THE LNM INSTITUTE OF INFORMATION TECHNOLOGY
DEPARTMENT OF MATHEMATICS
MATH-221: PROBABILITY AND STATISTICS
END TERM

Maximum Time: 3 Hours Date: 01/05/2018 Maximum Marks: 50

Instruction: You should attempt all questions. Your writing should be legible and neat. Marks awarded are shown next to the question. **NO USE OF CALCULATORS.** Please make an index showing the question number and page number on the front page of your answer sheet in the following format, otherwise you may be penalized by the deduction of **4 marks**.

Question No.				
Page No.				

---

1. (a) Compute the probability that if 10 married couples are seated at random at a round table, then no wife sits next to her husband. **[03 Marks]**

Solution: Let $E_i,\ i = 1, 2, \cdots, 10$ denote the event that the i^{th} couple sit next to each other. Then, the probability that at least one couple sit next to her husband is $p(\bigcup_1^{10} E_i)$. It follows that the desired probability that no wife sits next to her husband is $1 - p(\bigcup_1^{10} E_i)$. **[0.5 marks]**

Now,

$$p(\bigcup_1^{10}) = \sum_{i=1}^{10} P(E_i) - \cdots + (-1)^{n+1} \sum_{i_1<i_2<\cdots<i_n} P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_n}) + \cdots - P(E_1 \cap E_2 \cdots \cap E_{10}) \quad (1)$$

[0.5 marks]

We can arrange 19! ways of 20 people around a round table.
To obtain the number of arrangements that "10 men sitting next to their wives" can be obtained by taking each of the n married couples as being single entities. Then, we arrange $20 - n$ entities around a round table. Hence, total number of such arrangements are $(20 - n - 1)!$
As each of the n married couples can be arranged next to each other in one of two possible ways, it follows that there are $2^n(20-n-1)! = 2^n(19-n)!$ arrangements that result in a specified set of n men each sitting next to their wives.
So,

$$P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_n}) = \frac{2^n(19-n)!}{19!} \quad (2)$$

[0.5 marks]

So by (A), the probability that at least one married couple sits together

$$\begin{aligned} p(\bigcup_1^{10}) &= {}^{10}C_1 2^1 \frac{(18!)}{(19!)} - {}^{10}C_2 2^2 \frac{(17!)}{(19!)} + {}^{10}C_3 2^3 \frac{(16!)}{(19!)} - \cdots - {}^{10}C_{10} 2^{10} \frac{(9!)}{(19!)} &(3)\\ &\approx 0.6605 &(4) \end{aligned}$$

[0.5 marks]

Finally, the probability that no married couple sits together $= 1 - 0.6605 = 0.3395$ **[0.5 marks]**

(b) Write down the probability density function f of a general normal random variable X. Prove that f is indeed a probability density function. Additionally, find expectation and variance of the random variable X. **[05 Marks]**

Solution: The density function of a normal random variable X with parameters μ and σ^2 is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad -\infty < x < \infty$$

[01 Marks]

To prove that $f(x)$ is indeed a probability density function, we need to show that $\int_{-\infty}^{\infty} f(x)dx = 1$. In other words

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\, dx = 1$$

[0.5 Marks]

Making the substitution $y = (x-\mu)/\sigma$, we see that

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}\, dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-y^2/2}\, dy$$

Let $I = \int_{-\infty}^{\infty} e^{-y^2/2}\, dy$, then,
$I^2 = \int_{-\infty}^{\infty} e^{-y^2/2}\, dy \int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(y^2+x^2)/2} dydx$
Evaluate the double integral by means of a change of variables to polar coordinates. (That is, let $x = r\cos\theta, y = r\sin\theta$, and $dydx = rd\theta dr$.) Hence,

$$\begin{aligned} I^2 &= \int_0^{\infty}\int_0^{2\pi} e^{-r^2/2} r d\theta dr \\ &= 2\pi \int_0^{\infty} re^{-r^2/2} dr \\ &= 2\pi \end{aligned}$$

So, we get $I = \int_{-\infty}^{\infty} e^{-y^2/2}\, dy = \sqrt{2\pi}$ **[01 Marks]**

Which provides

$$\int_{-\infty}^{\infty} f(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2}\, dy = \frac{1}{\sqrt{2\pi}}\sqrt{2\pi} = 1$$

[0.5 Marks]

Expatiation and variance

Let first calculate for standard normal random variable $Z = (X-\mu)/\sigma$

$$\begin{aligned} E[Z] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-x^2/2}\, dx \\ &= \frac{1}{\sqrt{2\pi}} [e^{-x^2/2}]_{-\infty}^{\infty} \\ &= 0 \end{aligned}$$

[0.5 Marks]

$Var(Z) = E[z^2] - E[Z]^2 = E[z^2]$
Now, $E[Z^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2}\, dx$

Integrating by parts gives:
$E[Z^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2}\, dx = 1$
Hence $Var(Z) = E[z^2] = 1$ **[0.5 Marks]**
As, $Z = (X-\mu)/\sigma \Rightarrow X = \mu + \sigma$
This gives $E[X] = \mu + \sigma E[Z] = \mu$ **[0.5 Marks]**
Similarly, $Var(X) = \sigma^2 Var(Z) = \sigma^2$ **[0.5 Marks]**

2. (a) If random variables X and Y are independent and uniformly distributed on $(0,1)$, then calculate the probability density function of $X+Y$. **[04 Marks]**

Solution: We know that

$$f_X(a) = f_Y(a) = \begin{cases} 1, & \text{if } 0 < a < 1, \\ 0, & \text{if otherwise.} \end{cases}$$

Since X and Y are independent. So

$$f_{X+Y}(x,y) = f_X(x) f_Y(y) = \begin{cases} 1, & \text{if } 0 < x, y < 1, \\ 0, & \text{if otherwise.} \end{cases}$$

[0.5 Marks]

By convolution of the distribution

$$f_{X+Y}(a) = \int_{-\infty}^{+\infty} f_X(a-y) f_Y(y) dy$$

[0.5 Marks]

$$f_{X+Y}(a) = 0 + \int_0^1 f_X(a-y) f_Y(y) dy + 0$$

$$f_{X+Y}(a) = \int_0^1 f_X(a-y) f_Y(y) dy = \int_0^1 f_X(a-y) dy.$$

since $f_Y(a) = 1$ for $0 < a < 1$. **[0.5 Marks]**

For $0 \le a \le 1$, we have

$$f_{X+Y}(a) = \int_0^a f_X(a-y) dy + \int_a^1 f_X(a-y) dy = \int_0^a dy + 0 = a.$$

Since $(a-y) \le 0$. **[1 Marks]**
For $1 < a < 2$, we have

$$f_{X+Y}(a) = \int_0^{a-1} f_X(a-y) dy + \int_{a-1}^1 f_X(a-y) dy = 0 + \int_{a-1}^1 dy = 2 - a.$$

Since $(a-y) \ge 1$. **[1 Marks]**
So

$$f_{X+Y}(a) = \begin{cases} a, & \text{if } 0 \le a \le 1, \\ 2-a, & \text{if } 1 < a < 2, \\ 0, & \text{if otherwise.} \end{cases}$$

[0.5 Marks]

Alternate Solution :

We know that

$$f_X(a) = f_Y(a) = \begin{cases} 1, & \text{if } 0 < a < 1, \\ 0, & \text{if otherwise.} \end{cases}$$

Since X and Y are independent. So the joint pdf of X and Y is

$$f(x,y) = f_X(x)f_Y(y) = \begin{cases} 1, & \text{if } 0 < x, y < 1, \\ 0, & \text{if otherwise.} \end{cases}$$

[**0.5 Marks**]

Let $Z = X + Y$

$$F_z(a) = P\{Z \leq a\} = \int\int_A f(x,y)dxdy.$$

where $A = \{(x,y) : x + y \leq a\}$.

If $a < 0$, $\quad F_z(a) = 0.$ [**0.5 Marks**]

If $0 \leq a \leq 1$, $\quad F_z(a) = \frac{1}{2}a^2.$ [**0.5 Marks**]

If $1 \leq a \leq 2$, $\quad F_z(a) = \frac{(4a-a^2-2)}{2}.$ [**1.5 Marks**]

If $a > 2$, $\quad F_z(a) = 1.$ [**0.5 Marks**]

$$F_Z(a) = \begin{cases} 0, & \text{if } a < 0, \\ \frac{1}{2}a^2, & \text{if } 0 \leq a \leq 1, \\ \frac{(4a-a^2-2)}{2}, & \text{if } 1 < a < 2, \\ 1, & \text{if } a > 2. \end{cases}$$

We Know that $f_Z(a) = \frac{d}{da}F_Z(a)$

$$f_Z(a) = \begin{cases} 0, & \text{if } a < 0, \\ a, & \text{if } 0 \leq a \leq 1, \\ 2-a, & \text{if } 1 < a < 2, \\ 0, & \text{if } a > 2. \end{cases}$$

[**0.5 Marks**]

(b) The joint probability density function of random variables X and Y is given as

$$f(x,y) = \begin{cases} 6(1-x), & \text{if } 0 < y < x, 0 < x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Determine the probability density function of random variables X and Y. Are X and Y independent? Give reasons to your answers. [**04 Marks**]

Solution: **Density of X:** We have

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)dy, \forall x \in R.$$

Since $f(x,y)$ is zero for $x \geq 1$ and $x \leq 0$, therefore $f_X(x) = 0$ if $x \geq 1$ or $x \leq 0$.

$$f_X(x) = \int_0^x f(x,y)dy$$

$$f_X(x) = \int_0^x 6(1-x)dx = 6x(1-x).$$

$$f_X(x) \begin{cases} 6x(1-x), & \text{if } 0 < x < 1, \\ 0, & \text{if otherwise.} \end{cases}$$

[**1.5 Marks**]

Density of Y: We have

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y)dx, \forall y \in R.$$

Since $f(x,y)$ is zero for $y \geq 1$ and $y \leq 0$, therefore $f_Y(y) = 0$ if $y \geq 1$ or $y \leq 0$.

$$f_Y(y) = \int_0^1 f(x,y)dx = \int_0^y f(x,y)dx + \int_y^1 f(x,y)dx$$

$$f_Y(y) = 0 + \int_y^1 f(x,y)dx = \int_y^1 6(1-x)dx = 3(y-1)^2$$

$$f_Y(y) \begin{cases} 3(y-1)^2, & \text{if } 0 < y < 1, \\ 0, & \text{if otherwise.} \end{cases}$$

[**0.5 Marks**]

Independent: $f(x,y) = 6(1-x) \neq 6x(1-x) \times 3(y-1)^2 = f_X(x)f_Y(y)$ at point $(\frac{1}{2}, \frac{1}{4})$ which is a point of continuity of f and$f_X(x)f_Y(y)$. So X and Y are dependent. [**1 Marks**]

Note: If proper reason is not given for independency zero marks is given.

3. (a) Let X and Y be random variables with the joint pmf $P\{X = i, Y = j\} = \dfrac{1}{N^2}, i, j = 1, 2, \cdots, N$. Find the mean of the random variable $\min\{X, Y\}$. [**03 Marks**]

First Solution: Set $Z := \min\{X, Y\}$. Note that range of random variable Z is $\{1, 2, \cdots, N\}$. Therefore $E[Z]$ is defined

 Please Turn Over $\cdots$

and

$$
\begin{aligned}
E[Z] &= \sum_{x=1}^{N}\sum_{y=1}^{N}\min\{x,y\}P(X=x,Y=y) \quad [\textbf{0.5 Marks}] \\
&= \frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1}^{N}\min\{x,y\} = \frac{1}{N^2}\left[\sum_{x=1}^{N}\sum_{y=x}^{N}x + \sum_{x=2}^{N}\sum_{y=1}^{x-1}y\right] \\
&= \frac{1}{N^2}\left[\sum_{x=1}^{N}x\sum_{y=x}^{N}1 + \sum_{x=2}^{N}\frac{(x-1)(x-1+1)}{2}\right] = \frac{1}{N^2}\left[\sum_{x=1}^{N}x(N-x+1) + \frac{1}{2}\sum_{x=2}^{N}(x-1)x\right] \\
&= \frac{1}{N^2}\left[(N+1)\sum_{x=1}^{N}x - \sum_{x=1}^{N}x^2 + \frac{1}{2}\left(\sum_{x=2}^{N}x^2 - \sum_{x=2}^{N}x\right)\right] \\
&= \frac{1}{N^2}\left[(N+1)\sum_{x=1}^{N}x - \sum_{x=1}^{N}x^2 + \frac{1}{2}\left(\sum_{x=1}^{N}x^2 - 1 - \sum_{x=1}^{N}x + 1\right)\right] \\
&= \frac{1}{N^2}\left[(N+1)\sum_{x=1}^{N}x - \sum_{x=1}^{N}x^2 + \frac{1}{2}\sum_{x=1}^{N}x^2 - \frac{1}{2}\sum_{x=1}^{N}x\right] \\
&= \frac{1}{N^2}\left[\frac{(2N+1)}{2}\sum_{x=1}^{N}x - \frac{1}{2}\sum_{x=1}^{N}x^2\right] \\
&= \frac{1}{N^2}\left[\frac{(2N+1)N(N+1)}{4} - \frac{1}{2}\frac{N(N+1)(2N+1)}{6}\right] = \frac{(N+1)(2N+1)}{6N} \quad [\textbf{2.5 Marks}]
\end{aligned}
$$

econd Solution: One might compute the double sum in following way also.

$$\begin{aligned}
E[Z] &= \sum_{x=1}^{N}\sum_{y=1}^{N} \min\{x,y\} P(X=x, Y=y) \quad [\mathbf{0.5\ Marks}] \\
&= \frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1}^{N} \min\{x,y\} = \frac{1}{N^2}\left[\sum_{x=1}^{N}\sum_{y=x}^{N} x + \sum_{y=1}^{N-1}\sum_{x=y+1}^{N} y\right] \\
&= \frac{1}{N^2}\left[\sum_{x=1}^{N} x \sum_{y=x}^{N} 1 + \sum_{y=1}^{N-1} y \sum_{x=y+1}^{N} 1\right] \\
&= \frac{1}{N^2}\left[\sum_{x=1}^{N} x(N-x+1) + \sum_{y=1}^{N-1} y(N-(y+1)+1)\right] \\
&= \frac{1}{N^2}\left[\sum_{x=1}^{N} xN - \sum_{x=1}^{N} x^2 + \sum_{x=1}^{N} x + \sum_{y=1}^{N-1} Ny - \sum_{y=2}^{N-1} y^2\right] \\
&= \frac{1}{N^2}\left[N\sum_{x=1}^{N} x - \sum_{x=1}^{N} x^2 + \sum_{x=1}^{N} x + N\sum_{y=1}^{N-1} y - \sum_{y=2}^{N-1} y^2\right] \\
&= \frac{1}{N^2}\left[N\sum_{x=1}^{N} x - \sum_{x=1}^{N} x^2 + \sum_{x=1}^{N} x + N\sum_{y=1}^{N} y - N^2 - \sum_{y=1}^{N} y^2 + N^2\right] \\
&= \frac{1}{N^2}\left[2N\sum_{x=1}^{N} x - 2\sum_{x=1}^{N} x^2 + \sum_{x=1}^{N} x\right] \\
&= \frac{1}{N^2}\left[(2N+1)\sum_{x=1}^{N} x - 2\sum_{x=1}^{N} x^2\right] = \frac{1}{N^2}\left[(2N+1)\frac{N(N+1)}{2} - \frac{N(N+1)(2N+1)}{3}\right] \\
&= \frac{(2N+1)(N+1)}{6N} \quad [\mathbf{2.5\ Marks}]
\end{aligned}$$

Third Solution: Set $Z := \min\{X, Y\}$. Since X and Y are discrete random variable, hence Z will also be a discrete random variable. Also it is clear that range of random variable Z would be $\{1, 2, \cdots, N\}$. Now we find its pmf. Hence for given $i \in \{1, 2, \cdots, N\}$,

$$\begin{aligned}
P\{Z=i\} &= \sum_{(x,y):\min\{x,y\}=i} f(x,y) \\
&= \sum_{y=i}^{N} f(i,y) + \sum_{x=i+1}^{N} f(x,i) \\
&= \frac{(N-(i-1))}{N^2} + \frac{N-i}{N^2} \\
&= \frac{2N-2i+1}{N^2}, \text{ for } i = 1, 2, \cdots, N
\end{aligned}$$

[**01 Marks**]

Therefore

$$\begin{aligned}
E[Z] &= \sum_{i=1}^{N} iP(Z=i) = \frac{1}{N^2}\sum_{i=1}^{N} 2Ni - 2i^2 + i = \frac{1}{N^2}\left[2N\sum_{i=1}^{N} i - 2\sum_{i=1}^{N} i^2 + \sum_{i=1}^{N} i\right] \\
&= \frac{1}{N^2}\left[2NN\frac{N+1}{2} - 2\frac{N(N+1)(2N+1)}{6} + N\frac{N+1}{2}\right] \\
&= \frac{1}{N}\left[N(N+1) - \frac{(N+1)(2N+1)}{3} + \frac{N+1}{2}\right] = \frac{N+1}{N}\left[N - \frac{2N+1}{3} + \frac{1}{2}\right] \\
&= \frac{N+1}{6N}\left[6N - 4N - 2 + 3\right] \\
&= \frac{(N+1)(2N+1)}{6N}
\end{aligned}$$

[**02 Marks**]

(b) State the Jensen's Inequality. Hence, show that if the moment of order $q > 0$ exists for a random variable X, then moments of order p, where $0 < p < q$ exist. [**03 Marks**]

Solution: **Jensen's Inequality:** Let $f : I \to \mathbb{R}$ be a convex function where $I \subset \mathbb{R}$ is an interval such that it contains the range of random variable X. Suppose that X and $f(X)$ has finite mean. Then

$$f(EX) \leq E[f(X)].$$

[**01 Marks**]

Let $f : [0, \infty) \to \mathbb{R}$ be defined as $f(x) = x^r$, where $r > 1$ is a real number. Then $f'(x) = rx^{r-1}, f''(x) = r(r-1)x^{r-2}$. Since $r > 1$, $f''(x) \geq 0$ on $[0, \infty)$, i.e., f is a convex function on $[0, \infty)$. Hence by Jensen's inequality,

$$[E|X|]^r \leq E\left[|X|^r\right] \implies [E|X|] \leq \left(E\left[|X|^r\right]\right)^{\frac{1}{r}}. \tag{5}$$

Let $0 < p < q$. Then we take $r = \frac{q}{p} > 1$ in (5) and we get

$$[E|X|] \leq \left(E\left[|X|^{\frac{q}{p}}\right]\right)^{\frac{p}{q}}. \tag{6}$$

Now replacing $|X|$ by $|X|^p$ in (6), we get

$$[E|X|^p] \leq \left(E\left[|X|^q\right]\right)^{\frac{p}{q}}$$

If $E|X|^q < \infty$ then $(E|X|^q)^{\frac{p}{q}} < \infty$ and therefore $[E|X|^p] < \infty$. [**02 Marks**]

(c) Let $X \sim B(n, p)$. Estimate $P(X \geq \alpha n)$, where $p < \alpha < 1$, using the Chebyshev's inequality. [**03 Marks**]

Solution: **Chebyshev's inequality:** Let X be a random variable with finite mean μ and finite variance σ^2. Then for every $\epsilon > 0$,

$$P\{|X - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{\epsilon^2}$$

[**0.5 Marks**]

Note that $EX = np, \operatorname{var}(X) = np(1-p)$. [**0.5 Marks**]

Chebyshev's inequality gives estimate for $P(|X - EX| \geq \alpha n)$ so we have rewrite the event $\{X \geq \alpha n\}$ so that we can use the Chebushev's inequality.

$$\begin{aligned}
P\{X \geq \alpha n\} &= P\{X - np \geq \alpha n - np\} \\
&\leq P\left(|X - np| \geq \alpha n - np\right) \quad (\because \{|Y| \geq a\} = \{Y \leq -a\} \cup \{Y \geq a\}) \\
&\leq \frac{\text{var(X)}}{(\alpha n - np)^2} = \frac{np(1-p)}{n^2(\alpha - p)^2} = \frac{p(1-p)}{n(\alpha - p)^2}
\end{aligned}$$

[02 Marks]

4. (a) Find the characteristic function of a Poisson random variable with the parameter $\lambda > 0$. State the uniqueness theorem. Hence, prove that sum of two independent Poisson random variable is a Poisson random variable. **[04 Marks]**

Solution:

$$\begin{aligned}
\phi_X(t) = E\left[e^{itX}\right] &= \sum_{k=0}^{\infty} e^{itk} P(X=k) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(e^{it}\lambda\right)^k}{k!} = e^{-\lambda} \exp(e^{it}\lambda) \\
&= \exp\left(\lambda(e^{it} - 1)\right)
\end{aligned}$$

[01 Marks]

Uniqueness Theorem: Let X_1 and X_2 be two random variables such that $\phi_{X_1} = \phi_{X_2}$. Then X_1 and X_2 have same distribution. **[01 Marks]**

Let $X \sim$ Poisson(λ) and $Y \sim$ Poisson(μ) be two independent Poisson random variables. Then characteristic function of $X + Y$ is

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) = \exp\left(\lambda(e^{it} - 1)\right) \exp\left(\mu(e^{it} - 1)\right) = \exp\left((\lambda + \mu)(e^{it} - 1)\right)$$

RHS is a characteristic function of a Poisson($\lambda + \mu$) random variable, therefore by uniqueness theorem $X + Y \sim$ Poisson($\lambda + \mu$). **[02 Marks]**

(b) Let $X_1, X_2, \cdots, X_{25}$ be i.i.d. random variable with the following probability mass function

$$f(t) = \begin{cases} 0.6 & \text{if} \quad t = 1, \\ 0.4 & \text{if} \quad t = -1, \\ 0 & \qquad \text{otherwise.} \end{cases}$$

Let $Y = X_1 + X_2 + \cdots + X_{25}$. Using the central limit theorem compute approximate value of $P(4 \leq Y \leq 6)$. Use the standard normal table given in the end. **[04 Marks]**

Solution:

$$\begin{aligned}
\mu = EX_1 &= 1 \times P(X_1 = 1) - 1 \times P(X_1 = -1) = 0.2 \\
EX_1^2 &= 1^2 \times P(X_1 = 1) + (-1)^2 \times P(X_1 = -1) = 0.6 + 0.4 = 1 \\
\sigma^2 = Var(X_1) &= 1 - (0.2)^2 = 1 - 0.04 = 0.96
\end{aligned}$$

[01 Marks]

Here $n = 25$, hence $n\mu = 5, \sqrt{n}\sigma = 5 \times 0.98 = 4.9$. Now CLT tell us to treat Y as a normal random variable with mean 5 and standard deviation 4.9.

$$\begin{aligned}
P(4 \leq Y \leq 6) &= P(Y \leq 6) - P(Y < 4) = P(Y \leq 6) - P(Y \leq 4) \\
&= N\left(\frac{6-5}{4.9}\right) - N\left(\frac{4-5}{4.9}\right) = N\left(\frac{1}{4.9}\right) - N\left(\frac{-1}{4.9}\right) = N\left(\frac{1}{4.9}\right) - 1 + N\left(\frac{1}{4.9}\right) \\
&= 2N(0.204) - 1 \qquad \textbf{[02Marks]} \\
&= 2 \times 0.5793 - 1 = 0.1586 \qquad \textbf{[01Marks]}
\end{aligned}$$

5. (a) Let X_1, $X_2, \ldots, X_n$ be a random sample from the population with the following probability density function:

$$f(x) = \begin{cases} \theta\left(x - \frac{1}{2}\right) + 1, & 0 \leq x \leq 1, \\ 0, & \text{otherwise}, \end{cases}$$

where, $\theta \in [-2, 2]$ is an unknown parameter. We define the estimator $\hat{\Theta}_n$ as $\hat{\Theta}_n = 12\overline{X} - 6$ to estimate θ, where $\overline{X}$ is the sample mean.

(i) Is $\hat{\Theta}_n$ an unbiased estimator of θ?

(ii) Find the mean squared error (MSE) of $\hat{\Theta}_n$.

(ii) Is $\hat{\Theta}_n$ a consistent estimator of θ?

[6 marks]

Solution: **(i)** We know that bias of $\hat{\Theta}_n$ is defined as $B(\hat{\Theta}_n) = E(\hat{\Theta}_n) - \theta$. Thus, to find the bias of $\hat{\Theta}_n$, we need to calucualate

$$E(\hat{\Theta}_n) = E(12\overline{X} - 6) = 12E(\overline{X}) - 6, \text{ (By linearity of Expectation).} \tag{7}$$

Now

$$E(\overline{X}) = E\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n}E(X_1 + X_2 + \cdots + X_n) = \frac{1}{n}\left[E(X_1) + E(X_2) + \cdots + E(X_n)\right]$$
$$= \frac{1}{n}nE(X) = E(X), \text{ (As } X_1, X_2, \ldots, X_n \text{ are i.i.d. random variables).}$$

[01 Marks]

Now

$$E(X) = \int_0^1 x\left(\theta\left(x - \frac{1}{2}\right) + 1\right)dx = \int_0^1 \left(\theta x^2 - \frac{\theta x}{2} + x\right)dx = \frac{\theta}{3} - \frac{\theta}{4} + \frac{1}{2} = \frac{\theta + 6}{12}.$$

Thus, from Eq. (1), we have $E(\hat{\Theta}_n) = 12\left(\frac{\theta + 6}{12}\right) - 6 = \theta.$ **[01 Marks]**

$\Longrightarrow B(\hat{\Theta}_n) = E(\hat{\Theta}_n) - \theta = 0.$

$\Longrightarrow \hat{\Theta}_n$ is an unbiased estimator.

(ii) $MSE(\hat{\Theta}_n) = Var(\hat{\Theta}_n) + (B(\hat{\Theta}_n))^2 = Var(\hat{\Theta}_n), \text{ as } B(\hat{\Theta}_n) = 0.$
Now, $Var(\hat{\Theta}_n) = Var(12\overline{X} - 6) = 144Var(\overline{X})$.

Furthermore,

$$Var(\overline{X}) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n^2}Var(X_1 + X_2 + \cdots + X_n) \text{ [}\mathbf{0.5\ Marks}\text{]}$$
$$= \frac{1}{n^2}\left[Var(X_1) + Var(X_2) + \cdots + Var(X_n)\right] = \frac{1}{n^2}nVar(X) = \frac{Var(X)}{n}, \text{ (As } X_i's \text{ are i.i.d. random variables).}$$

[0.5 Marks]

Now $Var(X) = E(X^2) - E(X)^2$. Thus, to calculate $Var(X)$, we find $E(X^2)$ as

$$E(X^2) = \int_0^1 x^2\left(\theta\left(x - \frac{1}{2}\right) + 1\right)dx = \int_0^1 \left(\theta x^3 - \frac{\theta x^2}{2} + x^2\right)dx = \frac{\theta}{4} - \frac{\theta}{6} + \frac{1}{3} = \frac{2\theta + 8}{24}.$$

Thus,

$$Var(X) = \frac{2\theta+8}{24} - \left(\frac{\theta+6}{12}\right)^2 = \frac{12-\theta^2}{144}. \text{ [}\mathbf{01\ Marks}\text{]}$$

$$\Longrightarrow Var(\overline{X}) = \frac{Var(X)}{n}\frac{12-\theta^2}{144n}$$

$$\Longrightarrow Var(\hat{\Theta}_n) = 144Var(\overline{X}) = \frac{12-\theta^2}{n}$$

$$\Longrightarrow MSE(\hat{\Theta}_n) = \frac{12-\theta^2}{n}.$$ [**01 Marks**]

(iii) $MSE(\hat{\Theta}_n) \to 0$ as $n \to 0$, so $\hat{\Theta}_n$ is consistent estimator of θ. [**01 Marks**]

(b) Let $X_1,\ X_2,\ldots,X_n$ be a random sample taken from a population with the following density function

$$f(x) = \begin{cases} e^{\lambda-x}, & x>\lambda, \\ 0, & \text{otherwise.} \end{cases}$$

Define $X_{(1)} = \min\{X_1,\ X_2,\ \ldots, X_n\}$. Find $P\left(X_{(1)} > 3\lambda\right)$. [**3 marks**]

Solution: We know that if $X_{(1)},\ X_{(2)},\ldots,X_{(n)}$ be the order statistics of $X_1, X_2, \ldots, X_n$, then the p.d.f. of $X_{(i)}$ are given by

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!} f(x)\,[F(x)]^{i-1}\,[1-F(x)]^{n-i}.$$

Therefore,

$$f_{X_{(1)}}(x) = \frac{n!}{(1-1)!(n-1)!} f(x)\,[F(x)]^{1-1}\,[1-F(x)]^{n-1} = nf(x)\,[1-F(x)]^{n-1}. \text{ [}\mathbf{01\ Marks}\text{]} \tag{8}$$

Thus, to find p.d.f. $f_{X_{(1)}}(x)$ of random variable $X_{(1)}$, we need to find c.d.f. $F(x)$ of random variable X. $F(x) = P(X \le x) = 0$, $x < \lambda$. In the case of $x > \lambda$, we have

$$F(x) = \int_\lambda^x e^{\lambda-t}dt = -e^{\lambda-x} + 1.$$

Thus, from Eq. (8), we have

$$f_{X_{(1)}}(x) = \begin{cases} ne^{\lambda-x}[1+e^{\lambda-x}-1]^{n-1} = ne^{n(\lambda-x)}, & x>\lambda, \\ 0, & \text{otherwise.} \end{cases} \text{ [}\mathbf{01\ Marks}\text{]}$$

Thus

$$P(X_{(1)} > 3\lambda) = \int_{3\lambda}^\infty ne^{n(\lambda-x)}dx = e^{-2n\lambda} \text{ [}\mathbf{01\ Marks}\text{]}$$

6. (a) Let $X_1,\ X_2,\ldots,X_n$ be a random sample from a geometric distribution with the parameter p. Find the Maximum Likelihood Estimator (MLE) of p. [**4 marks**]

Solution: Note that a geometric(p) random variable has the pmf

$$f(x) = p(1-p)^{x-1},\quad x=1,2,\cdots \text{ [}\mathbf{01\ Marks}\text{]}$$

The likelihood function of the given random sample is

$$L(p|x_1,x_2,\cdots,x_n) = [p(1-p)^{x_1-1}][p(1-p)^{x_2-1}]\cdots[p(1-p)^{x_n-1}] = p^n(1-p)^{\sum_{i=1}^n x_i-n},\ \text{for } p\in(0,1).$$

[01 Marks]

Note that we have parameter range $(0,1)$ rather than $[0,1]$ (if $p=0$ or $p=1$ then we pmf becomes identically zero, which is absurd). Set $y=\sum_{i=1}^{n} x_i$, then $y \geq n$ and taking logarithm of likelihood function we obtain

$$f(p) = \log L(p) = n \log p + (y-n)\log(1-p).$$

$$f'(p) = \frac{n}{p} - \frac{y-n}{1-p} = 0 \implies \frac{n}{p} = \frac{y-n}{1-p} \implies \frac{1-p}{p} = \frac{y-n}{n} \implies \frac{1-p+p}{p} = \frac{y-n+n}{n}$$
$$\implies \frac{1}{p} = \frac{y}{n} \implies p = \frac{n}{y} \quad \textbf{[01 Marks]}$$

Note that if $p < \frac{n}{y}$ then

$$\frac{n}{p} > y. \text{ And also } -p > -\frac{n}{y} \implies 1-p > 1-\frac{n}{y} \implies 1-p > \frac{y-n}{y} \implies y > \frac{y-n}{1-p}.$$

Therefore $f'(p) > 0$ on the interval $(0, \frac{n}{y})$ and by reversing the inequality we obtain $f'(p) < 0$ on the interval $(\frac{n}{y}, 1)$. Therefore f is strictly increasing on $(0, \frac{n}{y})$ and strictly decreasing on $(\frac{n}{y}, 1)$. Alternatively, it is easy to see that $\lim_{p\to 0^+} f(p) = -\infty = \lim_{p \to 1^-} f(p)$. So $p = \frac{n}{y}$ is the point of global maximum. Hence $\hat{p} = \frac{n}{\sum_{i=1}^{n} X_i}$ the MLE of p for the given sample. **[01 Marks]**

(b) Define the pivotal quantity. Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\theta, 9)$. Check the random variables $Q_1 = \overline{X} - \theta$ and $Q_2 = \frac{\sqrt{n}(\overline{X}-\theta)}{3}$ for pivotal quantity. **[4 marks]**

Solution: Let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from a distribution with a parameter θ that is to be estimated. The random variable Q is said to be a pivot or a pivotal quantity, if it has the following properties:

i. It is a function of the observed data $X_1, X_2, X_3, \ldots, X_n$ and the unknown parameter θ, but it does not depend on any other unknown parameters: $Q = Q(X_1, X_2, \ldots, X_n, \theta)$.

ii. The probability distribution of Q does not depend on θ or any other unknown parameters.

[01 Marks]

We note that Q_1 and Q_2 by definitions are functions of $\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$ and θ, and thus we conclude that Q_1 and Q_2 both are the functions of the observed data $X_1, X_2, \ldots, X_n$ and the unknown parameter θ, and they do not depend on any other unknown parameters. **[01 Marks]**

Since $X_i \sim N(\theta, 9)$ and $X_1, X_2, \ldots, X_n$ are i.i.d. random variables, therefore we conclude that

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \sim N\left(\theta, \frac{9}{n}\right),$$

which implies that $Q_1 = \overline{X} - \theta \sim N\left(0, \frac{9}{n}\right)$. Also by normalising $\overline{X}$, we conclude that the random variable

$$\frac{\overline{X}-\theta}{\frac{3}{\sqrt{n}}} = \frac{\sqrt{n}(\overline{X}-\theta)}{3} = Q_2 \sim N(0,1)$$

[01 Marks]

Thus, distributions of Q_1 and Q_2 do not depend on θ or any other unknown parameters. We conclude that Q_1 and Q_2 are both valid pivots. **[01 Marks]**

TABLE 1 Values of the standard normal distribution function

x	0	1	2	3	4	5	6	7	8	9
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5363	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9430	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9648	.9656	.9664	.9671	.9678	.9686	.9693	.9700	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9762	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9874	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.	.9987	.9990	.9993	.9995	.9997	.9998	.9998	.9999	.9999	1.0000