Solutions to MSO205 Mid Semester Examination

Question 1. (3 marks) A secretary types 4 letters and prepares 4 corresponding envelopes. In a hurry, she places one letter in each envelope at random. What is the probability that at least one letter is in the correct envelope? You may write your answer as a fraction.

Answer: Accepted answer: $\frac{5}{8}$ ($\frac{15}{24}$ is also accepted)

For $i = 1, 2, 3, 4$, consider the events E_i that the i-th letter is in the correct envelope. We need to find $\mathbb{P}(E_1 \cup E_2 \cup E_3 \cup E_4)$.

Since one letter has been placed in each envelope, this is an experiment performed at random without replacement. Therefore for any $i = 1, 2, 3, 4$, in the event E_i, the other letters can be placed in any of the remaining $4 - 1 = 3$ envelopes in $3!$ ways. But the total number of ways in which 4 letters are distributed in 4 envelopes is $4!$ and hence $\mathbb{P}(E_i) = \frac{3!}{4!} = \frac{1}{4}$. (this argument – part marks: 1)

More generally, if $1 \leq k \leq 4$ letters are in the correct envelope, then remaining $4-k$ letters can be placed in the remaining $4-k$ envelopes with probability $\frac{(4-k)!}{4!}$. For example, for $1 \leq i < j \leq 4$, the probability that i and j-th letter are correctly placed is $\mathbb{P}(E_i \cap E_j) = \frac{2!}{4!}$. Similarly, for $1 \leq i < j < k \leq 4$, $\mathbb{P}(E_i \cap E_j \cap E_k) = \frac{1!}{4!}$ and $\mathbb{P}(E_1 \cap E_2 \cap E_3 \cap E_4) = \frac{0!}{4!}$. (this argument – part marks: 1)

By the Inclusion-Exclusion principle, (mentioning Inclusion-Exclusion – part marks: 0.5)

$$\mathbb{P}(E_1 \cup E_2 \cup E_3 \cup E_4) = \binom{4}{1} \times \frac{3!}{4!} - \binom{4}{2} \times \frac{2!}{4!} + \binom{4}{3} \times \frac{1!}{4!} - \binom{4}{4}\frac{0!}{4!} = 1 - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} = \frac{15}{24} = \frac{5}{8}.$$

(final calculation – part marks: 0.5)

Question 2. (2 marks) Show by an example that pairwise independence of events does not imply mutual independence of the same events.

Answer: Consider a random experiment $\mathcal{E}$ with sample space $\Omega = \{1, 2, 3, 4\}$ and event space $\mathcal{F} = 2^\Omega$. If the outcomes are equally likely, then we have the probability function/measure $\mathbb{P}$ determined by the information $\mathbb{P}(\{\omega\}) = \frac{1}{4}, \forall \omega \in \Omega$. Consider the events $A = \{1, 4\}, B = \{2, 4\}, C = \{3, 4\}$. Then $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$. (explicit description of the probability space, including the sample space and the specific events – part marks: 1)

Moreover, $A \cap B = B \cap C = C \cap A = \{4\}$ and hence $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap C) = \mathbb{P}(C \cap A) = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(B)\mathbb{P}(C) = \mathbb{P}(C)\mathbb{P}(A)$. Therefore, the events A, B, C are pairwise independent. (this verification – part marks: 0.5)

However, $A \cap B \cap C = \{4\}$ and hence $\mathbb{P}(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$. Here, the events A, B, C are not mutually independent. (this verification – part marks: 0.5)

Question 3. (2 marks) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space such that there exist events A and B with $\mathbb{P}(A) = 0.9, \mathbb{P}(B) = 0.2$. Is it true that $\mathbb{P}(A|B) \geq 0.3$? Justify your answer.

Answer: By Bonferroni's inequality, we have (this argument – part marks: 1)

$$\mathbb{P}(A \cap B) \geq \mathbb{P}(A) + \mathbb{P}(B) - 1 = 0.1.$$

Hence, $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \geq \frac{0.1}{0.2} = 0.5$. It is true that $\mathbb{P}(A|B) \geq 0.3$. (final calculation – part marks: 1)

Question 4. (3 + 1 marks) Let X be an RV with the distribution function F_X. Consider the following two functions $G_1, G_2 : \mathbb{R} \to \mathbb{R}$ defined by

$$G_1(x) := (F_X(x))^2, \quad G_2(x) := G_1(x) + \frac{1}{2}, \forall x \in \mathbb{R}.$$

Can G_1 and G_2 be distribution functions? Justify your answer.

Answer: Since, F_X is a given distribution function, we have that F_X is non-decreasing, right-continuous and

$$\lim_{x \to -\infty} F_X(x) = 0, \quad \lim_{x \to \infty} F_X(x) = 1.$$

Again, these conditions are sufficient to ensure that a function is a distribution function.

We verify these properties for G_1.

For $x \in \mathbb{R}$,

$$\lim_{y \downarrow x} G_1(y) = \lim_{y \downarrow x} (F_X(y))^2 = (F_X(x))^2 = G_1(x).$$

Here, we have used the right-continuity of F_X and have shown that G_1 is right-continuous. (this verification – part marks: 1)

Again, for any $x < y$, $0 \leq F_X(x) \leq F_X(y) \leq 1$ and consequently, $0 \leq G_1(x) = (F_X(x))^2 \leq (F_X(y))^2 = G_1(y) \leq 1$. Thus, G_1 is non-decreasing. (this verification – part marks: 0.5)

Now,

$$\lim_{x \to -\infty} G_1(x) = \lim_{x \to -\infty} (F_X(x))^2 = 0, \quad \lim_{x \to \infty} G_1(x) = \lim_{x \to \infty} (F_X(x))^2 = 1.$$

(this verification – part marks: 0.5 + 0.5)

Since, G_1 has the necessary properties, G_1 is a distribution function. (this statement – part marks: 0.5)

Now,

$$\lim_{x \to -\infty} G_2(x) = \lim_{x \to -\infty} \left[G_1(x) + \frac{1}{2} \right] = \frac{1}{2} \neq 0.$$

(this verification – part marks: 0.5)

Since the above necessary condition fails for G_2, it is not a distribution function. (this verification – part marks: 0.5)

An alternative: Now,

$$\lim_{x\to\infty} G_2(x) = \lim_{x\to\infty}\left[G_1(x)+\frac{1}{2}\right] = 1+\frac{1}{2} \neq 1.$$

(this verification – part marks: 0.5)

Since the above necessary condition fails for G_2, it is not a distribution function.

Note: Working with some explicit F_X will not be given credit.

Question 5. (4 + (1.5 + 1.5) + 3 + (2 + 1) marks) Let X be a continuous RV with the following p.d.f. $f_X : \mathbb{R} \to [0,\infty)$ defined by

$$f_X(x) := \begin{cases} \exp(-x), \text{ if } x > 0, \\ 0, \text{ otherwise} \end{cases}$$

Consider the RV $Y := |X-1|$. Answer the following sub-questions with full justification.

(a) Compute the distribution function of Y.

(b) Show that Y is a continuous RV and compute its p.d.f.

(c) Identify the support of Y.

(d) Compute a quantile of order 0.9 for Y. Is it unique? You may use the fact that $\exp(-2) \approx 0.1353$. You may write your answer in terms of a natural logarithm, i.e. $\log_e$ or ln.

Answer: We have, for all $y \in \mathbb{R}$, (part marks: 1)

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(|X-1| \le y) = \begin{cases} 0, \text{ if } y < 0, \\ \mathbb{P}(1-y \le X \le y+1), \text{ if } y \ge 0 \end{cases}$$

For $y \ge 0$,

$$\begin{aligned}\mathbb{P}(1-y \le X \le y+1) &= \int_{1-y}^{y+1} f_X(x)\,dx \\ &= \begin{cases} \int_{1-y}^{y+1} \exp(-x)\,dx, \text{ if } 0 \le y < 1, \\ \int_0^{y+1} \exp(-x)\,dx, \text{ if } y \ge 1 \end{cases} \\ &= \begin{cases} e^{-(1-y)} - e^{-(y+1)}, \text{ if } 0 \le y < 1, \\ 1 - e^{-(y+1)}\,dx, \text{ if } y \ge 1 \end{cases}\end{aligned}$$

Therefore, (part marks: 3)

$$F_Y(y) = \begin{cases} 0, \text{ if } y < 0, \\ e^{(y-1)} - e^{-(y+1)}, \text{ if } 0 \le y < 1, \\ 1 - e^{-(y+1)}\,dx, \text{ if } y \ge 1 \end{cases}$$

Now, F_Y is continuous with

$$F'_Y(y) = \begin{cases} 0, \text{ if } y < 0, \\ e^{(y-1)} + e^{-(y+1)}, \text{ if } 0 < y < 1, \\ e^{-(y+1)}, \text{ if } y > 1 \end{cases}$$

with possible discontinuities at the points 0 and 1. Also, (this verification – part marks: 1.5)

$$\int_{-\infty}^{\infty} F'_Y(y)\,dy = \int_0^1 [e^{(y-1)} + e^{-(y+1)}]\,dy + \int_1^{\infty} e^{-(y+1)}\,dy = (1-e^{-1}) + (-e^{-2}+e^{-1}) + (-0+e^{-2}) = 1.$$

Therefore, Y is a continuous RV with the p.d.f. (part marks: 1)

$$f_Y(y) = \begin{cases} 0, \text{ if } y \leq 0, \\ e^{(y-1)} + e^{-(y+1)}, \text{ if } 0 < y < 1, \\ 0, \text{ if } y = 1, \\ e^{-(y+1)}, \text{ if } y > 1 \end{cases}$$

(Note: values at 0 and 1 can be chosen arbitrarily (part marks for specifying values at 0 and 1 to complete the definition of a p.d.f.: 0.5))

The support of Y is given by (part marks: 0.5 for recalling the definition)

$$\{y \in \mathbb{R} : \mathbb{P}(y-h < Y \leq y+h) > 0, \forall h > 0\}.$$

Note that $\mathbb{P}(y-h < Y \leq y+h) = F_Y(y+h) - F(y-h) = \int_{y-h}^{y+h} f_Y(t)\,dt$ for all $h > 0$. We now check various cases. By looking at the sign of F'_Y, we conclude that F_Y is strictly increasing on $(0,\infty)$.

(a) If $y < 0$, then choosing $0 < h < |y|$, we have $F_Y(y+h) - F(y-h) = 0 - 0 = 0$. Therefore, such y is not in the support. (part marks: 1)

(b) If $y = 0$, then for any $h > 0$, $F_Y(y-h) = 0$ and $F_Y(y+h) > 0$. Consequently, 0 is in the support. (part marks: 0.5)

(c) As F_Y is strictly increasing on $(0,\infty)$, for $y \in (0,\infty)$, we have for all $h > 0$

$$\mathbb{P}(y-h < Y \leq y+h) \geq \mathbb{P}(y-h_0 < Y \leq y+h_0) = F_Y(y+h_0) - F(y-h_0) > 0$$

with $h_0 = \min\{y, h\}$. Therefore, $y \in (0,\infty)$ is in the support. (part marks: 1)

Therefore, the support of Y is $[0,\infty)$.

Observe that $F_Y(1) = 1-e^{-2} \approx 1-0.1353 \approx 0.8647 < 0.9$. As F_Y is non-decreasing, a quantile of order 0.9 for the continuous RV Y is greater than 1 and we need to solve $F_Y(y) = 1 - e^{-(y+1)}$ for $y > 1$ (this justification – part marks: 1). Now, a quantile y of order 0.9 needs to satisfy

$F_Y(y) = 0.9$, which gives

$$1 - e^{-(y+1)} = 0.9,$$

or $(y+1) = -\ln(0.1) = \ln 10$ or $y = \ln 10 - 1$ (part marks: 1). As we have only this solution to $F_Y(y) = 0.9$ (F_Y is strictly increasing on $(1, \infty)$), the quantile is unique. (part marks: 1)

Question 6. (2 + (1 + 1) marks) Consider a discrete RV X with the p.m.f.

$$f_X(x) := \begin{cases} \frac{1}{3}\left(\frac{2}{3}\right)^x, & \text{if } x \in \{0, 1, 2, \cdots\}, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Justify the existence of MGF of X and compute it.

(b) Using the MGF obtained in part (a), compute the mean and variance of X.

Answer: We have

$$\mathbb{E}e^{tX} = \sum_{k=0}^{\infty} e^{tk} \frac{1}{3}\left(\frac{2}{3}\right)^k = \frac{\frac{1}{3}}{1 - \left(\frac{2}{3}\right)e^t},$$

exists if $1 - \left(\frac{2}{3}\right)e^t > 0$ or equivalently, $t < \ln\left(\frac{3}{2}\right)$. In particular, these expectations exist for a neighbourhood $(-\ln\left(\frac{3}{2}\right), \ln\left(\frac{3}{2}\right))$ of 0 (this justification – part marks: 1). Hence, the MGF exists and equals (this verification, including the range of t clearly specified – part marks: 1)

$$M_X(t) = \frac{\frac{1}{3}}{1 - \left(\frac{2}{3}\right)e^t}, -\infty < t < \ln\left(\frac{3}{2}\right).$$

Note that for such values of t,

$$M_X'(t) = \frac{1}{3}\frac{1}{(1 - \left(\frac{2}{3}\right)e^t)^2}\left(\frac{2}{3}\right)e^t, \quad M_X''(t) = \frac{2}{3}\frac{1}{(1 - \left(\frac{2}{3}\right)e^t)^3}\left(\frac{2}{3}\right)^2 e^{2t} + \frac{1}{3}\frac{1}{(1 - \left(\frac{2}{3}\right)e^t)^2}\left(\frac{2}{3}\right)e^t.$$

Evaluating at $t = 0$, we have $\mathbb{E}X = M_X'(0) = 2, \mathbb{E}X^2 = M_X''(0) = 8 + 2 = 10$. Then $Var(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 6$. (part marks: 1 + 1)