

Time Allowed: 50+ min

Max Marks: 112

Answer all questions. All questions carry equal marks. Mark your answers on this sheet. Circle the appropriate letter for the True/False type choices. State and make any assumptions that you need.

Question 1. A bio-experiment consists of a sequence of SI trials. Each trial has 4 outcomes B, H, T and E . At the i -th trial, $i = 1, 2, \dots$, $P(B) = 0.2 - 1/(10i)$, $P(H) = 0.2 + 1/(10i)$, $P(T) = 0.3 + 1/(10i)$ and $P(E) = 0.3 - 1/(10i)$.

A) The trials are Bernoulli trials. **T or F**

B) The trials are Bernoulli trials if we consider only the occurrence of H . **T or F**

C) $S =$ 'success' is said to occur if either T or E occurs in any trial. The trials are Bernoulli trials if we consider only the occurrence of S . **T or F**

D) If 100 trials are performed, write an expression for the probability of getting at least 2 H .

A) Not so as more than 2 possible outcomes.

B) Not so as prob is not same for all trials.

C) $P(S) = 0.6$. All conditions of Bernoulli trials are satisfied.

D) $P(\text{getting } \geq 2 H) = 1 - P(0 H) - P(1 H)$.

$$P(0 H) = \prod_{i=1}^{100} \left(1 - \left(0.2 + \frac{1}{10i} \right) \right), \quad P(1 H) = \sum_{i=1}^{100} \left(0.2 + \frac{1}{10i} \right) \prod_{j=1, j \neq i}^{100} \left(1 - \left(0.2 + \frac{1}{10j} \right) \right).$$

$$P(0 H) = P(\text{Not } H \cdot \text{Not } H \cdot \dots \cdot \text{Not } H) = P(\text{Not } H) \cdot P(\text{Not } H) \cdot \dots \cdot P(\text{Not } H).$$

$$\text{Similarly, } P(1 H) = P([H \cdot \text{Not } H \cdot \dots \cdot \text{Not } H] \text{ OR } [\text{Not } H \cdot H \cdot \dots \cdot \text{Not } H] \text{ OR } \dots [\text{Not } H \cdot \text{Not } H \cdot \dots \cdot H]) \\ = P(H \cdot \text{Not } H \cdot \dots \cdot \text{Not } H) + P(\text{Not } H \cdot H \cdot \dots \cdot \text{Not } H) + \dots + P(\text{Not } H \cdot \text{Not } H \cdot \dots \cdot H)$$

Question 2. Let X_1, X_2, \dots , be a sequence of binomial RVs with $N = 10$ and $p = 0.002$; Y_1, Y_2, \dots , a sequence of Poisson RVs with $\lambda = 10$; and Z_1, Z_2, \dots , a sequence of standard normal RVs. All RVs are SI.

A) If M is large then $(X_1 + X_2 + \dots + X_M) + Z_{100}$ is Gaussian. **T or F**

B) If M is large then $(Y_1 + X_1) + (Y_2 + X_2) + \dots + (Y_M + X_M)$ is approximately Poisson. **T or F**

C) $X_{99} + (Z_1 + Z_2 + \dots + Z_M)$ is a discrete valued RV. **T or F**

D) If M is large and $A = \prod_{i=1}^M (X_i + Y_i + Z_i)^{1234}$, write an expression for the pdf of a RV R defined as $R = \log A$.

A) If M is large then $X = (X_1 + X_2 + \dots + X_M)$ is Gaussian (CLT) and hence $X + Z_{100}$ is Gaussian as well.

B) If M is large then $X = (X_1 + X_2 + \dots + X_M)$ which is binomial can be approximated by Poisson (N large, p small). Also $Y = (Y_1 + Y_2 + \dots + Y_M)$ is Poisson. Overall sum is also Poisson as sum of two SI Poisson RVs is also a Poisson RV.

C) X_{99} is discrete while $Z = (Z_1 + Z_2 + \dots + Z_M)$ is continuous. Overall pdf of the sum is continuous. You may verify this by writing the pdf of the sum as convolution of two pdfs, one discrete and other continuous.

D) $R = \log A = 1234 \sum_{i=1}^M (X_i + Y_i + Z_i)$. This is clearly a sum of M SI RVs R_1, R_2, \dots where $R_i = X_i + Y_i + Z_i$.

Hence for large M , R is a Gaussian RV. Its pdf is

$$f(R) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(r - \mu)^2}{2\sigma^2}\right], \quad -\infty < r < \infty.$$

Question 3. A bag contains 5 red, 10 black, and 15 white balls. 10 balls are drawn randomly one ball at a time, first 5 with replacement and the next 5 without replacement. Their color is noted.

A) $P(\text{first 5 balls are red}) > P(\text{last 5 balls are red})$. **T or F**

B) Find $P(\text{all 10 balls are red})$.

C) Write an expression for $P(\text{last 5 balls are red} \mid \text{first 7 balls are red})$.

D) Write an expression for $P(\text{first 7 balls are red} \mid \text{last 5 balls are red})$.

A) $(5/30)^5$ vs $1/^{30}C_5$. $P(5 \text{ balls are red without replacement}) = (5/30) \cdot (4/29) \cdot (3/28) \cdot (2/27) \cdot (1/26) = 1/^{30}C_5$.

B) $P(\text{all 10 balls are red}) = P(\text{1st 5 are red}) \cdot P(\text{last 5 are red}) = (5/30)^5 \cdot 1/^{30}C_5$.

C) The first 5 drawings have no impact on the last 5 drawings. Given first 7 balls are red, means 6th and 7th balls are red. Hence

$P(6\text{-}10^{\text{th}} \text{ balls are red} \mid 6\text{-}7 \text{ balls are red}) = P(6\text{-}10^{\text{th}} \text{ balls are red}) / P(6\text{-}7 \text{ balls are red}) = [1/^{30}C_5] / [(5/30)(4/29)] = P(8\text{-}10^{\text{th}} \text{ balls are red}) = (3/28) \cdot (2/27) \cdot (1/26)$.

D) This is simply $P(\text{first 5 are red}) = (5/30)^5$.

Question 4. Let X and Y be jointly Gaussian RV with $X \sim N(0.1, 0.05)$, $Y \sim N(0.2, 0.05)$, and $\text{cov}(X, Y) = 0.03$. Further, it is known that the sum of jointly Gaussian RV is Gaussian.

A) $E(X + Y) = 0.3$.

T or F

B) $E[(X + Y)^2] = 0.1$.

T or F

C) $P(X + Y < 0.3) = 0.35$.

T or F

D) Write an expression for the pdf of RV Z defined as $Z = X + Y$.

A) Expected value of the sum is sum of the expected values.

B) $E[(X + Y)^2] = E(X^2) + E(Y^2) + 2E(XY) = (0.05 + 0.1^2) + (0.05 + 0.2^2) + 2(0.03 + 0.1 \times 0.2) = 0.25$.

C) Since $X + Y$ is Gaussian with mean 0.3, $P(X + Y < 0.3) = 0.5$.

D) Z is Gaussian with mean equal to 0.3 and variance equal to $0.25 - 0.3^2 = 0.25 - 0.09 = 0.16$.

Question 5. In a data communication system, the transmitted signal S and received signal R are related as $R = S + N$, where N is the noise in the channel modeled as a standard normal RV. It is given that S is either 10 Volt for logical '1' or -10 Volt for logical '0'. These logical values occur with equal probability.

A) $E(R) = 0$.

T or F

B) The RVs R and S are SI.

T or F

C) Sketch the pdf of S and hence the pdf of R .

D) How should the receiver process R to declare whether '1' or '0' was transmitted?

A) $E(S) = E(N) = 0$; hence $E(R) = E(S + N) = E(S) + E(N) = 0$.

B) If $S = 10$, R is Gaussian with mean = 10 and variance equal to 1. If $S = -10$, R is Gaussian with mean = -10 and variance equal to 1. Hence the pdf of R changes for different values of S .

C) pdf of S : two impulses of height 0.5 each at $s = 10$ and $s = -10$, respectively. The pdf of R is convolution of pdf of S with pdf of N . This gives pdf of R as

$$f(R) = 0.5 \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(r-10)^2}{2}\right] + 0.5 \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(r+10)^2}{2}\right], -\infty < r < \infty.$$

D) If R is positive pick logical '1' and if R is negative, pick logical '0'.

Question 6. A fax machine is designed to communicate outcome of a black (B) and white (W) scan. Each sheet consists of 1000 pixels. These colors occur in an SI manner on a sheet. The design is optimised for $p = \text{prob}(B) = 0.01$. The fax machine input/output behaviour is as follows:

Input scan: BB, Output bits sent on channel: 111; Input scan: BW, Output bits sent on channel: 110

Input scan: WB, Output bits sent on channel: 10; Input scan: WW, Output bits sent on channel: 0

A) What is the expected number of bits per page sent on the channel?

B) Explain why a completely black sheet will require a lot more time than a completely white sheet.

C) If instead we use no coding, ie, output sent on the channel is same as the input scan, what is the expected number of bits per page.

D) Justify the above described coding used by the fax machine.

A) There are 500 pairs of input scans for a total of 1000 pixels. For 1 pair, the number of bits sent on the channel is a RV with pdf

$$L = 3, \text{prob}(L = 3) = \text{prob}(\text{BB or BW}) = \text{prob}(\text{BB}) + \text{prob}(\text{BW}) = 0.01 \times 0.01 + 0.01 \times 0.99 = 0.01$$

$$L = 2, \text{prob}(L = 2) = \text{prob}(\text{WB}) = 0.99 \times 0.01 = 0.0099 = 0.01 \text{ (approx)}$$

$$L = 1, \text{prob}(L = 1) \text{prob}(\text{WW}) = 0.99 \times 0.99 = 0.98$$

$$E(\text{bits per two scans}) = 3 \times 0.01 + 2 \times 0.01 + 1 \times 0.98 = 1.03$$

$$E(\text{bits per page}) = 500 \times 1.03 = 515 \text{ bits.}$$

B) A completely black page will require $500 \times 3 = 1500$ bits (111 is sent out for each pair of BB) while a completely white sheet will require $500 \times 1 = 500$ bits (0 is sent out for each pair of WW).

C) There are 1000 bits in one page, one for each pixel.

D) The coding used by fax machine minimizes the number of bits sent on the channel to an average of 515 bits/page. This is based on the assumption that in a scan $\text{prob}(\text{W}) \gg \text{prob}(\text{B})$ which is indeed the case in most real situations.

Question 7. Let $h(x)$ and $g(y)$ be marginal pdf of X and Y , respectively, and $f(x, y)$ be the joint pdf of (X, Y) .

A) If X and Y are SI then $\text{MGF}(X^2 + Y) = \text{MGF}(X^2) \text{MGF}(Y)$.

T or F

B) If $f(x, y) = A \exp(x^2 + y)$, $0 < x < 1$, $0 < y < 1 \Rightarrow X$ and Y are SI.

T or F

C) Five uncorrelated RVs that are jointly Gaussian are not SI.

T or F

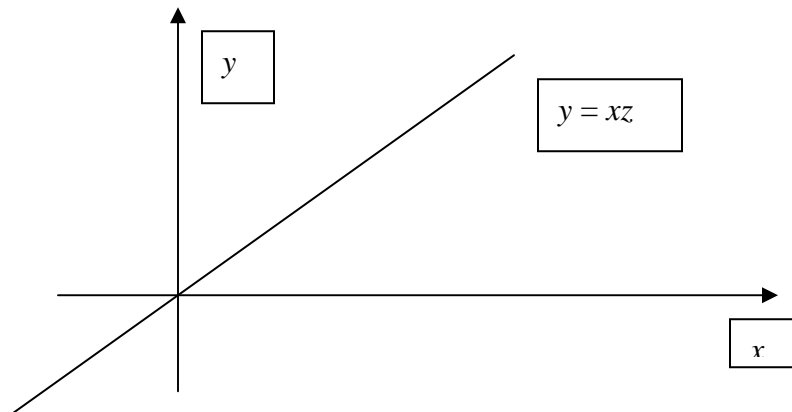
D) Define a RV $Z = Y/X$. Sketch the region corresponding to the event $Z \leq z$.

A) $\text{MGF}(X^2 + Y) = E(\exp[s(X^2 + Y)]) = E(\exp[sX^2] \cdot \exp[sY]) = E(\exp[sX^2]) \cdot E(\exp[sY]) = \text{MGF}(X^2) \text{MGF}(Y)$.

B) Straightforward

C) If Gaussian RVs are uncorrelated, they are SI.

D) $Z \leq z \Rightarrow Y/X \leq z \Rightarrow Y \leq Xz$. This is the region below the curve $Y = Xz$. This curve is a straight line that passes thru the origin and has a slope of z .



Matriculation Number: _____