Question 1: (20%) The probability assignment of a two dimensional random variable (X, Y) is as follows.

$P(X = x, Y = y)$	$Y = -1$	$Y = 0$
$X = 1$	0.2	0.5
$X = 0$	0.1	0.2

1. (5%) Find out the marginal pmfs $P(X = 0)$, $P(X = 1)$, $P(Y = -1)$, and $P(Y = 0)$.
2. (2%) Are X and Y independent? Use one sentence to justify your choice.
3. (5%) Let $Z = XY$. Find the probability mass function of Z.

   Hint: Find out the sample space of Z first.
4. (2%) Find the conditional probability that $P(X = 1|Z = 0)$.
5. (1%) Are X and Z independent? Use the result from the previous sub-question to justify your choice. One sentence is sufficient.
6. (5%) Find the correlation between X and Z, namely, find $\mathsf{E}(XZ)$. Are X and Z orthogonal?

1. $P(X=0) = 0.1+0.2 = 0.3$
$P(X=1) = 0.2+0.5 = 0.7$
$P(Y=-1) = 0.2+0.1 = 0.3$
$P(Y=0) = 0.5+0.2 = 0.7$

2. No: Since $P(X=1, Y=-1) = 0.2 \neq P(X=1)\cdot P(Y=-1) = 0.09$

3. Z takes value in $\{-1, 0\}$.
$P(Z=-1) = 0.2$
$P(Z=0) = 0.1+0.5+0.2 = 0.8$

4. $P(X=1|Z=0) = \frac{P(X=1, Z=0)}{P(Z=0)} = \frac{0.5}{0.8} = \frac{5}{8}$

5. No: since $P(X=1|Z=0) \neq P(X=1)$

6. $E(XZ) = E(X^2Y) = (-1)\times 0.2 + 0\times 0.5 + 0\times 0.1 + 0\times 0.2$
$= -0.2$
in which $(-1) = (1)^2\times(-1)$, $0 = 1^2\times 0$, $0 = 0^2\times(-1)$, $0 = 0^2\times 0$.
No, X & Z are not orthogonal

Question 2: (30%) X_1 is Bernoulli distributed with parameter p, and so are $X_2, X_3, \cdots, X_n$. Furthermore, all $X_1, \cdots, X_n$ are independent.

1. (5%) Find the mean $\mathsf{E}(X_1)$, variance $\mathsf{Var}(X_1)$, and the second moment $\mathsf{E}(X_1^2)$ of X_1.
2. (7%) Let $Z = \sum_{i=1}^{n} X_i$. Find the mean $\mathsf{E}(Z)$, variance $\mathsf{Var}(Z)$, and the second moment $\mathsf{E}(Z^2)$ of Z. Hint: you can do it by either one of the following two methods. First, by the properties of sum of random variables. Or second, using the fact that Z is a binomial distribution with parameters (n, p), and the pmf of a binomial distribution is $\mathsf{P}(Z = k) = \binom{n}{k} p^k (1-p)^{n-k}$.
3. (4%) Find the characteristic function $\Phi_{X_1}(\omega)$ of X_1, which is the same as $\Phi_{X_2}(\omega)$, $\Phi_{X_3}(\omega)$, etc. Use that result to find the characteristic function $\Phi_Z(\omega)$ of Z.
4. (7%) Use the moment theorem to verify your answer of $\mathsf{E}(Z)$ computed in the second sub-question.
5. (7%) Let $Y = aZ + b$, where a and b are constants. Find the characteristic function $\Phi_Y(\omega)$ of Y in terms of the characteristic function of $\Phi_Z(\omega)$.

1. $E(X_1) = (1-p)\cdot 0 + p \times 1 = p$

$E(X_1^2) = (1-p)\times 0^2 + p \times 1^2 = p$

$Var(X_1) = E(X_1)^2 - (E(X))^2 = p(1-p)$

2. $E(Z) = \sum_{i=1}^{n} E(X_i) = np$

$Var(Z) = \sum_{i=1}^{n} Var(X_i) = np(1-p)$

$E(Z^2) = Var(Z) + (E(Z))^2 = np(1-p) + n^2p^2$

3. $\Phi_{X_1}(w) = E(e^{jwX_1}) = e^{jw\cdot 0} \times (1-p) + e^{jw\cdot 1} \times p$

$= (1-p) + e^{jw} \times p$

$\Phi_Z(w) = (\Phi_{X_1}(w))^n = (1-p+pe^{jw})^n$

4. $E(Z) = \frac{1}{j}\left[\frac{d}{dw}\Phi_Z(w)\right]_{w=0}$

$= \frac{1}{j} \times \left[n(1-p+pe^{jw}) \times p \times e^{jw} \times j\right]_{w=0}$

$= np$ — the same as in 2.

5. $\Phi_Y(\omega) = E(e^{j\omega Y})$

$$= E(e^{j\omega(aZ+b)})$$

$$= E(e^{j\omega aZ} \cdot e^{j\omega b})$$

$$= e^{j\omega b}\, \Phi_Z(a\omega)$$ #

Question 3: (15%) Professor Wang's computer crashed, and all the data about the grades of the first midterm was lost. He only remembered that the average of the first midterm was 50, and the total number of enrolled students was 60.

1. (5%) What is the maximum number of students having grades better than 90 based on the Markov inequality? The answer can be expressed as a decimal or fractional number. For example, you can write the maximum number of students is 2.75 or 11/4.

2. (5%) The TA reminded Professor Wang that the standard deviation of the first midterm grade distribution was 10, namely, the variance was 10^2. With this additional information, answer the first question again using the Chebyshev inequality.

3. (5%) One student also remembered that the grade distribution was very much like a Gaussian distribution. So if Professor Wang assumed that the distribution was Gaussian with mean 50 and variance $10^2 = 100$, what is the approximate number of students having grades better than 90?

   Hint: express your result using $F_Z(z) = \mathsf{P}(Z \leq z)$, the cumulative distribution function (cdf) of a standard Gaussian with mean 0 and variance 1.

1. $$P(X \geq 90) \leq \frac{E(X)}{90} = \frac{50}{90} = \frac{5}{9}.$$

   # of students is less than $\frac{5}{9} \times 60 = \frac{100}{3}$

2. $$P(X \geq 90) \leq P(X \leq 10 \text{ or } X \geq 90)$$
   $$= P(|X-50| \geq 40)$$
   $$\leq \frac{Var(X)}{40^2} = \frac{100}{1600} = \frac{1}{16}.$$

   # of students is less than $\frac{1}{16} \times 60 = \frac{15}{4}$

3. $$P(X \geq 90) = P\left(\frac{X-E(X)}{\sqrt{Var(X)}} \geq \frac{90-E(X)}{\sqrt{Var(X)}}\right)$$
   $$= P\left(\frac{X-E(X)}{\sqrt{Var(X)}} \geq \frac{90-50}{10}\right)$$
   $$\approx P(Z \geq 4)$$ where Z is a standard Gaussian
   $$= 1 - F_Z(4)$$, # of students: $60(1-F_Z(4))$

Question 4: (25%) Consider an additive Gaussian noise channel $Y = X + N$, in which the input X is Gaussian distributed with mean 0 and variance a^2, and N is a standard Gaussian with mean 0 and variance 1. X and N are independent.

1. (2%) Whether is Y a Gaussian random variable? Make your yes-no choice and justify it by one sentence.
2. (5%) Find $\mathsf{E}(Y)$ and $\mathsf{Var}(Y)$, and then write down the marginal pdf of Y, $f_Y(y)$, using the previous result.
3. (8%) Find the covariance $\mathsf{Cov}(X, Y)$ and the correlation coefficient ρ between X and Y.
4. (5%) Find the coefficients a^* and b^* of the minimum mean square error "linear" estimator for X, $f_{MMSE,lin}(Y) = a^*Y + b^*$.
5. (5%) Find out the mean square error achieved by $f_{MMSE,lin}(Y)$. Namely, evaluate $\mathsf{E}\left((X - f_{MMSE,lin}(Y))^2\right)$.

1. Yes: since Y is a linear combination of X & N.

2. $E(Y) = E(X) + E(N) = 0 + 0 = 0.$

$$Var(Y) = Var(X) + Var(N) + 2\underbrace{Cov(X,N)}_{=0}$$

$$= a^2 + 1$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi(a^2+1)}} e^{-\frac{y^2}{2(a^2+1)}}$$

3. $$Cov(X,Y) = E((X - E(X))(Y - E(Y)))$$

$$= E(XY) = E(X(X+N))$$

$$= \underbrace{E(X^2)}_{=Var(X) = a^2} + \underbrace{E(XN)}_{=E(X)E(N) = 0 \cdot 0}$$ since X & N are indep.

$$= a^2$$

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{a^2}{\sqrt{a^2 \times (1+a^2)}} = \frac{a}{\sqrt{1+a^2}}$$

4. $a^* = \dfrac{Cov(X,Y)}{Var(Y)}$ $\qquad b^* = E(X) - a^* E(Y)$

$= \dfrac{a^2}{a^2+1}$ $\qquad = 0 - a^* \times 0 = 0.$

5. $E\left[(X - f_{MMSE,lin}(Y))^2\right]$

Method 1.

$$= E\left[\left(X - \frac{a^2}{a^2+1} Y\right)^2\right]$$

$$= E\left(X^2 - \frac{2a^2}{(a^2+1)} XY + \left(\frac{a^2}{a^2+1}\right)^2 Y^2\right)$$

$$= E(X^2) - 2a^2 \times \frac{1}{a^2+1} E(XY) + \left(\frac{a^2}{a^2+1}\right)^2 E(Y^2)$$

$$= a^2 - \frac{2a^2 \times a^2}{a^2+1} + \frac{a^4}{(a^2+1)^2} \times (a^2+1)$$

$$= \frac{a^2}{a^2+1}$$

Method 2:

$$E\left[(X - f_{MMSE,lin}(Y))^2\right]$$

$$= Var(X)(1-\rho^2) = a^2 \times \left(1 - \frac{a^2}{1+a^2}\right)$$

$$= \frac{a^2}{1+a^2}$$

Question 5: (15%) The number of customers that arrive at a convenience store during a time t (unit: minute) is a Poisson random variable with parameter αt. A person is flipping a fair coin every minute, and use T (min) to denote the waiting time before the first head faces up. Let N denote the number of customers arrive during the waiting time T. Assume that the customer arrivals are independent to the coin flipping.

1. (6%) Find $\mathsf{E}(N)$ the expected number of customers arrive during the waiting time.
2. (9%) Find the variance $\mathsf{Var}(N)$.

Hint 1: Use the following formula. Suppose X is a Poisson random variable with parameter λ, and Y is a geometric random variable with parameter p. Then

$$\begin{aligned}\mathsf{E}(X) &= \lambda\\ \mathsf{Var}(X) &= \lambda\\ \mathsf{E}(Y) &= \frac{1-p}{p}\\ \mathsf{Var}(Y) &= \frac{1-p}{p^2}.\end{aligned}$$

Hint 2: Use the equality that $\mathsf{E}(X) = \mathsf{E}(\mathsf{E}(X|Y))$.

1.
$$E(N) = E(E(N|T))$$ || since N given T is a Poisson
$$= E(\alpha T)$$
$$= \alpha E(T) = \alpha \times \frac{1-\frac{1}{2}}{\frac{1}{2}} = \alpha$$ || since T is a geometric random variable with parameter $p = \frac{1}{2}$

2.
$$\begin{aligned}E(N^2) &= E(E(N^2|T))\\ &= E(\mathrm{Var}(N|T) + (E(N|T))^2)\\ &= E(\alpha T + (\alpha T)^2)\\ &= E(\alpha T) + E(\alpha^2 T^2)\\ &= \alpha E(T) + \alpha^2(\mathrm{Var}(T) + (E(T))^2)\end{aligned}$$

since $E(T) = \frac{1-\frac{1}{2}}{\frac{1}{2}} = 1$ $\quad \mathrm{Var}(T) = \frac{1-\frac{1}{2}}{(\frac{1}{2})^2} = 2$

$$= \alpha + 3\alpha^2$$

(Conti)

$$\Rightarrow \mathrm{Var}(N) = E(N^2) - (E(N))^2$$

$$= \alpha + 3\alpha^2 - \alpha^2$$

$$= \alpha + 2\alpha^2$$

Question 6: (10%) A binary transmission channel introduces bit error with probability 0.10. Estimate the probability that there are 25 or fewer errors in 100 bit transmissions. Express your results in concrete numbers by looking up the following table: $\Phi(0) = 0.5$, $\Phi(1) = 0.8413$, $\Phi(2) = 0.9772$, $\Phi(3) = 0.99865$, $\Phi(4) = 1 - 3.16 \times 10^{-5}$, and $\Phi(5) = 1 - 2.87 \times 10^{-6}$, where $\Phi(x) = P(X \le x)$ is the cumulative distribution function of a standard Gaussian random variable X with mean 0 and variance 1.

let X be the total # of error bits.

$$\Rightarrow E(X) = np = 100 \times 0.1 = 10.$$

$$Var(X) = np(1-p) = 100 \times 0.1 \times 0.9 = 9$$

$$\Rightarrow P(X \le 25)$$

$$= P\left(\frac{X-10}{\sqrt{9}} \le \frac{25-10}{\sqrt{9}}\right)$$

$$= P\left(\frac{X-10}{\sqrt{9}} \le 5\right)$$

By the central limit theorem

$\frac{X-10}{\sqrt{9}}$ can be approximated by a Gaussian with mean 0 & variance 1

$$\Rightarrow P\left(\frac{X-10}{\sqrt{9}} \le 5\right)$$

$$\approx P(Z \le 5) = \Phi(5) = 1 - 2.87 \times 10^{-6}$$

#