Problem 1a (20pts) Let f be the pdf of a random variable X. The variance of X is denoted by σ and f is an even function (that is, $f(-x) = f(x)$). Define another random variable $Z = aX^2 + bX + c$.

1. For which value of a, b and c, X and Z are uncorrelated?

2. For which value of a, b and c, X and Z are independent?

1) Having $f(-x) = f(x)$ - distribution is symmetric. $\Rightarrow$

$\Rightarrow \quad E(Z) = E(Z^3) = E(Z^5) = \dots = 0$ as all odd central moments $= 0$

X and Z are uncorrelated $\iff$ $Cov(Z, X) = 0$

$Cov(Z,X) = Cov(aX^2 + bX + c, X) = a\,Cov(X^2, X) + b\,Cov(X,X) + Cov(c,X) = 0$

$a(\underbrace{E[X^2 \cdot X]}_{=0} - E[X^2]\underbrace{E[X]}_{=0}) + b(E[X \cdot X] - E[X]\underbrace{E[X]}_{=0}) + 0 = 0$

$\Rightarrow b\,E[X^2] = b\,\sigma^2 = 0 \Rightarrow b = 0$

Therefore, X and Z are uncorrelated, if $b = 0$ and a, c can be any

2) If X and Z are correlated then they are not independent.

But if X and Z are not correlated, then $b = 0 \Rightarrow$
$\Rightarrow Z = aX^2 + c$, so X and Z will be independent only if $aX = 0$
Therefore, $a = 0$ and c can be any

Problem 1.6

16.1

Statement :

If X and Z are uncorrelated, then X and Z are independent.

The statement is false.

Let X be a $N(0,1)$-distributed Random Variable, and $Y = X^2$ be another Random Variable.

Since $Y = X^2$ we know that X and Y are depentent

To counterproof the statement we must show that $\mathrm{corr}(X,Y) = 0$ which is equivalent of showing that $\mathrm{cov}(X,Y) = 0$

$$\mathrm{cov}(X,Y) = E[(X-\bar{x})(Y-\bar{y})]$$

$$= E[XY] - \underbrace{\bar{y}E[X]}_{0} - \underbrace{\bar{x}E[Y]}_{0} + \underbrace{\bar{x}\bar{y}}_{0} \quad | \text{ Since } E[X]=0$$

$$= E[XY] = E[XX^2] = E[X^3] = 0$$

| Since the density of X is even, all odd moments have an expectation of 0

$$\Rightarrow \mathrm{cov}(X,Y) = 0$$

$$\text{thus: } \mathrm{corr}(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{0}{\sigma_X \sigma_Y} = 0$$

16.2

Statement :

If X and Z are independent, then X and Z are uncorrelated.

The statement is true.

To prove that, we have to show, that if X and Y are two independent Random Variables, they are uncorrelated.

Note: For independent RV it holds that $P(X=x, Y=y) = P(X=x)\cdot P(Y=y)$

so the pdf $f_{XY} = \pi_X(x)\cdot \pi_Y(y)$

let X and Y be two Random Variables with mean $\bar{x}$ and $\bar{y}$ bar, then the covariance of X,Y is:

$$\mathrm{cov}(X,Y) = E[(X-\bar{x})\cdot(Y-\bar{y})]$$

$$= E[XY - X\bar{y} - \bar{x}Y + \bar{x}\bar{y}]$$

$$E[XY] - \bar{y}E[X] - \bar{x}E[Y] + \bar{x}\bar{y} \quad | \text{ Note } E[X]=\bar{x} \text{ and } E[Y]=\bar{y}$$

$$= E[XY] - \bar{y}\bar{x} \underbrace{- \bar{y}\bar{x} + \bar{y}\bar{x}}_{0}$$

$$= E[XY] - E[X]\cdot E[Y]$$

$$= E[X]E[Y] - E[X]\cdot E[Y] = 0$$

$$= E[XY] \int x\cdot y \cdot d\mu_{(X,Y)} \quad | \text{ since X and Y are independent, } d\mu_{(X,Y)} = d\mu_X \cdot d\mu_Y$$

$$= \int x\cdot d\mu_X \cdot y \cdot d\mu_Y = \left(\int x \cdot d\mu_X\right)\cdot\left(\int y \cdot dy\right)$$

$$= E[X]\cdot E[Y]$$

16.3

Statement

Suppose var(X) and var(Z) are finite. Then var(X + Z) = var(X) + var(Z) if and only if X and Z are independent.

Let X,Y be two Random Variables with mean $\bar{x},\bar{y}$

$$\text{The Variance of } X+Y = V(X+Y) = E[(X+Y)^2 - E[X+Y]^2]$$

$$= E[X^2] + 2E[XY] + E[Y^2] - (E[X] + E[Y])^2$$

$$= E[X^2] + E[Y^2] + \underbrace{2\cdot E[X\cdot Y] - 2E[X]E[Y]}_{2\cdot \mathrm{cov}(X,Y)} - E[X]^2 \cdot E[Y]^2$$

$$= E[X^2] + E[Y^2] - E[X]^2 - E[Y]^2 + 2\cdot \mathrm{cov}(X,Y)$$

$$= \underbrace{E[X^2] - E[X]^2}_{\mathrm{Var}(X)} + \underbrace{E[Y^2] - E[Y]^2}_{\mathrm{var}(Y)} + 2\cdot \mathrm{cov}(X,Y)$$

$$\Rightarrow \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{cov}(X,Y)$$

we can follow, that $\mathrm{Var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)$ only if X and Y are uncorrelated.

From 16.1 we learned, that beeing uncorrelated is not a sufficient property to inferre that two RV are independent.

Thus the statement is false since for all X and Y that are uncorrelated the statement is true, and not only for independence

Page No.

Date

Problem 2:

$$X = \begin{cases} 0, & \text{w.p } 1/2 \\ 1 & \text{w.p } 1/2 \end{cases} \qquad Y = \begin{cases} 0 & \text{w.p } 1/3 \\ 1 & \text{w.p. } 2/3 \end{cases}$$

$$Z = \begin{cases} 0 & \text{w.p. } 1 \\ 1 & \text{w.p. } 0. \end{cases}$$

→ 1) To characterize all the possible couplings in X & Y, we need to find a parametric joint distr^n bet^n X & Y

Let μ_X & μ_Y be the marginal prob. of X & Y

$\therefore \mu_X(0) = 1/2 \qquad \mu_Y(0) = 1/3$

$\mu_X(1) = 1/2 \qquad \mu_Y(1) = 2/3$

& we need to define
$\mu_{XY}(x, y)$ such that

$$\sum_{x \in R} \mu_{X,Y}(x, y) = \mu_Y(y) \quad \text{—(i)}$$

$$\sum_{y \in R} \mu_{X,Y}(x, y) = \mu_X(x) \quad \text{—(ii)}$$

$\because$ X & Y are discrete here

We can represent all the possible couplings in terms of a 2×2 matrix

This matrix also should satisfy the property of marginals i.e. (i) & (ii)

$$T = \begin{pmatrix} t_{00} & t_{01} \\ t_{10} & t_{11} \end{pmatrix}$$ where t_{ij} is the $\mu_{xy}(x=i, y=j)$

So for T to be valid joint probability distr^n from (i) & (ii)

$t_{00} + t_{10} = \mu_y(0) = 1/3$ ... a
$t_{01} + t_{11} = \mu_y(1) = 2/3$ ... b
$t_{00} + t_{01} = \mu_x(0) = 1/2$ ... c
$t_{10} + t_{11} = \mu_x(1) = 1/2$ ... d

We can convert eq^n to a single variable by simplifying further

$t_{00} = \frac{1}{3} - t_{10} = 1/2 - t_{01} \Rightarrow t_{01} = \frac{1}{2} - t_{00}$
$t_{11} = \frac{1}{2} - t_{10} = \frac{2}{3} - t_{01} \Rightarrow t_{11} = \frac{2}{3} - \frac{1}{2} + t_{00}$
$t_{01} = \frac{2}{3} - t_{11} = \cancel{\frac{2}{3}} - \cancel{\frac{2}{3}} - \frac{1}{2} + t_{00}$

$t_{10} = \frac{1}{2} - t_{11} = \frac{1}{3} - t_{00}$

∴ We have
$t_{00} = \alpha$
$t_{01} = \frac{1}{2} - \alpha$
$t_{10} = \frac{1}{3} - \alpha$
$t_{11} = \frac{1}{6} + \alpha$

Here α can be at max $(1/3)$ to satisfy marginal distr^n of $\mu_y(0)$ in a

Page No.
Date

∴ We can have various couplings from j.d. matrix

$$T = \begin{pmatrix} \alpha & 1/2 - \alpha \\ 1/3 - \alpha & 1/6 + \alpha \end{pmatrix} \quad \text{where } \alpha \in [0, 1/3]$$

where $t_{ij} = \mu_{X,Y}(x = i, y = j)$

→ 2) From the given Bernoulli distr^{n}s,

$\bar{X} = 1/2 \quad \bar{Y} = 2/3 \quad \& \quad \bar{Z} = 0$.

$\bar{X} \cdot \bar{Y} = 1/3$.

$$\text{Cov}[X,Y] = E[XY] - E[X] \cdot E[Y]$$
$$= E[XY] - 1/3$$

Now here $E[XY]$ can be expanded as

$$\cancel{\sum_{i=0, j=0}^{i=1, j=1} x \cdot y \cdot \mu_{X,Y}(x=i, y=j)}$$

$$\sum x \cdot y \cdot \mu_{X,Y}(x, y)$$

for every pair of X & Y.

Since our X & Y are sparse

$E[XY]$ can only be evaluated at $x = 1, y = 1$

From our joint distr^n matrix

$$t_{11} = \frac{1}{6} + \alpha = \mu_{X,Y}(1,1)$$

$$\therefore \text{Cov}[X,Y] = \cancel{\Sigma} \cdot 1 \cdot 1 \cdot \mu_{X,Y}(1,1) - 1/3$$
$$= \frac{1}{6} + \alpha - \frac{1}{3} = \alpha - \frac{1}{6}$$

Page No
Date

Since

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$$ ... from Bernoulli distrⁿ

$\sigma_X^2 = 1/4$

$\sigma_Y^2 = 2/9$

$\sigma_X \sigma_Y = \frac{1}{3\sqrt{2}}$

So corr ↓ or ↑es linearly with cov. as σ_X & σ_Y are const.

∴ Correlation is maximized if α is max. i.e. $\alpha = 1/3$ & its minimum when α is min. i.e. $\alpha = 0$.

For $\alpha = 1/3$ corr. is $\frac{1}{6} \cdot 3\sqrt{2} = \frac{1}{\sqrt{2}}$

$\alpha = 0$ corr. is $-\frac{1}{6} \cdot 3\sqrt{2} = -\frac{1}{\sqrt{2}}$.

→ 3) For X & Y to be non-correlated

$Cov(X,Y) = 0$

$\therefore \alpha - \frac{1}{6} = 0$

$\therefore \alpha = \frac{1}{6}$.

∴ Coupling is $\begin{pmatrix} 1/6 & 1/3 \\ 1/6 & 1/3 \end{pmatrix}$ which also follows the marginals.

→ 4) Since Z is a const r.v. always o/p is 0 w.p = 1

The T can be simply constructed by conserving sum of marginals.

$$T = \begin{pmatrix} \alpha & 0 \\ 1-\alpha & 0 \end{pmatrix}$$ where $t_{ij} \in T$ such that $t_{ij} = \mu_{X,Z}(X=i, Z=j)$

And to conserve marginals i.e. $\mu_X(0) = 1/2$

α has to be $1/2$ $\therefore t_{00} = 1/2 = \mu_X(0) \cdot \mu_Z(0)$

Also $E[X.Z]$ is again only valid for $x=1$ & $z=1$)

Since $\mu_{X,Z}(1\cdot 1) = 0$
& $E[X] = 1/2$ $E[Z] = 0$

$\therefore E[X \cdot Z] = E[X] \cdot E[Z]$ holds.

Even if we do this for every pair of x, z the same holds true. Hence x & z are independent.

EXERCISE 3

$\boxed{1}$ We have that

$$W(\pi_{X_1}, \pi_{X_2})^2 = \mathbb{E}_Z[\|X_1 - X_2\|^2] \quad \text{where}$$

μ_Z, Z's law, is such that

$$\mu_Z = \arg \operatorname*{Inf}_{\pi_Z \in \Pi(\pi_{X_1}, \pi_{X_2})} \mathbb{E}_Z[\|X_1 - X_2\|^2]$$

$\Longleftrightarrow$

$\mu_{X_1} = N(\bar{x}_1, \sigma_1^2)$
$\mu_{X_2} = N(\bar{x}_2, \sigma_2^2)$

$$W(\pi_{X_1}, \pi_{X_2})^2 = \operatorname*{Inf}_{X_1 \sim \mu_{X_1},\, X_2 \sim \mu_{X_2}} \mathbb{E}[\|X_1 - X_2\|^2] =$$

$X_1, X_2 \in \mathbb{R}$

$$= \operatorname*{Inf}_{X_1 \sim \mu_{X_1}, X_2 \sim \mu_{X_2}} \mathbb{E}[(X_1 - X_2)^2] =$$

$$= \operatorname*{Inf}_{X_1 \sim \mu_{X_1},\, X_2 \sim \mu_{X_2}} \mathbb{E}[X_1^2 + X_2^2 - 2X_1X_2] \overset{\text{linearity}}{=}$$

$$= \operatorname*{Inf}_{X_1 \sim \mu_{X_1},\, X_2 \sim \mu_{X_2}} \left(\mathbb{E}[X_1^2] + \mathbb{E}[X_2^2] - 2\mathbb{E}[X_1X_2]\right) =$$

$$= \inf_{\substack{X_1 \sim \mu_{X_1},\\ X_2 \sim \mu_{X_2}}} \left(\bar{x}_1^2 + \sigma_1^2 + \sigma_2^2 + \bar{x}_2^2 - 2\left(\mathrm{Cov}[X_1, X_2] + \bar{x}_1\bar{x}_2\right)\right) =$$

$$\mu_{X_1} = N(\bar{x}_1, \sigma_1^2), \quad \mu_{X_2} = N(\bar{x}_2, \sigma_2^2)$$

and $V[X_1] = \mathbb{E}[X_1^2] - \bar{x}_1^2$ Finally,

$$\mathrm{Cov}[X_1, X_2] = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]$$

$$= \inf_{\substack{X_1 \sim \mu_{X_1}\\ X_2 \sim \mu_{X_2}}} \left((\bar{x}_1 - \bar{x}_2)^2 + \sigma_1^2 + \sigma_2^2 - 2\,\mathrm{Cov}[X_1, X_2]\right) =$$

To get the Inf we need to maximize the covariance since the rest is fixed wrt μ_{X_1}, μ_{X_2} and ≥ 0

Given that $\rho = \dfrac{\mathrm{Cov}[X_1, X_2]}{\sqrt{\sigma_1^2 \sigma_2^2}}$ and $|\rho| \leq 1$

$$\max_{\substack{X_1 \sim \mu_{X_1},\\ X_2 \sim \mu_{X_2}}} \mathrm{Cov}[X_1, X_2] = \rho_{max} \sqrt{\sigma_1^2 \sigma_2^2} = \sqrt{\sigma_1^2 \sigma_2^2}$$

$$\Longrightarrow W(\pi_{x_1}, \pi_{x_2})^2 = (\bar{x}_1 - \bar{x}_2)^2 + \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2 = (\bar{x}_1 - \bar{x}_2)^2 + (\sigma_1 - \sigma_2)^2$$

$$\Updownarrow$$

$$W(\pi_{x_1}, \pi_{x_2}) = \sqrt{(\bar{x}_1 - \bar{x}_2)^2 + (\sigma_1 - \sigma_2)^2}$$

$\boxed{2}$ Let be $X_1 \sim N(\bar{X}_1, \sigma_1^2)$, $X_2 \sim N(\bar{X}_2, \sigma_2^2)$

$$D_{KL}(\pi_{X_1} \| \pi_{X_2}) = \int_{\mathbb{R}} \ln\left(\frac{\pi_{X_1}(x)}{\pi_{X_2}(x)}\right) \pi_{X_1}(x) dx =$$

$$= \int_{\mathbb{R}} (\ln(\pi_{X_1}(x)) - \ln(\pi_{X_2}(x))) \pi_{X_1}(x) dx \quad (*)$$

$$\ln(\pi_{X_1}(x)) - \ln(\pi_{X_2}(x)) \overset{\text{Gaussian densities}}{=} \ln\left(\frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\bar{X}_1)^2}{2\sigma_1^2}}\right) -$$

$$- \ln\left(\frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(x-\bar{X}_2)^2}{2\sigma_2^2}}\right) = -\frac{1}{2}\ln(2\pi\sigma_1^2) - \frac{(x-\bar{X}_1)^2}{2\sigma_1^2} +$$

$$+ \frac{1}{2}\ln(2\pi\sigma_2^2) + \frac{(x-\bar{X}_2)^2}{2\sigma_2^2} = -\frac{1}{2}\ln(2\pi\sigma_1^2) - \frac{1}{2\sigma_1^2}(x^2 - 2x\bar{X}_1 + \bar{X}_1^2) +$$

$$+ (x^2 - 2\bar{X}_2 x + \bar{X}_2^2)\frac{1}{2\sigma_2^2} + \ln(2\pi\sigma_2^2)\frac{1}{2} =$$

$$= \frac{1}{2}\ln\left(\frac{\sigma_2^2}{\sigma_1^2}\right) - \frac{1}{2\sigma_1^2}x^2 + \frac{x\bar{X}_1}{\sigma_1^2} - \frac{\bar{X}_1^2}{2\sigma_1^2} + \frac{x^2}{2\sigma_2^2} + \frac{\bar{X}_2^2}{2\sigma_2^2} - \frac{x\bar{X}_2}{\sigma_2^2} =$$

$$= \ln\left(\frac{\sigma_2}{\sigma_1}\right) + x^2\left(\frac{1}{2\sigma_2^2} - \frac{1}{2\sigma_1^2}\right) + x\left(\frac{\bar{X}_1}{\sigma_1^2} - \frac{\bar{X}_2}{\sigma_2^2}\right) + \frac{\bar{X}_2^2}{2\sigma_2^2} - \frac{\bar{X}_1^2}{2\sigma_1^2}$$

$\Longrightarrow$

$$(*) = \int_{\mathbb{R}}\left[\ln\left(\frac{\sigma_2}{\sigma_1}\right) + x^2\frac{(\sigma_1^2 - \sigma_2^2)}{2\sigma_2^2\sigma_1^2} + x\left(\frac{\sigma_2^2\bar{X}_1 - \sigma_1^2\bar{X}_2}{\sigma_1^2\sigma_2^2}\right) + \frac{\bar{X}_2^2\sigma_1^2 - \bar{X}_1^2\sigma_2^2}{2\sigma_1^2\sigma_2^2}\right]$$

$$\pi_{X_1}(x) dx \underset{\downarrow}{=}$$

$\int_{\mathbb{R}} \pi_{X_1}(x) dx = 1$ (density) + integrating wrt x

$$= \ln\left(\frac{\sigma_2}{\sigma_1}\right) + E[X_1^2]\left(\frac{\sigma_1^2-\sigma_2^2}{2\sigma_2^2\sigma_1^2}\right) + E[X_1]\left(\frac{\sigma_2^2\bar{x}_1 - \sigma_1^2\bar{x}_2}{\sigma_1^2\sigma_2^2}\right) +$$

$$+ \frac{\bar{x}_2^2\sigma_1^2 - \bar{x}_1^2\sigma_2^2}{2\sigma_1^2\sigma_2^2} = \ln\left(\frac{\sigma_2}{\sigma_1}\right) + (\sigma_1^2+\bar{x}_1^2)\left(\frac{\sigma_1^2-\sigma_2^2}{2\sigma_1^2\sigma_2^2}\right) +$$

$$\sigma_1^2 = E[X_1^2] - \bar{x}_1^2$$

$$+ \frac{\bar{x}_1^2}{\sigma_1^2} - \frac{\bar{x}_1\bar{x}_2}{\sigma_2^2} + \frac{\bar{x}_2^2}{2\sigma_2^2} - \frac{\bar{x}_1^2}{2\sigma_1^2} = \ln\left(\frac{\sigma_2}{\sigma_1}\right) + \frac{(\sigma_1^2-\sigma_2^2)}{2\sigma_2^2} + \frac{(\bar{x}_1-\bar{x}_2)^2}{2\sigma_2^2} =$$

$$= \ln\left(\frac{\sigma_2}{\sigma_1}\right) + \frac{1}{2\sigma_2^2}\left[\sigma_1^2 - \sigma_2^2 + (\bar{x}_1-\bar{x}_2)^2\right] =$$

$$= \frac{1}{2}\left[\ln\left(\frac{\sigma_2^2}{\sigma_1^2}\right) + \frac{\sigma_1^2}{\sigma_2^2} - 1 + \frac{(\bar{x}_1-\bar{x}_2)^2}{\sigma_2^2}\right]$$