

## Termin 1 - Karty kontrolne

20.11.2024, g. 8:00 - 10:15, s. 623, 623a.

Podsekcja:

Filip Bazarnicki

Kacper Wach

Zadania:

1) Napisać funkcję do wizualizacji danych z próbek pomiarowych.

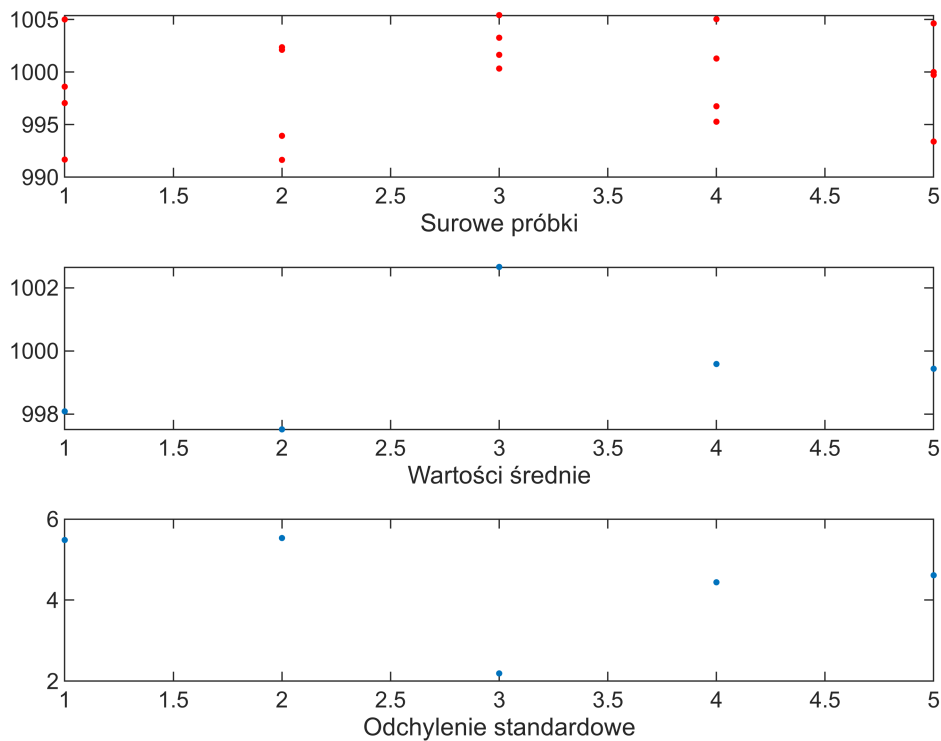
Wejściem funkcji jest macierz  $N \times M$  zawierająca wyniki pomiarów  $M$ -próbek  $N$ -elementowych. Funkcja powinna tworzyć wykres zawierający trzy podwykresy ustawione jeden pod drugim: pierwszy do prezentacji wyników surowych próbek, drugi przedstawiający wartości średnie dla próbek (karta pomiarów wartości średniej), trzeci przedstawiający odchylenia standardowe próbek (karta pomiarów odchylenia standardowego). Oś odciętych w każdym wykresie powinna przedstawiać nr próbki, na osi rzędnych - wartości wskaźników. Przetestować funkcję na wygenerowanych danych.

### Rozwiązanie zapisano w funkcji 'lab1\_zad1(A)'

```
A=randn(4,5)*5+1000;A=sort(A), mean(A), min(min(A)),
```

```
A = 4x5
103 ×
    0.9917    0.9916    1.0003    0.9953    0.9934
    0.9970    0.9939    1.0016    0.9967    0.9997
    0.9986    1.0021    1.0033    1.0013    1.0000
    1.0050    1.0024    1.0054    1.0050    1.0046
ans = 1x5
103 ×
    0.9981    0.9975    1.0027    0.9996    0.9994
ans = 991.6490
```

```
lab1_zad1(A);
```



2) Zapoznać się z funkcją 'boxplot'. Przetestować funkcję 'boxplot' używając danych symulowanych za pomocą funkcji 'rand', 'randn'. Zmodyfikować napisaną powyżej funkcję wykorzystując w niej polecenie 'boxplot'.

help **boxplot**

**boxplot** Displays box plots of multiple data samples.

**boxplot(X)** produces a box plot of the data in X. If X is a matrix there is one box per column, and if X is a vector there is just one box. On each box, the central mark is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme datapoints the algorithm considers to be not outliers, and the outliers are plotted individually.

**boxplot(X,G)** specifies one or more grouping variables G, producing a separate box for each set of X values sharing the same G value or values. Grouping variables must have one row per element of X, or one row per column of X. Specify a single grouping variable in G by using a vector, a character array, a cell array of character vectors, a string array, or a vector categorical array; specify multiple grouping variables in G by using a cell array of these variable types, such as {G1 G2 G3}, or by using a matrix. If multiple grouping variables are used, they must all be the same length. Groups that contain a NaN or an empty string ('') in a grouping variable are omitted, and are not counted in the number of groups considered by other parameters.

By default, character and string grouping variables are sorted in the order they initially appear in the data, categorical grouping variables are sorted by the order of their levels, and numeric grouping variables are sorted in numeric order. To control the order of the groups, you can either use categorical variables in G and specify the order of their levels, or use the 'positions' argument.

`boxplot(AX, X, ...)` produces a box plot in axes with handle AX.

`boxplot(..., 'PARAM1', val1, 'PARAM2', val2, ...)` specifies optional parameter name/value pairs.

'plotstyle'	'traditional' (default), or 'compact' to specify a box style designed for plots with many groups. The plotstyle changes the defaults for some other parameters, as described below.
'boxstyle'	'outline' (default) to draw an unfilled box with dashed lines for whiskers, or 'filled' to draw a narrow filled box with solid lines for whiskers.
'colorgroup'	One or more grouping variables, of the same type as permitted for G, specifying that the box color should change when the specified variables change. Default is [] for no box color change.
'colors'	Colors for boxes, specified as a single color (such as 'r' or [1 0 0]) or multiple colors (such as 'rgbm' or a three-column matrix of RGB values). The sequence is replicated or truncated as required, so for example 'rb' gives boxes that alternate in color. Default when no 'colorgroup' is specified is to use the same color scheme for all boxes. Default with 'colorgroup' is a modified hsv colormap.
'datalim'	A two-element vector containing lower and upper limits, used by 'extrememode' to determine which points are extreme. Default is [-Inf Inf].
'extrememode'	'clip' (default) to move data outside the 'datalim' limits to the limit, or 'compress' to distribute such points evenly in a region just outside the limit, retaining the relative order of the points. A dotted line marks the limit if any points are outside it, and two gray lines mark the compression region if any points are compressed. Values at +/-Inf can be clipped or compressed, but NaNs still do not appear on the plot. Box notches are drawn to scale and may extend beyond the bounds if the median is inside the limit; they are not drawn if the median is outside the limits.
'factordirection'	'data' (default) to arrange the factors with the first value next to the origin, 'list' to arrange the factors left-to-right if on the x axis or top-to-bottom if on the y axis, or 'auto' to use 'data' for numeric grouping variables and 'list' for strings.
'fullfactors'	'off' (default) to have one group for each unique row of G, or 'on' to create a group for each possible combination of group variable values, including combinations that do not appear in the data.
'factorseparator'	Specifies which factors should have their values separated by a grid line. The value may be 'auto' or a vector of grouping variable numbers. For example, [1 2] adds a separator line when the first or second grouping variable changes value. 'auto' is [] for one grouping variable and [1] for two or more grouping variables. Default is [].
'factorgap'	Specifies an extra gap to leave between boxes when the corresponding grouping factor changes value, expressed as a percentage of the width of the plot. For example, with [3 1], the gap is 3% of the width of the plot between groups with different values of the first grouping variable, and 1% between groups with the same value of the first grouping variable but different values for the second. 'auto'

	specifies that <b>boxplot</b> should choose a gap automatically. Default is [].
'grouporder'	Order of groups for plotting, specified as a cell array of strings. With multiple grouping variables, separate values within each string with a comma. Using categorical arrays as grouping variables is an easier way to control the order of the boxes.
'jitter'	Maximum distance D to displace outliers along the factor axis by a uniform random amount, in order to make duplicate points visible. D = 1 makes the jitter regions just touch between the closest adjacent groups. The default is 0.
'labels'	Character array, cell array of strings, or numeric vector of box labels. May have one label per group or per X value. Multiple label variables may be specified via a numeric matrix or a cell array containing any of these types.
'labelorientation'	'horizontal' (default) for horizontal labels, or 'inline' to draw the labels vertically when 'orientation' has its default 'vertical' value.
'labelverbosity'	'all' (default) to display every label, 'minor' to display a label for a factor only when that factor has a different value from the previous group, or 'majorminor' to display a label for a factor when that factor or any factor major to it has a different value from the previous group.
'medianstyle'	'line' (default) to draw a line for the median, or 'target' to draw a black dot inside a white circle.
'notch'	'on' to draw comparison intervals using notches ('plotstyle' is 'traditional') or triangular markers ('plotstyle' is 'compact'), 'marker' to draw them using triangular markers, or 'off' (default) to omit them. Two medians are significantly different at the 5% level if their intervals do not overlap. The interval endpoints are the extremes of the notches or the centers of the triangular markers. When the sample size is small, notches may extend beyond the end of the box.
'orientation'	'vertical' (default) to plot X on the y axis, or 'horizontal' to plot X on the x axis.
'outliersize'	Size of marker used for outliers, in points. Default is 6.
'positions'	Box positions specified as a numeric vector with one entry per group or X value (default 1:NGROUPS when the number of groups is NGROUPS).
'symbol'	Symbol and color to use for outliers, using the same values as the LineSpec parameter S in PLOT. Default is 'r+'. If the symbol is omitted then the outliers are invisible; if the color is omitted then the outliers have the same color as their corresponding box. Any line specification in S is ignored.
'whisker'	Maximum whisker length W. Default is W=1.5. Points are drawn as outliers if they are larger than $Q3+W*(Q3-Q1)$ or smaller than $Q1-W*(Q3-Q1)$ , where Q1 and Q3 are the 25th and 75th percentiles, respectively. The default value 1.5 corresponds to approximately $\pm 2.7$ sigma and 99.3 coverage if the data are normally distributed. The plotted whisker extends to the adjacent value, which is the most extreme data value that is not an outlier. Set 'whisker' to 0 to give no whiskers and to make every point outside of Q1 and Q3 an outlier.
'widths'	A scalar or vector of box widths to use when the 'boxstyle' is 'outline'. The default is half of the

minimum separation between boxes, which is .5 when the 'positions' argument takes its default value. The list of values is replicated or truncated as necessary.

When the 'plotstyle' parameter takes the value 'compact', then the default values for other parameters are the following:

boxstyle - 'filled'	labelverbosity - 'majorminor'
factorgap - 'auto'	medianstyle - 'target'
factorseparator - 'auto'	outliersize - 4
jitter - 0.5	symbol - 'o'
labelorientation - 'inline'	

You can see the data values and group names by using the data cursor tool, available from the figure window. The data cursor shows the original values of any points affected by the 'datalim' parameter. You can label the specific group to which an outlier belongs using the gname function.

To modify the properties of box components, use findobj using tags to find their handles as in one of the examples below. The tag names depend on the plotstyle and are:

```
all styles: 'Box', 'Outliers'
traditional: 'Median', 'Upper Whisker', 'Lower Whisker',
             'Upper Adjacent Value', 'Lower Adjacent Value',
compact:    'Whisker', 'MedianOuter', 'MedianInner'
when 'notch' is 'marker':
             'NotchLo', 'NotchHi'
```

Examples:

```
% Box plot of car gas mileage grouped by country
load carsmall
boxplot(MPG, Origin)
boxplot(MPG, Origin, 'sym','r*', 'colors',hsv(7))
boxplot(MPG, Origin, 'grouporder', ...
        {'France' 'Germany' 'Italy' 'Japan' 'Sweden' 'USA'})

% Plot by median gas mileage
[sortedMPG,sortedOrder] = sort(grpstats(MPG,Origin,@median));
pos(sortedOrder) = 1:6;
boxplot(MPG, Origin, 'positions', pos)

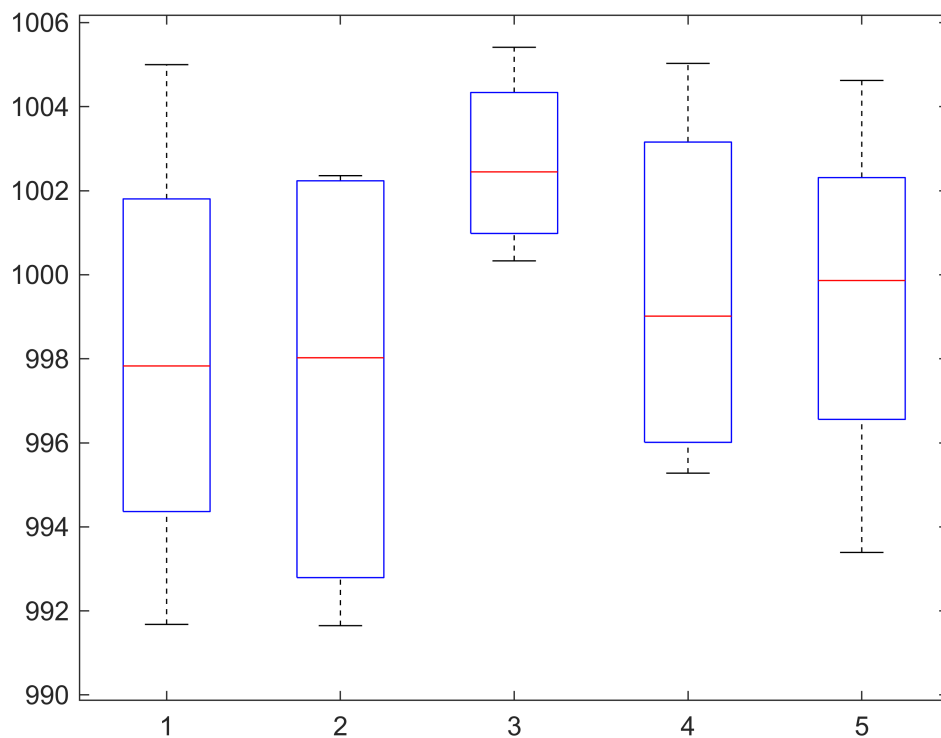
% Change some graphics properties
boxplot(chi2rnd(1,100,10)); % Generate box plot
h=findobj(gca,'tag','Outliers'); % Get handles for outlier lines.
set(h,'Marker','o'); % Change symbols for all the groups.
set(h(1),'MarkerEdgeColor','b'); % Change color for one group
```

See also anova1, kruskalwallis, multcompare.

Documentation for boxplot

```
srednia = mean(A);
odchylene = std(A);

figure
%subplot(3,1,1)
boxplot(A)
```



```
% subplot(3,1,2)
% plot(srednia, '.', 'MarkerSize', 8)
%
%
% subplot(3,1,3)
% plot(odchylenie, '.', 'MarkerSize', 8)
```

3) Napisać funkcję wykreślającą kartę kontrolną wartości średnich wraz z dolnymi i górnymi liniami kontrolnymi. Użyć danych z tabeli 1. Wejściem funkcji jest macierz NxM zawierająca wyniki pomiarów M próbek N-elementowych. Ograniczenie na N: 2 do 25, dla próbek o liczności nie występującej w tab. dokonać interpolacji z sąsiednich danych w tabeli 1.

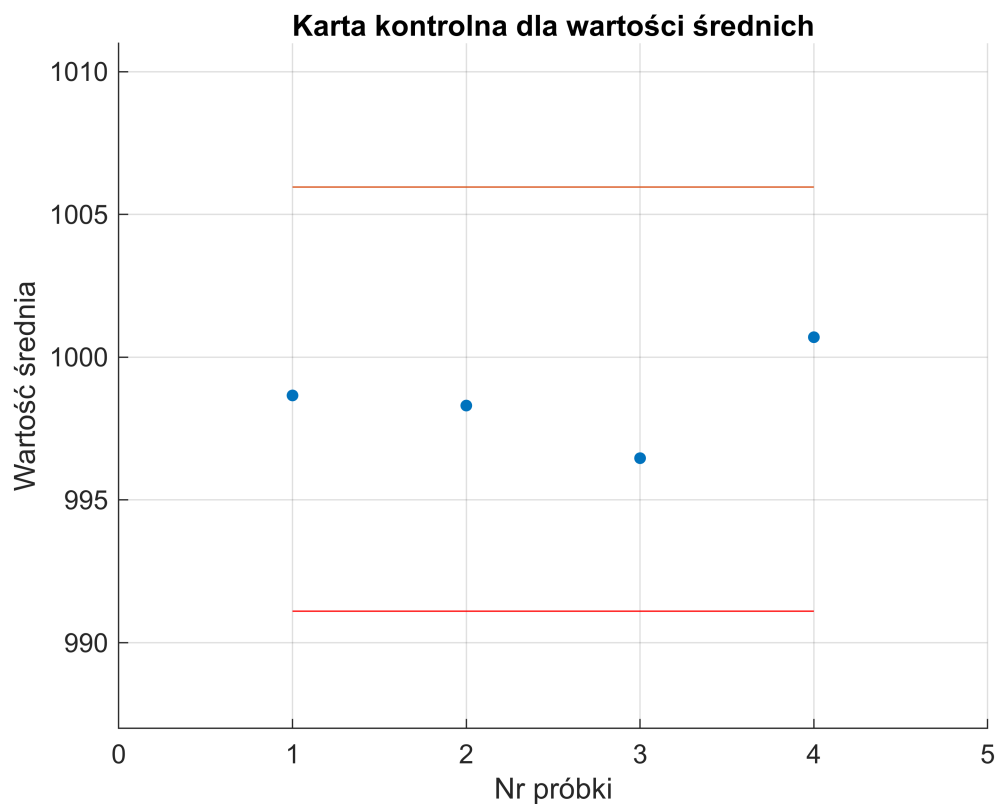
N	A <sub>3</sub> (N)	B <sub>3</sub> (N)	B <sub>4</sub> (N)
2	2.659	0	3.267
3	1.954	0	2.568
4	1.628	0	2.266
5	1.427	0	2.089
6	1.287	0.030	1.970
7	1.182	0.118	1.882
8	1.099	0.185	1.815
9	1.032	0.239	1.761
10	0.975	0.284	1.716
15	0.789	0.428	1.572
20	0.680	0.510	1.490
25	0.606	0.565	1.435

Tabela 1. Mnożniki do wyznaczania górnych i dolnych linii kontrolnych. A<sub>3</sub> – mnożniki dla linii kontrolnych dla wartości średniej; B<sub>3</sub>, B<sub>4</sub> – mnożniki dla dolnej i górnej (odpowiednio) granicy kontrolnej dla odchylenia standardowego. Dla wartości średniej  $\bar{D}(G)LK = \bar{\bar{X}} \pm A_3(N) * \bar{S}$ , Dla odchylenia standardowego  $DLK = B_3(N) * \bar{S}$ ,  $GLK = B_4(N) * \bar{S}$ .

pwd

Zadanie zrealizowano w postaci funkcji **cw1\_zad3(A)**, którą wywołano poniżej:

```
A=randn(5,4)*5+1000;  
lab1_zad3(A)
```



4) Napisać funkcję wykreślającą kartę kontrolną odchylenia standardowego wraz z dolnymi i górnymi liniami kontrolnymi. Użyć danych z poniższej tabeli. Wejściem funkcji jest macierz NxM zawierająca wyniki pomiarów M-próbek N-elementowych. Ograniczenie na N: 2 do 25, dla próbek o liczności nie występującej w tab. dokonać interpolacji z sąsiednich danych w tabeli.

N	A <sub>3</sub> (N)	B <sub>3</sub> (N)	B <sub>4</sub> (N)
2	2.659	0	3.267
3	1.954	0	2.568
4	1.628	0	2.266
5	1.427	0	2.089
6	1.287	0.030	1.970
7	1.182	0.118	1.882
8	1.099	0.185	1.815
9	1.032	0.239	1.761
10	0.975	0.284	1.716
15	0.789	0.428	1.572
20	0.680	0.510	1.490
25	0.606	0.565	1.435

Tabela 1. Mnożniki do wyznaczania górnych i dolnych linii kontrolnych. A<sub>3</sub> – mnożniki dla linii kontrolnych dla wartości średniej; B<sub>3</sub>, B<sub>4</sub> – mnożniki dla dolnej i górnej (odpowiednio) granicy kontrolnej dla odchylenia standardowego. Dla wartości średniej  $D(G)LK = \bar{\bar{X}} \pm A_3(N) \cdot \bar{S}$ , Dla odchylenia standardowego  $DLK = B_3(N) \cdot \bar{S}$ ,  $GLK = B_4(N) \cdot \bar{S}$ .

Zadanie zrealizowano w postaci funkcji **cw1\_zad4(A)**, którą wywołano poniżej:

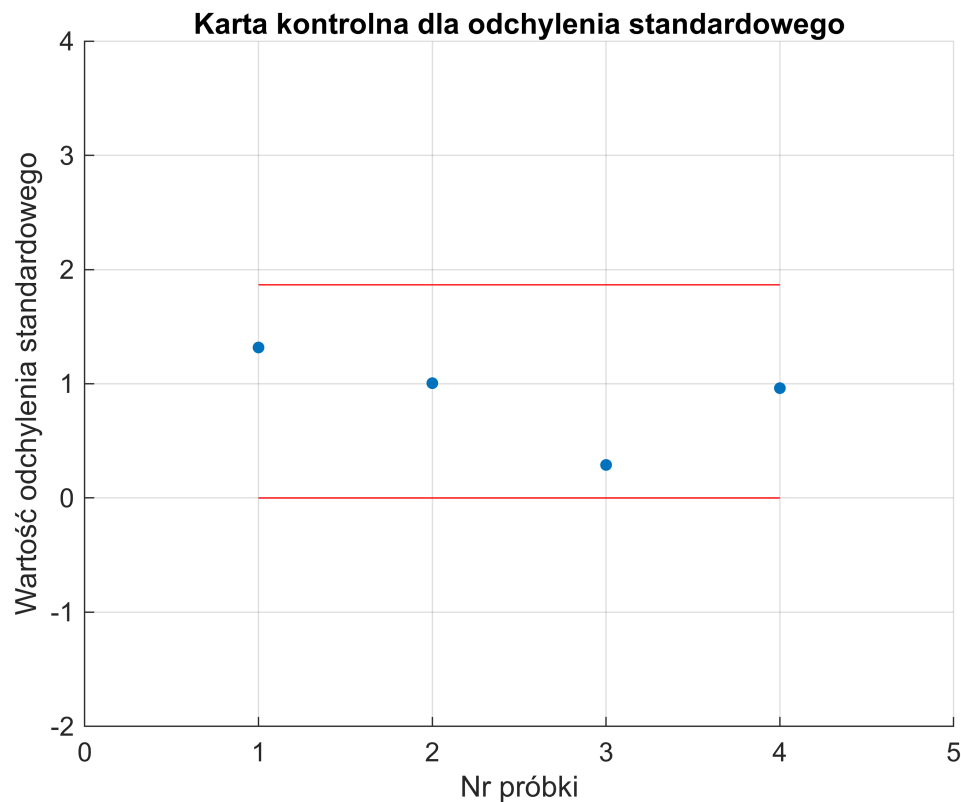
```
A=randn(5,4)
```

```
A = 5x4
```

```
 2.4245    1.8779   -0.5583   -0.2099
 0.9594    0.9407   -0.3114   -1.6989
 -0.3158    0.7873   -0.5700    0.6076
 0.4286   -0.8759   -1.0257   -0.1178
 -1.0360    0.3199   -0.9087    0.6992
```

```
lab1_zad4(A)
```





5) Załóżmy, że zmienne losowe zakłócające stabilny przebieg procesu dodają się do zmiennej odpowiadającej stabilnemu, czyli uregulowanemu, procesowi. Próbkę pobieraną w chwili  $t$  odpowiada obserwacjom pewnej zmiennej losowej  $Y(t)$ ,

$$Y(t) = X_0 + \sum_{i=1}^k I_i(t) \cdot X_i,$$

gdzie  $X_i$  pochodzi z rozkładu  $F_i$  o wartości oczekiwanej  $\mu_i$  i wariancji  $\sigma_i^2$

2, a funkcja wskaźnikowa  $I(t)$  ma postać:  $I(t)=1$  z prawdopodobieństwem  $\pi_i$  lub  $I(t)=0$  z prawdopodobieństwem  $1-\pi_i$ . Jeżeli podany model

jest właściwy i mamy  $k$  możliwych wyznaczalnych przyczyn rozregulowania, to w danej próbce obecna jest jedna z  $2^k$  możliwych kombinacji zmiennych losowych odpowiadających tym przyczynom rozregulowania. Skrajnymi możliwościami spośród wszystkich  $2^k$  kombinacji są: niewystąpienie żadnej przyczyny rozregulowania oraz wystąpienie wszystkich przyczyn rozregulowania. Zakładamy, że próbki pobierane są dostatecznie rzadko, by postać zmiennej  $Y$  dla jednej próbki była niezależna od jej postaci dla innej próbki. Jednocześnie zakładamy, że każda funkcja wskaźnikowa  $I_i$  zachowuje stałą wartość (0 lub 1) w czasie pobierania jednej próbki.

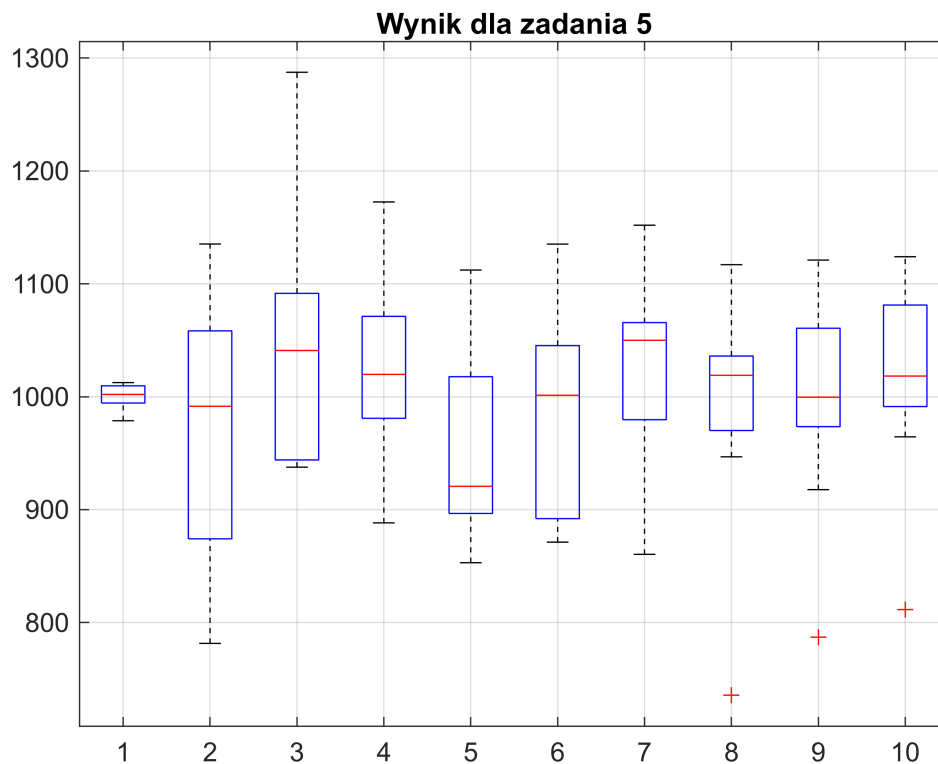
W przypadku szczególnym, gdy rozkłady wszystkich zmiennych losowych są normalne, obserwowana zmienna losowa  $Y(t)$  może być przedstawiona jako:

$$Y(t) = N(\mu_0 + \sum_{i \in I} (i \cdot \mu_i), \sigma_0^2 + \sum_{i \in I} (i \cdot \sigma_i^2)),$$

gdzie  $X_0 = N(\mu_0, \sigma_0^2)$ .

Dla opisanego powyżej modelu napisać funkcję w środowisku Matlab realizującą M-próbek N elementowych (macierz NxM) zmiennej Y dla założonej liczby k przyczyn rozregulowania procesu X0, prawdopodobieństw realizacji zmiennych zakłócających Xi. Założyć, że proces odstawowy oraz procesy zakłócające mają rozkład normalny o zadanych wartościach oczekiwanych i wariancjach.

```
mi0 = 1000;  
sigma0 = 10;  
pz = [0 0.7];  
miz = [-250 0];  
sigmaz = [0 100];  
M=10;  
N=10;  
  
A=lab1_zad5(mi0,sigma0,pz,miz,sigmaz,M,N);  
  
figure;boxplot(A);title('Wynik dla zadania 5');grid on
```



## Funkcje realizujące poszczególne zadania

```
function wy=lab1_zad1(A)

[N,M] = size(A);

srednie = mean(A);
odchylenieStandard =std(A);

figure
subplot(3,1,1)
plot(1:M,A, '.', 'Color', 'r', 'MarkerSize', 8)
xlabel('Surowe próbki')

subplot(3,1,2)
plot(1:M,srednie, '.', 'MarkerSize', 8)
xlabel('Wartości średnie')
% wizualizacja danych z macierzy A
% A: N x M - zawiera M-próbkę po N elementów każda

subplot(3,1,3)
```

```
plot(1:M,odchylenieStandard, '.', 'MarkerSize', 8)
xlabel('Odchylenie standardowe')
```

```
end % of zad 1
```

```
function wy=lab1_zad3(A)
```

```
[n,m] = size(A);
```

```
% Tabela ze współczynnikami do zad. 3:
```

```
N=[2 3 4 5 6 7 8 9 10 15 20 25];
```

```
A3=[2.659 1.954 1.628 1.427 1.287 1.182 1.099 1.032 0.975 0.789 0.680 0.606];
```

```
interp1 = interp1(N,A3,n);
```

```
figure
```

```
hold on
```

```
grid on
```

```
title('Karta kontrolna dla wartości średnich');
```

```
plot(mean(A), '.', 'MarkerSize', 15);
```

```
xlim([0 m+1]);
```

```
ylim([floor(min(min(A))) ceil(max(max(A)))]);
```

```
xlabel('Nr próbek');
```

```
ylabel('Wartość średnia');
```

```
odchylenie=mean(std(A));
```

```
srednia=mean(mean(A));
```

```
nr=1:m;
```

```
plot(nr,ones(size(nr))*(srednia+interp1*odchylenie),nr,
ones(size(nr))*(srednia-interp1*odchylenie), 'r')
```

```
%xmm = mean(mean(A));
```

```
end % zadanie 3
```

```
function wy=lab1_zad4(A)
```

```
[n,m] = size(A);
```

```
% Tabela ze współczynnikami do zad. 4.
```

```
N=[2 3 4 5 6 7 8 9 10 15 20 25];
```

```
B3=[0 0 0 0 0.030 0.118 0.185 0.239 0.284 0.428 0.510 0.565];
```

```
B4=[3.267 2.568 2.266 2.089 1.970 1.882 1.815 1.761 1.716 1.572 1.490 1.435];
```

```

interB3 = interp1(N,B3,n);

interB4 = interp1(N,B4,n);
odchylenie=mean(std(A));

nr=1:m;
figure
hold on
grid on
title('Karta kontrolna dla odchylenia standardowego');
std_values = std(A); % Obliczenie odchylenia standardowego
plot(std_values, '.', 'MarkerSize', 15);
xlim([0 m+1]);
ylim([floor(min(std_values)-2) ceil(max(std_values)+2)]); % Automatyczne skalowanie
osi Y
xlabel('Nr próbek');
ylabel('Wartość odchylenia standardowego');
plot(nr, ones(size(nr))*(interB3*odchylenie), 'r');
plot(nr, ones(size(nr))*(interB4*odchylenie), 'r');

%Vq = interp1(X,V,Xq) interpolates to find Vq, the values of the
%   underlying function V=F(X) at the query points Xq.

end % zad 4

```

```

function A=lab1_zad5(mi0,sigma0,pz,miz,sigmaz,M,N)

xo = mi0 + randn(N,M)*sigma0;

dlu = length(pz);

A=[];
for j=1:M
    I=zeros(dlu,1);
    for i=1:dlu
        if rand<pz(i)
            I(i)=1;
        end
    end
end

Z = [];

for i=1:dlu
    Z=[Z (randn(N,1)*sigmaz(i)+miz(i))*I(i)];
end
if dlu == 1
    K = Z;
end

```

```
    else
        K = (sum(Z'))';
    end
A = [A, K];
end
A = xo + A;
end
```