{-\lambda}}{x!}$, $x = 0, 1, \ldots$	$e^{\lambda(e^t - 1)}$	λ	λ	Number of events occurring in a unit interval, events are occurring randomly at a mean rate of λ per unit interval
Uniform $m > 0$	$\dfrac{1}{m}$, $x = 1, 2, \ldots, m$		$\dfrac{m+1}{2}$	$\dfrac{m^2 - 1}{12}$	Select an integer randomly from $1, 2, \ldots, m$

Table XIII: Continuous Distributions

Probability Distribution and Parameter Values	Probability Density Function	Moment-Generating Function	Mean $E(X)$	Variance $\mathrm{Var}(X)$	Examples
Beta $0 < \alpha$ $0 < \beta$	$\dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, $0 < x < 1$		$\dfrac{\alpha}{\alpha+\beta}$	$\dfrac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$	$X = X_1/(X_1 + X_2)$, where X_1 and X_2 have independent gamma distributions with same θ
Chi-square $r = 1, 2, \ldots$	$\dfrac{x^{r/2-1}e^{-x/2}}{\Gamma(r/2)2^{r/2}}$, $0 < x < \infty$	$\dfrac{1}{(1-2t)^{r/2}}$, $t < \dfrac{1}{2}$	r	$2r$	Gamma distribution, $\theta = 2$, $\alpha = r/2$; sum of squares of r independent $N(0, 1)$ random variables
Exponential $0 < \theta$	$\dfrac{1}{\theta} e^{-x/\theta}$, $0 \le x < \infty$	$\dfrac{1}{1 - \theta t}$, $t < \dfrac{1}{\theta}$	θ	θ^2	Waiting time to first arrival when observing a Poisson process with a mean rate of arrivals equal to $\lambda = 1/\theta$
Gamma $0 < \alpha$ $0 < \theta$	$\dfrac{x^{\alpha-1}e^{-x/\theta}}{\Gamma(\alpha)\theta^\alpha}$, $0 < x < \infty$	$\dfrac{1}{(1 - \theta t)^\alpha}$, $t < \dfrac{1}{\theta}$	$\alpha\theta$	$\alpha\theta^2$	Waiting time to αth arrival when observing a Poisson process with a mean rate of arrivals equal to $\lambda = 1/\theta$
Normal $-\infty < \mu < \infty$ $0 < \sigma$	$\dfrac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma\sqrt{2\pi}}$, $-\infty < x < \infty$	$e^{\mu t + \sigma^2 t^2/2}$	μ	σ^2	Errors in measurements; heights of children; breaking strengths
Uniform $-\infty < a < b < \infty$	$\dfrac{1}{b-a}$, $a \le x \le b$	$\dfrac{e^{tb} - e^{ta}}{t(b-a)}$, $t \ne 0$; 1, $t = 0$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	Select a point at random from the interval $[a, b]$

Appendix C Tables