Task 1. Multivariate normal distribution with missing completely at random

1. Generate sample X with 100 observations from two-dimensional normal distribution. Parameters of the distribution are the following:

$$\mu = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

2. Replace some random values from sample X by missing values. In each variable (X_1 and X_2) approximately 25% of all observations should be missed. We will call the sample with missing values as X_{mis}.

3. Consider X_{mis}. Estimate mean-vector and covariance-matrix model parameters by using the following methods: a) Complete-case method; b) Available-case method; c) Unconditional mean imputation; d) Regression imputation. Compare results.

4. Replace all values from sample X that less than -1 by missing values (left censoring). We will call this sample as X_{cen}.

5. Consider X_{cen}. Estimate mean-vector and covariance-matrix model parameters by using the following methods: a) Complete-case method; b) Available-case method; c) Unconditional mean imputation; d) Regression imputation. Compare results.

Task 2. Maximum likelihood method for censored data.

1. Generate sample X with 100 observations from exponential distribution (choose the distribution parameter λ by yourself).

2. Replace all values from sample X that less than c by missing values (left censoring). Choose the censoring level c by yourself. We will call this sample as $X^{(1)}$.

3. Consider $X^{(1)}$. a) Estimate model parameter λ by using maximum likelihood method for censored data. b) Remove all missing observations and estimate model parameter λ by using classical estimator for complete data. Compare results.

4. Replace all values from sample X that greater than c by missing values (right censoring). Choose the censoring level c by yourself. We will call this sample as $X^{(2)}$.

5. Consider $X^{(2)}$. a) Estimate model parameter λ by using maximum likelihood method for censored data. b) Remove all missing observations and estimate model parameter λ by using classical estimator for complete data. Compare results.

Task 3. EM-algorithm

1. Generate sample X with 100 observations from two-dimensional normal distribution. Parameters of the distribution are the following:

$$\mu = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

2. Replace some random values from sample X by missing values. In each variable (X_1 and X_2) approximately 25% of all observations should be missed. We will call the sample with missing values as X_{mis}.

3. Estimate Consider X_{mis}. Estimate mean-vector and covariance-matrix model parameters by using EM-algorithm.