Problem 1. The random vector $(X, Y, Z)^T$ has distribution $\mathcal{N}\left(\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}2 & -2 & 0\\-2 & 3 & 0\\0 & 0 & 4\end{pmatrix}\right)$. Are there constants a, b, c such that the random variable $X + Y + aZ + 3$ is independent of the random vector $(X + 2Z + 4,\ X - bY + Z + c)^T$? If so, determine these constants.

Problem 2. The random vector (X, Y) has the normal disribution with the parameters $EX = 0 = EY = 0$, $D^2X = D^2Y = 1$ and $Cov(X, Y) = \rho \in [-1, 1]$. Calculate $D^2(XY)$ 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 the intervals constructed in this way turn out to be disjoint.

Problem 4. We have two independent observations X_1 and X_2 from the normal distribution. One of them has the distribution with the parameters (μ, σ^2), and the other one with the parameters $(2\mu, 2\sigma^2)$. Unfortunately, we have lost information about which observation comes from which distribution. Parameters are unknown. In this situation we consider the following estimator $\widehat{\sigma}^2$ of the parameter σ^2:
$\widehat{\sigma}^2 = a(X_1 - X_2)^2 + b(X_1 + X_2)^2$, where a and b are real numbers. Find a and b if the estimator $\widehat{\sigma}^2$ is unbiased.

Problem 5. We assume that $(X_1, X_2, \ldots, X_n)$ is a simple sample from the normal distribution $\mathcal{N}(\mu, \gamma^2\mu^2)$, where $\gamma^2 > 0$ is known and $\mu \in \mathbb{R}$ is unknown. Find an estimator of μ of the form $\hat{\mu} = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ that would have the smallest mean squared error $E(\hat{\mu} - \mu)^2$ among such estimators.

Problem 1. The random vector $(X,Y,Z)^T$ has distribution $\mathcal{N}\left(\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}2 & -2 & 0\\ -2 & 3 & 0\\ 0 & 0 & 4\end{pmatrix}\right)$. Are there constants a, b, c such that the random variable $X+Y+aZ+3$ is independent of the random vector $(X+2Z+4,\ X-bY+Z+c)^T$? If so, determine these constants.

Let $K_1 = X+Y+aZ+3 \rightsquigarrow \mathcal{N}_1$

$$K = \begin{pmatrix} K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} X+2Z+4 \\ X-bY+Z+C \end{pmatrix}$$

$$\begin{pmatrix} K_1 \\ K_2 \\ K_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & a \\ 1 & 0 & 2 \\ 1 & -b & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \\ C \end{pmatrix} \rightsquigarrow \mathcal{N}$$

K_1 ind. from $K \Leftrightarrow \operatorname{cov}(K_1, K_2) = \operatorname{cov}(K_1, K_3) = 0$

$$\operatorname{cov}(K_1, K_2) = \operatorname{cov}(X+Y+aZ+3,\ X+2Z+4) =$$

$$\operatorname{cov}(X,X) + 2\cancelto{0}{\operatorname{cov}(X,Z)} + \operatorname{cov}(Y,X) + 2\cancelto{0}{\operatorname{cov}(Y,Z)} + a\cancelto{0}{\operatorname{cov}(Z,X)} + 2a\operatorname{cov}(Z,Z)$$

$$= \operatorname{Var}[X] + \operatorname{cov}(X,Y) + 2a\operatorname{Var}[Z] = 2 - 2 + 8a = 0 \Rightarrow a = 0$$

$$\operatorname{cov}(K_1, K_3) = \operatorname{cov}(X+Y+aZ+3,\ X-bY+Z+C) =$$

$$\operatorname{Var}[X] + (1-b)\operatorname{cov}(X,Y) - b\operatorname{var}[Y] + a\operatorname{var}[Z] = 2 - 2(1-b) - 3b =$$

$$-b = 0 \Rightarrow b = 0$$

Sol: K_1 ind. from $K \Leftrightarrow a = b = 0,\ c \in \mathbb{R}$

Problem 2. The random vector (X, Y) has the normal disribution with the parameters $EX = 0 = EY = 0$, $D^2X = D^2Y = 1$ and $Cov(X, Y) = \rho \in [-1, 1]$. Calculate $D^2(XY)$ and $Cov(X^2, Y^2)$.

Let $Z = \begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$

We know $Var[X] = E[X^2] - E[X]^2$, $COV(X, Y) = E[XY] - E[X]E[Y]$

• $Var[XY] = E[\overset{**}{X^2Y^2}] - E[\overset{*}{XY}]^2 = -\rho^2 + 3\rho + 1 - \rho^2 = -2\rho^2 + 3\rho + 1$

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

$**$ For $E[X^2Y^2]$, let's standarize.

We know $Z \sim N(\mu, A) \iff Z = LK + \mu$, where $A = LL^T$, $K \sim N(0, I)$

Let $K = \begin{pmatrix} U \\ V \end{pmatrix} \sim N(0, 1)$, $L = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$LL^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \Rightarrow$

We can fix $a = 1, b = 0 \Rightarrow \begin{cases} c = \rho \\ c^2 + d^2 = 1 \Rightarrow d = \sqrt{1-\rho^2} \end{cases}$

$Z = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} U \\ \rho U + \sqrt{1-\rho^2}\, V \end{pmatrix}$

$E[X^2Y^2] = E[U^2(\rho U + \sqrt{1-\rho^2}\, V)^2] = E[\rho U^4 + (1-\rho^2)U^2V^2 + 2\rho\sqrt{1-\rho^2}\, U^3V] =$
$\rho E[U^4] + (1-\rho^2)E[U^2V^2] + 2\rho\sqrt{1-\rho^2}\, E[U^3V] \overset{***}{=} 3\rho + 1 - \rho^2 = -\rho^2 + 3\rho + 1$

$***$ We know $U, V \sim N(0,1)$ ind. $\iff \begin{pmatrix} U \\ V \end{pmatrix} \sim N(0, I)$

$E[U^4] = \dfrac{4!}{\frac{4}{2}! \cdot 2^{4/2}} = \dfrac{4 \cdot 3 \cdot 2}{2 \cdot 4} = 3$

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

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

• $COV(X^2, Y^2) = E[X^2Y^2] - \overset{****}{E[X^2]E[Y^2]} = -\rho^2 + 3\rho + 1 - 1 = -\rho^2 + 3\rho$

$****$ $E[X^2] = E[(X-0)^2] = E[(X-\mu_1)^2] = Var[X] = 1$
$E[Y^2] = E[(Y-0)^2] = E[(Y-\mu_2)^2] = Var[Y] = 1$

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 the intervals constructed in this way turn out to be disjoint.

We know $X_1 \dots X_{16}, Y_1 \dots Y_{25} \sim N(\mu, \sigma^2)$

Confidence level $= 0'8 = 1-\alpha \Rightarrow \alpha = 0'2$

We can take the simmetric confidence intervals for

$$\hat{\mu}_X = \frac{1}{16}\sum_{i=1}^{16} X_i = \bar{X} \sim N\left(\mu, \frac{\sigma^2}{16}\right)$$

$$\hat{\mu}_Y = \frac{1}{25}\sum_{i=1}^{25} Y_i = \bar{Y} \sim N\left(\mu, \frac{\sigma^2}{25}\right)$$

$$I_1 = \left]\bar{X} - \frac{\sigma}{4}u_{1-0'1}, \bar{X} + \frac{\sigma}{4}u_{1-0'1}\right[= \left]\bar{X} - \frac{\sigma}{4}u_{0'9}, \bar{X} + \frac{\sigma}{4}u_{0'9}\right[$$

$$I_2 = \left]\bar{Y} - \frac{\sigma}{5}u_{1-0'1}, \bar{Y} + \frac{\sigma}{5}u_{1-0'1}\right[= \left]\bar{Y} - \frac{\sigma}{5}u_{0'9}, \bar{Y} + \frac{\sigma}{5}u_{0'9}\right[$$

$$I_1 \cap I_2 = \emptyset \iff \bar{X} + \frac{\sigma}{4}u_{0'9} < \bar{Y} - \frac{\sigma}{5}u_{0'9} \ \vee\ \bar{Y} + \frac{\sigma}{5}u_{0'9} < \bar{X} - \frac{\sigma}{4}u_{0'9}$$

$$\iff \underset{1}{\bar{X} - \bar{Y} < -\sigma u_{0'9}\cdot\frac{9}{20}} \ \vee\ \underset{2}{\bar{X} - \bar{Y} > \sigma u_{0'9}\cdot\frac{9}{20}}$$

$$P = P[1 \cup 2] = 1 - P[\bar{1} \cap \bar{2}] =$$

$$1 - P\left[\bar{X} - \bar{Y} \geq -\frac{9}{20}\sigma u_{0'9} \wedge \bar{X} - \bar{Y} \leq \frac{9}{20}\sigma u_{0'9}\right] =$$

$$1 - P\left[-\frac{9}{20}\sigma u_{0'9} \leq \bar{X} - \bar{Y} \leq \frac{9}{20}\sigma u_{0'9}\right] \overset{*}{=} 1 - P\left[\frac{-9\cancel{\sigma}u_{0'9}}{\cancel{20}\frac{\sqrt{41}}{\cancel{20}}\cancel{\sigma}} \leq \frac{\bar{X} - \bar{Y}}{\frac{\sqrt{41}}{20}\sigma} \leq \frac{9u_{0'9}}{\sqrt{41}}\right]$$

$$= 1 - P\left[\left|\frac{\bar{X} - \bar{Y}}{\frac{\sqrt{41}}{20}\sigma}\right| \leq \frac{9u_{0'9}}{\sqrt{41}}\right] = 1 - 2P\left[\frac{\bar{X} - \bar{Y}}{\frac{\sqrt{41}}{20}\sigma} \leq \frac{9u_{0'9}}{\sqrt{41}}\right] + 1 \approx \sim$$

$u_{0'9} = 1'29$

Search in table

* Let's see the distribution of $\bar{X} - \bar{Y}$

$$E[\bar{X} - \bar{Y}] = E[\bar{X}] - E[\bar{Y}] = \mu - \mu = 0$$

$$Var[\bar{X} - \bar{Y}] = Var[\bar{X}] + Var[\bar{Y}] - 2\cancel{cov(\bar{X}, \bar{Y})}^{0} = \frac{\sigma^2}{16} + \frac{\sigma^2}{25} = 0'1025\sigma^2$$

$$\bar{X} - \bar{Y} \sim N(0, 0'1025\sigma^2) \Rightarrow \frac{\bar{X} - \bar{Y}}{\frac{\sqrt{41}}{20}\sigma} \sim N(0,1)$$

Problem 4. We have two independent observations X_1 and X_2 from the normal distribution. One of them has the distribution with the parameters (μ, σ^2), and the other one with the parameters $(2\mu, 2\sigma^2)$. Unfortunately, we have lost information about which observation comes from which distribution. Parameters are unknown. In this situation we consider the following estimator $\widehat{\sigma}^2$ of the parameter σ^2:
$\widehat{\sigma}^2 = a(X_1 - X_2)^2 + b(X_1 + X_2)^2$, where a and b are real numbers. Find a and b if the estimator $\widehat{\sigma}^2$ is unbiased.

Problem 5. We assume that $(X_1, X_2, \dots, X_n)$ is a simple sample from the normal distribution $\mathcal{N}(\mu, \gamma^2\mu^2)$, where $\gamma^2 > 0$ is known and $\mu \in \mathbb{R}$ is unknown. Find an estimator of μ of the form $\hat{\mu} = a_1X_1 + a_2X_2 + \cdots + a_nX_n$ that would have the smallest mean squared error $E(\hat{\mu} - \mu)^2$ among such estimators.

$$E[(\hat{\mu}-\mu)^2] = Var[\hat{\mu}-\mu] + E[\hat{\mu}-\mu]^2 \overset{*}{=} \gamma^2\mu^2\sum_{i=1}^n a_i^2 + \mu^2\left(\sum_{i=1}^n a_i - 1\right)^2 =$$

$$\mu^2\left[\gamma^2\sum_{i=1}^n a_i^2 + \left(\sum_{i=1}^n a_i - 1\right)^2\right] = \mu^2 f(a_1 \dots a_n) \rightsquigarrow \min$$

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

$$Var[\hat{\mu}-\mu] = Var[\hat{\mu}] = \sum_{i=1}^n a_i^2 Var[X_i] = \sum_{i=1}^n a_i^2\gamma^2\mu^2 = \gamma^2\mu^2\sum_{i=1}^n a_i^2$$

By Cauchy-Schwartz inequality, $\sum_{i=1}^n b_i c_i \le \sqrt{\sum_{i=1}^n b_i^2}\sqrt{\sum_{i=1}^n c_i^2}$

Take $b_i = a_i\sqrt{\gamma^2} = a_i\gamma$, $c_i = \frac{1}{\sqrt{\gamma^2}} = \frac{1}{\gamma}$

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

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

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

Let $\lambda = \sum_{i=1}^n a_i$, $g(\lambda) = \frac{\lambda^2}{n/\gamma^2} + (\lambda - 1)^2$

$$g'(\lambda) = \frac{2\gamma^2}{n}\lambda + 2(\lambda - 1) = \frac{2(\gamma^2+n)}{n}\lambda - 2 = 0 \iff$$

$\lambda = \frac{n}{(\gamma^2+n)} \Rightarrow$ the minimum is reached for $\lambda = \sum_{i=1}^n a_i = \frac{n}{\gamma^2+n}$

the C-S inequality reaches equality $\Longleftrightarrow b_i = k c_i \quad \forall i = 1..n$

$\Longleftrightarrow a_i \gamma = k \cdot \frac{1}{\gamma} \Longleftrightarrow a_i = \frac{k}{\gamma^2} \quad \forall i = 1..n$

Then, the error is minimized for $\sum_{i=1}^{n} a_i = \frac{nk}{\gamma^2} = \frac{n}{\gamma^2 + n} \Longleftrightarrow$

$$k = \frac{\gamma^2}{\gamma^2 + n}$$

Finally our estimator is $\hat{\mu} = \sum_{i=1}^{n} a_i X_i = \frac{1}{\gamma^2(\gamma^2+n)} \sum_{i=1}^{n} X_i$