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

Zadanie 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 $Cov(X^2,Y^2)$.

Zadanie 3. Let (X_1,Y_1), (X_2,Y_2), ..., (X_n,Y_n) be independent random vectors with the same normal distribution with the following parameters: unknown expectation $EX_i=EY_i=m$, the variance $Var\ X_i=\frac{1}{4}Var\ Y_i=1$ and the correlation coefficient $corr(X_i,Y_i)=\frac{1}{2}$. Separately, based on random samples $X_1,X_2,\ldots,X_n$ and $Y_1,Y_2,\ldots,Y_n$, two confidence intervals were built for the expected value m, each at the confidence level of 0.8. Calculate the probability that the intervals constructed in this way turn out to be disjoint.

Zadanie 4.

We have independent observations from three normal distributions with the same unknown variance. For each of the distributions we have a certain group of observations. For each of these groups, the mean and variance from the sample were determined separately. The obtained values, together with the sizes of individual groups, are given in the table:

n_l	$\overline{X}_l$	$S_l^2=\frac{1}{n_l-1}\sum_{i=1}^{n_l}(X_i^l-\overline{X}_l)^2$
5	1,8	0,8
8	2,6	0,5
7	0,9	0,6

We perform an analysis of variance test at the significance level of $\alpha=0.05$, where the null hypothesis is that the expected values in all three groups are the same. What is the result of this test?

Zadanie 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.

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

let $K = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \rightsquigarrow \mathcal{N}\left(\mu = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}\right)$

let $Z_1 = X + Y + aZ + 3$, $K' = \begin{pmatrix} Z_2 \\ Z_3 \end{pmatrix} = \begin{pmatrix} X + 2Z + 4 \\ X - bY + Z \end{pmatrix}$

Z_1 independent from $K' \iff Z_1$ ind. from $Z_2 \wedge Z_3 \iff$

$\mathrm{cov}(Z_1, Z_2) = \mathrm{cov}(Z_1, Z_3) = 0$

$\mathrm{cov}(Z_1, Z_2) = \mathrm{cov}(X + Y + aZ + 3, X + 2Z + 4) =$

$\mathrm{cov}(X, X) + 2\mathrm{cov}(X, Z) + \mathrm{cov}(Y, X) + 2\mathrm{cov}(Y, Z) + a\mathrm{cov}(Z, X) + 2a\mathrm{cov}(Z, Z) =$

$2 + 2\cdot 0 + (-1) + 2\cdot 0 + a\cdot 2\cdot 0 + 2a\cdot 3 = 1 + 6a = 0 \Rightarrow a = -1/6$

$\mathrm{cov}(Z_1, Z_3) = \mathrm{cov}(X + Y - \frac{1}{6}Z + 3, X - bY + Z) =$

$\mathrm{cov}(X, X) - b\mathrm{cov}(X, Y) + \mathrm{cov}(X, Z) + \mathrm{cov}(Y, X) - b\mathrm{cov}(Y, Y) + \mathrm{cov}(Y, Z)$

$-\frac{1}{6}\left(\mathrm{cov}(Z, X) - b\mathrm{cov}(Z, Y), + \mathrm{cov}(Z, Z)\right) =$

$2 + b + 0 - 1 - 4b + 0 - \frac{1}{6}(0 - b\cdot 0 + 3) = \frac{1}{2} - 3b = 0 \Rightarrow b = 1/6$

Z_1, K ind. $\iff b = -a = 1/6$

Zadanie 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 $Z = \begin{pmatrix} X \\ Y \end{pmatrix} \rightsquigarrow N\left(\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, A = \begin{pmatrix} 1 & p \\ p & 1 \end{pmatrix}\right)$

we know $var[X] = E[X^2] - E[X]^2$, $cov(X, Y) = E[XY] - E[X]E[Y]$

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

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

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

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

let $K = \begin{pmatrix} U \\ V \end{pmatrix} \rightsquigarrow N(0, I)$, $L = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

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

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

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

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

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

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

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

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

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

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

$E[Y^2] = E[(Y-0)^2] = E[(Y-\mu_2)^2] = var[Y] = 1$

Zadanie 3. Let (X_1,Y_1), (X_2,Y_2), $\dots$, (X_n,Y_n) be independent random vectors with the same normal distribution with the following parameters: unknown expectation $EX_i = EY_i = m$, the variance $Var\ X_i = \frac{1}{4} Var\ Y_i = 1$ and the correlation coefficient $corr(X_i,Y_i) = \frac{1}{2}$. Separately, based on random samples $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_n$, two confidence intervals were built for the expected value m, each at the confidence level of 0.8. Calculate the probability that the intervals constructed in this way turn out to be disjoint.

$Var[X_i]=1$, $Var[Y_i]=4$

$$\rho_{X_iY_i} = \frac{cov(X_i,Y_i)}{\sqrt{Var[X_i]Var[Y_i]}} = 1/2 \Rightarrow cov(X_i,Y_i) = \frac{1}{2}\cdot\sqrt{4\cdot 1} = 1$$

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \rightsquigarrow N\left(\mu_i = \begin{pmatrix} m \\ m \end{pmatrix}, A_i = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}\right)$$

Conf. $= 0'8 \Rightarrow \alpha = 1-0'8 = 0'2$

We can take confidence intervals for the unbiased point estimator of m:

$$I_1 = \left]\overline{X} - \frac{\sqrt{Var[X_i]}}{\sqrt{n}} u_{1-\frac{\alpha}{2}}, \overline{X} + \frac{\sqrt{Var[X_i]}}{\sqrt{n}} u_{1-\frac{\alpha}{2}}\right[= \left]\overline{X} - \frac{1}{\sqrt{n}} u_{0'9}, \overline{X} + \frac{1}{\sqrt{n}} u_{0'9}\right[$$

$$I_2 = \left]\overline{Y} - \frac{\sqrt{Var[Y_i]}}{\sqrt{n}} u_{1-\frac{\alpha}{2}}, \overline{Y} + \frac{\sqrt{Var[Y_i]}}{\sqrt{n}} u_{1-\frac{\alpha}{2}}\right[= \left]\overline{Y} - \frac{2}{\sqrt{n}} u_{0'9}, \overline{Y} + \frac{2}{\sqrt{n}} u_{0'9}\right[$$

$$I_1 \cap I_2 = \emptyset \Leftrightarrow \overline{X} - \frac{1}{\sqrt{n}} u_{0'9} \geq \overline{Y} + \frac{2}{\sqrt{n}} u_{0'9} \vee \overline{X} + \frac{1}{\sqrt{n}} u_{0'9} \leq \overline{Y} - \frac{2}{\sqrt{n}} u_{0'9} \Leftrightarrow$$
$$\overline{X} - \overline{Y} \geq \frac{3}{\sqrt{n}} u_{0'9} \quad \vee \quad \overline{X} - \overline{Y} \leq -\frac{3}{\sqrt{n}} u_{0'9}$$

We have to calculate $P = P\underbrace{\left[\overline{X}-\overline{Y} \geq \frac{3}{\sqrt{n}} u_{0'9}\right]}_{1} + P\underbrace{\left[\overline{X}-\overline{Y} \leq -\frac{3}{\sqrt{n}} u_{0'9}\right]}_{2} - \cancel{P[1 \wedge 2]} \to 0$

Let $\overline{Z}_i = \overline{X}_i - \overline{Y}_i \rightsquigarrow N(\dots,\dots)$, as $\overline{X}_i, \overline{Y}_i \rightsquigarrow N(\dots,\dots)$

$E[Z_i] = E[X_i] - E[Y_i] = m - m = 0 \quad \forall i=1\dots n \Rightarrow E[\overline{Z}] = 0$

$Var[Z_i] = Var[X_i - Y_i] = Var[X_i] + Var[Y_i] - 2cov(X_i,Y_i) = 1+4-2=3$

$\Rightarrow Var[\overline{Z}] = 3/n$

$$\overline{Z} \rightsquigarrow N\left(0, \frac{3}{n}\right) \Rightarrow \frac{\overline{Z}-0}{\sqrt{3/n}} = \frac{\overline{Z}}{\sqrt{3/n}} \rightsquigarrow N(0,1)$$

$$P\left[\overline{Z} \geq \frac{3}{\sqrt{n}} u_{0'9}\right] = P\left[\frac{\overline{Z}}{\sqrt{3/n}} \geq \frac{3}{\sqrt{n}} \cdot \frac{\sqrt{n}}{\sqrt{3}} u_{0'9}\right] = 1 - P\left[\frac{\overline{Z}}{\sqrt{3/n}} < u_{0'9}\sqrt{3}\right] =$$
$$1 - P\left[\frac{Z}{\sqrt{3/n}} < 1'29\cdot\sqrt{3}\right] = 0'0132$$

$$P\left[\overline{Z} \leq \frac{3}{\sqrt{n}} u_{0'9}\right] = 1 - P\left[\frac{\overline{Z}}{\sqrt{3/n}} \leq \frac{3}{\sqrt{n}} \cdot \frac{\sqrt{n}}{\sqrt{3}} u_{0'9}\right] = 1 - P\left[\frac{\overline{Z}}{\sqrt{3/n}} \leq u_{0'9}\sqrt{3}\right] =$$
$$1 - P\left[\frac{Z}{\sqrt{3/n}} \leq 1'29\cdot\sqrt{3}\right] = 0'0132$$

As a result, $P = 0'0132 \cdot 2 = 0'0264$

Zadanie 4.

We have independent observations from three normal distributions with the same unknown variance. For each of the distributions we have a certain group of observations. For each of these groups, the mean and variance from the sample were determined separately. The obtained values, together with the sizes of individual groups, are given in the table:

n_l	$\overline{X}_l$	$S_l^2 = \frac{1}{n_l-1}\sum_{i=1}^{n_l}(X_i^l - \overline{X}_l)^2$
5	1,8	0,8
8	2,6	0,5
7	0,9	0,6

We perform an analysis of variance test at the significance level of $\alpha = 0.05$, where the null hypothesis is that the expected values in all three groups are the same. What is the result of this test?

Let $H_0: \mu_1 = \dots \mu_3$, $H_1: \exists l,m \in \{1 \dots 3\} / \mu_l \neq \mu_m$

Let $k=3$, $n = \sum_{l=1}^{k=3} n_l = 5+8+7 = 20$, $1-\alpha = 0'95$,

$$\overline{X} = \frac{1}{n}\sum_{l=1}^{k}\sum_{i=1}^{n_l} X_{li} = \frac{1}{n}\sum_{l=1}^{k} n_l \overline{X}_l = \frac{1}{20}(5\cdot 1'8 + 8\cdot 2'6 + 7\cdot 0'9) = 1'805$$

We will reject $H_0 \Leftrightarrow$

$$\frac{SSB/(k-1)}{SSW/(n-k)} = \frac{SSB/2}{SSW/17} = \frac{17}{2}\frac{SSB}{SSW} > F_{2,17}(0'95) \approx 3'915$$

$$SSW = \sum_{l=1}^{k}\sum_{i=1}^{n_l}(\overline{X}_{li} - \overline{X}_l)^2 = \sum_{l=1}^{k}(n_l - 1)S_l^2 = 4\cdot 0'8 + 7\cdot 0'5 + 6\cdot 0'6 = 10'3$$

$$SSB = \sum_{l=1}^{k}\sum_{i=1}^{n_l}(\overline{X}_l - \overline{X})^2 = \sum_{l=1}^{k} n_l(\overline{X}_l - \overline{X})^2 =$$

$$5\cdot(1'8 - 1'805) + 8(2'6 - 1'805) + 7(0'9 - 1'805) = 10'7895$$

$$\frac{17}{2}\frac{SSB}{SSW} = 8'5\,\frac{10'7895}{10'3} = 8'904 > F_{2,17}(0'95) = 3'915 \Rightarrow$$

We reject H_0.

Zadanie 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.

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

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

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

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

$$\frac{\partial f}{\partial a_k}(a_1 \ldots a_n) = \gamma^2 2a_k + 2\Big(\sum_{i=1}^{n} a_i - 1\Big) = 0 \iff \gamma^2 a_k + \sum_{i=1}^{n} a_i - 1 = 0$$

$$\iff a_k = \frac{1 - \sum_{i=1}^{n} a_i}{\gamma^2}$$

As a result, $a_j = \dfrac{1 - \sum_{i=1}^{n} a_i}{\gamma^2}$ $\forall j = 1 \ldots n \implies a_i = \dfrac{1 - n a_i}{\gamma^2} \implies a_i = \dfrac{1}{n+\gamma^2}$

Then $\hat{\mu} = \dfrac{1}{n+\gamma^2}\sum_{i=1}^{n} X_i$