MSO201A 2021-22-II END SEM EXAMINATION SOLUTIONS

Question 1. (1.5 + 1.5 marks) Let $X \sim t_2$. Then $\boxed{\mathbb{E}X =\quad}$ and variance of X $\boxed{\quad}$

Answer: $\mathbb{E}X = 0$ and variance of X 'does not exist'.

Question 2. (3 marks) Which of the following RVs Y have zero mean? Put a tick (✓) beside all the correct option(s) to get credit.

(a) $Y \sim Cauchy(0, 1)$ (b) $Y \sim Exponential(5)$

(c) $Y \sim N(0, \sqrt{2})$, (d) $Y \sim Uniform(-3, 3)$

Answer: Accepted answers: (c) and (d)

Question 3. (1.5 + 1.5 marks) Let $X \sim Poisson(1)$ and $Y \sim Poisson(2)$ be independent. Then $\boxed{\mathbb{E}[X \mid X + Y = 5] =\quad}$ and $\boxed{Var[X \mid X + Y = 5] =\quad}$

Answer: The conditional distribution of X given $X + Y = 5$ is $Binomial(5, \frac{1}{1+2})$ (see Rohatgi & Saleh book). The conditional expectation is $\frac{5}{3}$ and conditional variance $\frac{10}{9}$.

Question 4. (2 + 1 marks) For $\alpha \in \mathbb{R}$, consider the functions $f, g : \mathbb{R} \to \mathbb{R}$ below.

$$f_\alpha(x) := \begin{cases} \alpha\, 4^{-x}, \forall x \in \{-1, 0, 1, 2, 3, \cdots\}, \\ 0, \text{ otherwise.} \end{cases}$$

$$g_\alpha(x) := \alpha \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-1)^2}{2}\right) + (1-\alpha)\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-3)^2}{2}\right), \forall x.$$

If f_α is a p.m.f., then $\boxed{\alpha =\quad}$ For this value of α, is g_α a p.d.f.? $\boxed{\text{Yes/No}}$

Answer: We have $\sum_{x=-1}^{\infty} 4^{-x} = 4 + \sum_{x=0}^{\infty} 4^{-x} = 4 + \frac{4}{3} = \frac{16}{3}$. Then $\alpha = \frac{3}{16}$, which also ensures f_α takes non-negative values and hence becomes a p.m.f.. For this value of α, g_α is a convex combination of two p.d.f.s and hence is a p.d.f.

Question 5. (3 marks) Let $X \sim Geometric(\frac{1}{3})$. Let $\alpha = \mathbb{E}\frac{1}{\Gamma(X+1)}$, where Γ denotes the Gamma function. Then $\boxed{\frac{1}{e}\alpha^3 - 2\mathbb{E}X =\quad}$

Answer:
$$\alpha = \sum_{n=0}^{\infty} \frac{1}{\Gamma(n+1)} \frac{1}{3} \left(\frac{2}{3}\right)^n = \frac{1}{3} \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n \frac{1}{n!} = \frac{1}{3} \exp\left(\frac{2}{3}\right).$$
Using results from Lecture notes, $\mathbb{E}X = 2$. The answer is
$$\frac{e}{27} - 4$$

Question 6. (3 marks) The following list contains some statements involving parametric statistical hypothesis testing problems. Put a tick (✓) beside all the correct option(s) to get credit.

(a) The null hypothesis H_0 is always true.
(b) If the observed sample realization falls inside the critical region of a test, then we accept the null hypothesis based on the test.
(c) The Neyman-Pearson Lemma gives the existence and uniqueness of an MP test of a given level α.
(d) If a test has level α, then type-I error of the test can not exceed α.

Answer: Accepted answers: (c) and (d)

Question 7. (1.5 + 1.5 marks) The DF F defined by
$$F(x) = \begin{cases} 0, \text{ if } x < 0, \\ \frac{1}{5} + \frac{x}{3}, \text{ if } 0 \leq x < 1, \\ \frac{1}{3} + \frac{x}{5}, \text{ if } 1 \leq x < 3, \\ 1, \text{ if } x \geq 3. \end{cases}$$
has jump(s) at $\boxed{}$ with jump size(s) $\boxed{}$

Answer: jumps at 0 and 3, with jump sizes $F(0) - F(0-) = \frac{1}{5} - 0 = \frac{1}{5}$ and $F(3) - F(3-) = 1 - \frac{1}{3} - \frac{3}{5} = \frac{1}{15}$.

Question 8. (3 marks) Let X, Y, Z be independent $N(0,1)$ RVs. Put a tick (✓) beside all correct statement(s) to get credit.

(a) Let F_X denote the DF of X. Then, $F_X(X) \sim Uniform(0,1)$.
(b) $\frac{X^2+Z^2}{2Y^2} \sim F_{2,2}$. (c) $X^2 + Z^2 \sim \chi^2_3$. (d) $\frac{X}{|Y|} \sim Cauchy(0,1)$

Answer: Accepted answers: (a) – Probability integral transform and (d) – Practice problem set 9 Question 5.
Note that $\frac{X^2+Z^2}{2Y^2} \sim F_{2,1}$ and $X^2 + Z^2 \sim \chi^2_2$.

Question 9. (3 marks) Let X be a non-negative RV defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that the MGF M_X exists on $\mathbb{R}$. Which of the following statement(s) is/are necessarily true? Put a tick (✓) beside all correct statement(s) to get credit.

(a) $\mathbb{P}(X > 1) = 1$.	(b) $\exp(\mathbb{E}X) \leq M_X(1)$.	(c) $\left(\mathbb{E}X^3\right)^2 > \mathbb{E}X^6$
(d) $\mathbb{P}(X \geq \alpha) \leq e^{-\lambda\alpha} M_X(\lambda)$ for all $\alpha > 0, \lambda > 0$.		

Answer: Accepted answers: (b) – by Jensen's inequality and (d) – by Chernoff's inequality Counter-example to (a): $X \sim Uniform(0,1)$. For (c): note that for any non-negative RV Y, we have $(\mathbb{E}Y)^2 \leq \mathbb{E}Y^2$; setting $Y = X^3$ we get the opposite inequality.

Question 10. (3 marks) Let $X \sim Uniform(-1, 3)$. Then, for any $p \in (0,1)$, the quantile of order p equals $\boxed{}$

Answer: We need to solve for unknown x in the equation $F_X(x) = p$. Now,

$$F_X(x) := \begin{cases} 0, \text{ if } x \leq -1, \\ \frac{x+1}{4}, \text{ if } -1 < x < 3, \\ 1, \text{ if } x \geq 3. \end{cases}$$

Solving the equation, we get the quantile of order p as $4p - 1$.

Question 11. (2 + 2 marks)

(a) Suppose that a random sample of size 16 has been drawn from a $N(\mu_1, 4)$ population, with $\mu_1 \in \mathbb{R}$ being unknown. If the sample mean obtained from the above sample is 4.2, then write down the equal-tailed 95% confidence interval for μ_1 based on the above random sample. $\boxed{}$

(b) Suppose that two random samples, one of size 16 and other of size 9 have been drawn from $N(\theta_1, 5.76)$ and $N(\theta_2, 5.76)$ populations respectively, with $\theta_1, \theta_2 \in \mathbb{R}$ being unknown. If the sample means for these random samples are 4 and 8 respectively, write down the equal-tailed 99% confidence interval for $\theta_1 - \theta_2$. $\boxed{}$

Answer: (a) For $\alpha = 0.05$, we have the solution to $\Phi(z_{\frac{1}{2}\alpha}) = 1 - \frac{\alpha}{2}$ as $z_{\frac{1}{2}\alpha} = 1.96$. The confidence interval is given by

$$\left[\bar{X}_n - \frac{\sigma}{\sqrt{n}} z_{\frac{1}{2}\alpha}, \bar{X}_n + \frac{\sigma}{\sqrt{n}} z_{\frac{1}{2}\alpha}\right] = [4.2 - 0.98, 4.2 + 0.98] = [3.22, 5.18]$$

open interval $(3.22, 5.18)$ is also acceptable as an answer

$[4.2 - \frac{1}{2} z_{\frac{1}{2}\alpha}, 4.2 + \frac{1}{2} z_{\frac{1}{2}\alpha}]$ with $\Phi(z_{\frac{1}{2}\alpha}) = 0.975$ or $1 - \frac{\alpha}{2}$ with $\alpha = 0.05$ gets 1 mark.

$[4.2 - \frac{1}{2} z_{\frac{1}{2}\alpha}, 4.2 + \frac{1}{2} z_{\frac{1}{2}\alpha}]$ without $\Phi(z_{\frac{1}{2}\alpha}) = 0.975$ gets no marks.

(b) For $\alpha = 0.01$, we have the solution to $\Phi(z_{\frac{1}{2}\alpha}) = 1 - \frac{\alpha}{2}$ as $z_{\frac{1}{2}\alpha} = 2.575$. The confidence interval is given by

$$\left[\bar{X}_n - \bar{Y}_m - \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} z_{\frac{1}{2}\alpha}, \bar{X}_n - \bar{Y}_m + \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} z_{\frac{1}{2}\alpha}\right].$$

Now, $\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} = \sqrt{\frac{5.76}{16} + \frac{5.76}{9}} = 1$. The confidence interval is given by

$$[4 - 8 - 2.575, 4 - 8 + 2.575] = [-6.575, -1.425]$$

open interval $(-6.575, -1.425)$ is also acceptable as an answer

$[-4 - z_{\frac{1}{2}\alpha}, -4 + z_{\frac{1}{2}\alpha}]$ with $\Phi(z_{\frac{1}{2}\alpha}) = 0.995$ or $1 - \frac{\alpha}{2}$ with $\alpha = 0.01$ gets 1 mark.

$[-4 - z_{\frac{1}{2}\alpha}, -4 + z_{\frac{1}{2}\alpha}]$ without $\Phi(z_{\frac{1}{2}\alpha}) = 0.995$ gets no marks.

Question 12. (2 + 2 marks) Let $X_1, X_2, \cdots, X_n$ be a random sample from $Binomial(1000, \theta)$ distribution, where $\theta \in (0, 1)$ is unknown. Write the MLE of θ^2 in the box.

Is the MLE of θ^2 consistent? Yes/No

Answer: Accepted answers: $\left(\frac{\bar{X}_n}{1000}\right)^2$ and 'Yes'.

The MLE of θ is $\frac{\bar{X}_n}{1000}$ (see Quiz 3 Question 4). Then the MLE of θ^2 is $\left(\frac{\bar{X}_n}{1000}\right)^2$. By WLLN, $\bar{X}_n \xrightarrow[n\to\infty]{P} 1000\theta$. By continuous mapping theorem for convergence in probability, we have $(\bar{X}_n)^2 \xrightarrow[n\to\infty]{P} 10^6\theta^2$ and hence the MLE of θ^2 is consistent.

Question 13. (4 marks) A 2-dimensional random vector $\binom{X}{Y}$ has the MGF $M_{X,Y}(t, s) = \exp(2t + s + \frac{1}{2}t^2 + 2s^2), \forall (t, s) \in \mathbb{R}^2$. Fix $\theta \in (0, \frac{\pi}{2})$ and define

$$U := (X - 2)\sin\theta + \frac{Y - 1}{2}\cos\theta, V := (X - 2)\cos\theta - \frac{Y - 1}{2}\sin\theta.$$

Put a tick (✓) beside all the correct option(s) to get credit.

(a) $X \sim N(0, 1)$. (b) $Cov(U, V) = 0$. (c) $X + Y \sim N(3, 5)$
(d) U and V are independent.

Answer: Accepted answers: (b), (c) and (d).

By identifying the joint distribution of $\binom{X}{Y}$ by looking at the joint MGF, we conclude that $\binom{X}{Y} \sim N_2(2, 1, 1, 4, 0)$ and X, Y are also independent. Then $X + Y \sim N(2+1, 1+4)$ and hence (c) is true. (b) and (d) follow from Practice set 9 Question 6.

Question 14. (1 + 1 + 2 marks) Let $\{X_n\}_n$ be a sequence of i.i.d. RVs defined on the same probability space.

(a) If $X_1 \sim Bernoulli(\frac{3}{4})$, then define $Y_n := \left(\frac{3n}{16}\right)^{-\frac{1}{2}} \left(X_1 + X_2 + \cdots + X_n - \frac{3n}{4}\right), \forall n$. It is known that $Y_n \xrightarrow[n\to\infty]{d} Z$. Then $Z \sim$

(b) If $X_1 \sim N(1,4)$, then define $Y_n := \frac{X_1+X_2+\cdots+X_n}{n}, \forall n$. It is known that $Y_n \xrightarrow[n\to\infty]{P} U$. Then the support of U equals $\boxed{}$

(c) If $X_1 \sim N(2,9)$, then define $Y_n := \sqrt{n}\left[\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)^2 - 4\right], \forall n$. It is known that $Y_n \xrightarrow[n\to\infty]{d} W$. Then $\boxed{W \sim }$

Answer: (a) by CLT, $Z \sim N(0,1)$.

(b) by WLLN, U is degenerate at 1 and hence the support is $\{1\}$.

Accepted answer: $\{1\}$. Just '1' is not acceptable.

(c) Take $g(x) = x^2, \forall x \in \mathbb{R}$. We have $g'(2) = 4 \neq 0$. By the delta method, $\frac{Y_n}{\sqrt{9}} \xrightarrow[n\to\infty]{d} V \sim N(0,4^2)$. Hence $W \stackrel{d}{=} \sqrt{9}V \sim N(0, 9\times 4^2) = N(0,144)$.

Question 15. (2 + 2 marks) Let X_1 and X_2 be independent *Geometric*(p) RVs for some $p \in (0,1)$. Let $f_{X_{(2)}}$ denote the p.m.f. of $X_{(2)}$.

(a) For integers $k > 0$, $\boxed{f_{X_{(2)}}(k) = }$

(b) Are $X_{(1)}$ and $X_{(2)} - X_{(1)}$ independent? $\boxed{\text{Yes/No}}$

Answer: Accepted answers: $\left(1-(1-p)^{k+1}\right)^2 - \left(1-(1-p)^k\right)^2$ and 'Yes'.

(a) The common p.m.f. of X_1 and X_2 is

$$f(x) = \begin{cases} p(1-p)^x, \text{ if } x \in \{0,1,2,\cdots\}, \\ 0, \text{ otherwise.} \end{cases}$$

Now, for any non-negative integer k, using the independence of X_1 and X_2

$$\mathbb{P}(X_{(2)} \leq k) = \mathbb{P}(X_1 \leq k, X_2 \leq k) = (\mathbb{P}(X_1 \leq k))^2.$$

Since,

$$\mathbb{P}(X_1 \leq k) = \sum_{x=0}^{k} p(1-p)^x = 1-(1-p)^{k+1},$$

we have,

$$\mathbb{P}(X_{(2)} \leq k) = \left(1-(1-p)^{k+1}\right)^2$$

and hence for positive integers k,

$$\mathbb{P}(X_{(2)} = k) = \mathbb{P}(X_{(2)} \leq k) - \mathbb{P}(X_{(2)} \leq k-1) = \left(1-(1-p)^{k+1}\right)^2 - \left(1-(1-p)^k\right)^2.$$

(b) For non-negative integers m, n, we have

$$
\begin{aligned}
&\mathbb{P}(X_{(1)} = n, X_{(2)} - X_{(1)} = m) \\
&= \mathbb{P}(X_{(1)} = n, X_{(2)} = m + n) \\
&= \begin{cases} \mathbb{P}(X_1 = n, X_2 = m + n) + \mathbb{P}(X_2 = n, X_1 = m + n), \text{ if } m > 0, \\ \mathbb{P}(X_1 = X_2 = n), \text{ if } m = 0 \end{cases} \\
&= \begin{cases} 2p(1-p)^n \, p(1-p)^{m+n}, \text{ if } m > 0, \\ \left(p(1-p)^n\right)^2, \text{ if } m = 0 \end{cases} \\
&= \begin{cases} 2p^2(1-p)^{2n}(1-p)^m, \text{ if } m > 0, \\ \left(p(1-p)^n\right)^2, \text{ if } m = 0 \end{cases}
\end{aligned}
$$

Since, the joint p.m.f. is in a product form, we conclude that $X_{(1)}$ and $X_{(2)} - X_{(1)}$ are independent.