Problem 1. Random variables $X, Y, Z \sim \mathcal{N}(0,1)$ are independent. Are there constants a, b, c, $d \in \mathbb{R}$ such that the random vectors $(X+Y+1, X+aY+bZ)^T$ and $(X-Y, X+cY+dZ+2)^T$ are independent? If so, determine these constants.

Problem 2. The random vector (X, Y) has the normal disribution with the parameters $EX = 0$, $EY = 1$, $Var\ X = 4$, $Var\ Y = 10$ and $Cov(X, Y) = -2$. Calculate $Var(XY)$ and $Cov(X^2, Y^2)$.

Problem 3. Let $(X_1, X_2, \ldots, X_{16})$ and $(Y_1, Y_2, \ldots, Y_{25})$ be two independent simple samples from the same normal distribution with the unknown expectation μ and the known variance σ^2 Separately, based on random samples $X_1, X_2, \ldots, X_{16}$ and $Y_1, Y_2, \ldots, Y_{25}$, two confidence intervals were built for the expected value μ, each at the confidence level of 0.8. Calculate the probability that one of these intervals is a subinterval of the other one.

Problem 4. We have two independent simple samples $X_1, \ldots, X_m \sim \mathcal{N}(\mu, \sigma^2)$ and $Y_1, \ldots, Y_n \sim \mathcal{N}(\mu, 2\sigma^2)$. Let $Z = \frac{2\sum_{i=1}^m X_i + \sum_{i=1}^n Y_i}{2m+n}$ (i.e. $Z = \frac{2m\overline{X}+n\overline{Y}}{2m+n}$). Find $a \in \mathbb{R}$ such that $S_a^2 = a\left(2\sum_{i=1}^m (X_i - Z)^2 + \sum_{i=1}^n (Y_i - Z)^2\right)$ is an unbiased estimator of the parameter σ^2.

Problem 5. Let X_1, X_2, X_3, X_4 be independent random variables such that $X_i \sim \mathcal{N}(m, im^2)$ for $i = 1, 2, 3, 4$. We consider estimators of the unknown parameter $m \in \mathbb{R}$ of the form $\widehat{m} = a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4$. Find $a_1, a_2, a_3, a_4 \in \mathbb{R}$ ensuring that the mean squared error $E(\hat{\mu} - \mu)^2$ of the estimator $\hat{\mu}$ is the smallest possible.

Problem 1. Random variables $X, Y, Z \sim \mathcal{N}(0,1)$ are independent. Are there constants a, b, c, $d \in \mathbb{R}$ such that the random vectors $(X+Y+1, X+aY+bZ)^T$ and $(X-Y, X+cY+dZ+2)^T$ are independent? If so, determine these constants.

$X, Y, Z \sim \mathcal{N}(0,1)$ ind. $\iff K = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \sim \mathcal{N}\left(\mu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)$

Let $K^1 = \begin{pmatrix} X+Y+1 \\ X+aY+bZ \end{pmatrix}$, $K^2 = \begin{pmatrix} X-Y \\ X+cY+dZ+2 \end{pmatrix}$

K^1, K^2 ind. $\iff \operatorname{cov}(K_i^j, K_k^l) = 0$, $i, j, k, l \in \{1,2\}$, $j \neq l$

$\operatorname{cov}(K_1^1, K_1^2) = \operatorname{cov}(X+Y+1, X-Y) = \operatorname{Var}[X] - \operatorname{Var}[Y] = 1 - 1 = 0$

$\operatorname{cov}(K_1^1, K_2^2) = \operatorname{cov}(X+Y+1, X+cY+dZ+2) = \operatorname{Var}[X] + c\operatorname{Var}[Y] = 1 + c = 0 \Rightarrow c = -1$

$\operatorname{cov}(K_2^1, K_1^2) = \operatorname{cov}(X+aY+bZ, X-Y) = \operatorname{Var}[X] - a\operatorname{Var}[Y] = 1 - a = 0 \Rightarrow a = 1$

$\operatorname{cov}(K_2^1, K_2^2) = \operatorname{cov}(X+aY+bZ, X+cY+dZ+2) = \operatorname{Var}[X] + ac\operatorname{Var}[Y] + bd\operatorname{Var}[Z] =$
$1 - 1 + bd = bd = 0 \Rightarrow b = 0, d \in \mathbb{R} \lor$
$d = 0, b \in \mathbb{R}$

Sol: $\begin{cases} a = 1 \\ b = 0 \\ c = -1 \\ d \in \mathbb{R} \end{cases} \lor \begin{cases} a = 1 \\ b \in \mathbb{R} \\ c = -1 \\ d = 0 \end{cases}$

Problem 2. The random vector (X, Y) has the normal disribution with the parameters $EX = 0$, $EY = 1$, $Var\ X = 4$, $Var\ Y = 10$ and $Cov(X, Y) = -2$. Calculate $Var(XY)$ and $Cov(X^2, Y^2)$.

$$Z\begin{pmatrix} X \\ Y \end{pmatrix} \rightsquigarrow N\left(\mu=\begin{pmatrix} 0 \\ 1 \end{pmatrix}, A=\begin{pmatrix} 4 & -2 \\ -2 & 10 \end{pmatrix}\right)$$

- $Var[XY] = E[(XY)^2]^{**} - E[XY]^{*2} = 52 - 4 = 48$

$* E[XY] = COV(X, Y) + E[X]E[Y] = -2 + 0 \cdot 1 = -2$

$**$ For $E[X^2Y^2]$ let's standardize :

$Z \rightsquigarrow N(\mu, A) \Leftrightarrow Z = LK + \mu,\ K = \begin{pmatrix} U \\ V \end{pmatrix} \rightsquigarrow N(0, I),\ LL^T = A$

$$\text{Let } LL^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -2 & 10 \end{pmatrix} \Rightarrow \begin{cases} a=2 \\ b=0 \\ c=-1 \\ d=3 \end{cases}$$

$$Z = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2U \\ -U+3V+1 \end{pmatrix}$$

$$E[X^2Y^2] = E[4U^2(-U+3V+1)^2] =$$

$$4[E[U^4] - 6E[U^3]E[V] - 2E[U^3] + 9E[U^2]E[V^2] + 6E[U^2]E[V] + E[U^2]]$$

$$= 4[3 + 9 + 1] = 52$$

$***$ $U, V \rightsquigarrow N(0,1)$ ind $\Leftrightarrow \begin{pmatrix} U \\ V \end{pmatrix} \rightsquigarrow N(0, I)$

$$E[U^4] = \frac{4!}{\left(\frac{4}{2}\right)!\, 2^{4/2}} = \frac{4 \cdot 3 \cdot 2}{2 \cdot 4} = 3$$

$$E[U^2V^2] = E[U^2]E[V^2] = \left(\frac{2!}{\left(\frac{2}{2}\right)!\, 2^{2/2}}\right)^2 = 1$$

$$E[U^3V] = E[U^3]E[V] = 0 \cdot 0 = 0$$

- $COV(X^2, Y^2) = E[X^2Y^2] - \underbrace{E[X^2]E[Y^2]}_{****} = 52 - 4 \cdot 11 = 18$

$****$

$var[X] = E[X^2] - E[X]^2 \Rightarrow E[X^2] = 4 + 0^2 = 4$

$Var[Y] = E[Y^2] - E[Y]^2 \Rightarrow E[Y^2] = 10 + 1^2 = 11$

Problem 3. Let $(X_1, X_2, \dots, X_{16})$ and $(Y_1, Y_2, \dots, Y_{25})$ be two independent simple samples from the same normal distribution with the unknown expectation μ and the known variance σ^2 Separately, based on random samples $X_1, X_2, \dots, X_{16}$ and $Y_1, Y_2, \dots, Y_{25}$, two confidence intervals were built for the expected value μ, each at the confidence level of 0.8. Calculate the probability that one of these intervals is a subinterval of the other one.

Problem 4. We have two independent simple samples $X_1, \ldots, X_m \sim \mathcal{N}(\mu, \sigma^2)$ and $Y_1, \ldots, Y_n \sim \mathcal{N}(\mu, 2\sigma^2)$. Let $Z = \frac{2\sum_{i=1}^m X_i + \sum_{i=1}^n Y_i}{2m+n}$ (i.e. $Z = \frac{2m\overline{X} + n\overline{Y}}{2m+n}$). Find $a \in \mathbb{R}$ such that $S_a^2 = a\left(2\sum_{i=1}^m (X_i - Z)^2 + \sum_{i=1}^n (Y_i - Z)^2\right)$ is an unbiased estimator of the parameter σ^2.

It's clear Z also follows a normal distribution.

$$E[Z] = \frac{1}{2m+n}\left(2m E[\overline{X}] + n E[\overline{Y}]\right) = \frac{2m\mu + n\mu}{2m+n} = \mu$$

$\overline{X}, \overline{Y}$ ind.

$$\mathrm{var}[Z] = \frac{1}{(2m+n)^2}\left(4m^2 \mathrm{var}[\overline{X}] + n^2 \mathrm{var}[\overline{Y}]\right) = \frac{4m^2 \frac{\sigma^2}{m} + n^2 \frac{2\sigma^2}{n}}{(2m+n)^2} = \frac{2\sigma^2(2m+n)}{(2m+n)^2} = \frac{2\sigma^2}{2m+n}$$

then, $Z \sim N\left(\mu, \frac{2\sigma^2}{2m+n}\right) \Longrightarrow X_i - Z$, $Y_i - Z$ also follow a normal dist.

S_a^2 unbiased estimator of $\sigma^2 \Longleftrightarrow E[S_a^2] = \sigma^2 \Longleftrightarrow$

$$a\Big[2\sum_{i=1}^m \underbrace{E[(X_i - Z)^2]}_{1)} + \sum_{i=1}^n \underbrace{E[(Y_i - Z)^2]}_{2)}\Big] =$$

$$a\left[2m\left(\sigma^2 - \frac{2\sigma^2}{2m+n}\right) + n\left(2\sigma^2 - \frac{2\sigma^2}{2m+n}\right)\right] =$$

$$a\left[2m\sigma^2 - \frac{4m\sigma^2}{2m+n} + 2n\sigma^2 - \frac{2n\sigma^2}{2m+n}\right] = 2a\sigma^2\left[m+n - \frac{2m+n}{2m+n}\right] =$$

$$2a\sigma^2[m+n-1] = \sigma^2 \Longleftrightarrow a = \frac{1}{2(m+n-1)}$$

1) $\mathrm{Var}[X_i - Z] = E[(X_i - Z)^2] - E[X_i - Z]^2 \Longrightarrow$

$$E[(X_i - Z)^2] = \mathrm{Var}[X_i - Z] + E[X_i - Z]^2 \overset{*}{=} \frac{\sigma^2(2m+n-2)}{2m+n}$$

$*E[X_i - Z] = E[X_i] - E[Z] = \mu - \mu = 0$

$*\ \mathrm{Var}[X_i - Z] = \mathrm{Var}[X_i] + \mathrm{Var}[Z] - 2\underset{**}{\mathrm{Cov}(X_i, Z)}$

$$= \sigma^2 + \frac{2\sigma^2}{2m+n} - \frac{4\sigma^2}{2m+n} = \sigma^2 - \frac{2\sigma^2}{2m+n}$$

$$**\ \mathrm{Cov}(X_i, Z) = \frac{1}{2m+n}\left(2\sum_{j=1}^{m}\mathrm{Cov}(X_i, X_j) + \sum_{i=1}^{n}\cancelto{0}{\mathrm{Cov}(X_i, Y_i)}\right) = \frac{2\,\mathrm{Var}[X_i]}{2m+n} = \frac{2\sigma^2}{2m+n}$$

2) $\mathrm{Var}[Y_i - Z] = E[(Y_i - Z)^2] - E[Y_i - Z]^2 \Longrightarrow$

$$E[(Y_i - Z)^2] = \mathrm{Var}[Y_i - Z] + E[Y_i - Z]^2 \overset{*}{=} \frac{2\sigma^2(2m+n-1)}{2m+n}$$

$*E[Y_i - Z] = E[Y_i] - E[Z] = \mu - \mu = 0$

$*\ \mathrm{Var}[Y_i - Z] = \mathrm{Var}[Y_i] + \mathrm{Var}[Z] - 2\underset{**}{\mathrm{Cov}(Y_i, Z)}$

$$= 2\sigma^2 + \frac{2\sigma^2}{2m+n} - \frac{4\sigma^2}{2m+n} = 2\sigma^2 - \frac{2\sigma^2}{2m+n}$$

$$**\ \mathrm{Cov}(Y_i, Z) = \frac{1}{2m+n}\left(2\sum_{i=1}^{m}\cancelto{0}{\mathrm{Cov}(Y_i, X_i)} + \sum_{j=1}^{n}\mathrm{Cov}(Y_i, Y_j)\right) = \frac{\mathrm{Var}[Y_i]}{2m+n} = \frac{2\sigma^2}{2m+n}$$

Problem 5. Let X_1, X_2, X_3, X_4 be independent random variables such that $X_i \sim \mathcal{N}(m, im^2)$ for $i = 1, 2, 3, 4$. We consider estimators of the unknown parameter $m \in \mathbb{R}$ of the form $\widehat{m} = a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4$. Find $a_1, a_2, a_3, a_4 \in \mathbb{R}$ ensuring that the mean squared error $E(\hat{\mu} - \mu)^2$ of the estimator $\hat{\mu}$ is the smallest possible.

Let $n=4$.

$$E[(\hat{m}-m)^2] = \mathrm{Var}[\hat{m}-m] + E[\hat{m}-m]^2 \overset{*}{=}$$

$$m^2\sum_{i=1}^n i a_i^2 + m^2\left(\sum_{i=1}^n a_i - 1\right)^2 = m^2\left(\sum_{i=1}^n i a_i^2 + \left(\sum_{i=1}^n a_i - 1\right)^2\right) = m^2 f(a_1 \dots a_n)$$

$$* \quad E[\hat{m}-m] = E[\hat{m}] - E[m] = \sum_{i=1}^n a_i E[X_i] - m = m\left(\sum_{i=1}^n a_i - 1\right)$$

$$\mathrm{Var}[\hat{m}-m] = \mathrm{Var}[\hat{m}] = \sum_{i=1}^n a_i^2 \mathrm{Var}[X_i] = \sum_{i=1}^n a_i^2\, i m^2 = m^2\sum_{i=1}^n i a_i^2$$

Let's minimize f. Using C-S inequality: $\left(\sum_{i=1}^n b_i c_i\right)^2 \le \sum_{i=1}^n b_i^2 \sum_{i=1}^n c_i^2$

Take $b_i = \sqrt{i}\, a_i$, $c_i = \frac{1}{\sqrt{i}}$

$$\left(\sum_{i=1}^n a_i\right)^2 \le \sum_{i=1}^n i a_i^2 \sum_{i=1}^n \frac{1}{i} = \frac{25}{12}\sum_{i=1}^n i a_i^2 \iff$$

$$\frac{12}{25}\left(\sum_{i=1}^n a_i\right)^2 \le \sum_{i=1}^n i a_i^2$$

$$f(a_1 \dots a_n) \ge \frac{12}{25}\left(\sum_{i=1}^n a_i\right)^2 + \left(\sum_{i=1}^n a_i - 1\right)^2$$

Let $x = \sum_{i=1}^n a_i$. Let's minimize $g(x) = \frac{12}{25}x^2 + (x-1)^2$

$$g'(x) = \frac{24}{25}x + 2(x-1) = \frac{74}{25}x - 2 = 0 \Rightarrow x = \sum_{i=1}^n a_i = \frac{25}{37}$$

The equality in the C-S inequality is reached when $b_i = k c_i \ \forall i = 1 \dots n$, $k \in \mathbb{R}$

$$b_i = \sqrt{i}\, a_i = \frac{k}{\sqrt{i}} = c_i \iff a_i = \frac{k}{i} \Rightarrow$$

$$\sum_{i=1}^n a_i = \frac{25}{37} = k\sum_{i=1}^n \frac{1}{i} = k \cdot \frac{25}{12} \Rightarrow k = \frac{12}{37} \Rightarrow a_i = \frac{12}{37 i}$$

$$\hat{m} = \frac{12}{37}\sum_{i=1}^4 \frac{X_i}{i}$$