

Sprawozdanie

Metody Probabilistyczne w Informatyce

Paweł Morozewicz 236421 3i3

SPIS TREŚCI:

1. Dane
2. Histogramy
3. Test Serii
4. Eliminacja błędów grubych
5. Rozkład normalny
6. Rozkład logarytm normalny

1. Dane

W celu wykonania ćwiczeń z przedmiotu wykorzystałem przypisany do mnie zestaw danych oznaczony numerem 6. Składał się on z 1000 pomiarów. Poniżej zrzut ekranu przedstawiający wygląd zbioru danych nr.6.

Rys.1

399,3	401	400,9	395,6	396,8	396,1	389,5	396,3	393,6	399,2	398,6	399,1	399,8	399,3	396,3	397,7	404,3	402,8	402,3
399,3	399,9	401,2	395,4	396,7	392,9	391,5	396,2	393	395,9	398,6	399,2	400,1	398,2	396,4	396,5	404,8	402,8	401,1
399,4	400,7	402,1	395,4	396,6	392,4	395,9	396,4	392,7	392	398,7	399	399,4	398,6	396,3	398,1	404	402,3	401,5
400,1	400,7	402,1	396	396,9	393,6	396	396,7	391,6	390,9	398,9	398,7	399	397,7	395,6	398,3	403,4	402,3	402
399,8	400,3	401,4	395,3	396,6	392,9	395,9	395,7	392,2	390,6	398,8	398,6	398,2	397,8	396	397,8	404,2	402,1	400,7
400,5	401,1	401,9	395,5	396,6	394,4	394,6	396,6	392,2	390,4	398,7	398,1	397,1	397,1	395,8	397,8	402,9	401,5	401
398,6	401	402,2	395,6	397,9	393,6	395	396,7	392,6	399,3	399	400,9	397,7	397,6	395,6	399,1	403,2	402,5	399,6
398,5	400,4	398	394,9	396	393,7	394,6	396,4	391,8	399,2	398,5	399,3	397,3	397,7	396,6	399	403,6	402,3	401,4
398,9	401,7	399,2	395,3	396	393,3	392,4	393,5	396,9	397	398,9	399,7	397,3	397,7	394,7	398,6	403,3	403	401
399,6	401,2	399,6	395	396,6	394,2	395,3	388,5	397,6	399	399,3	400,3	396,9	397,9	395,5	401,6	402,9	403,1	401,5
402,9	400,4	399,9	395,6	396,6	394,4	394,4	387,1	397,7	400,2	398,9	400,8	396,9	398,2	395,6	400,1	403,3	403,2	401,4
399,2	401,8	399,6	394,6	397,2	394,3	394,6	387,4	397,7	399,9	399,1	400	397,2	397,7	395,5	400,3	403,7	402,3	401,4
397,7	400,9	400,9	394,3	398,8	394,7	391,5	386,6	399,4	398,6	398,9	399,9	397,2	397,5	395,5	400,9	403,6	402,4	402,1
391,5	400,9	399,1	394,2	396,5	394	390	390,9	399,5	400,1	398,6	399,5	397,2	396,9	395,5	400,3	403,6	402,3	402,4
396,4	401,5	399,3	397,9	396,9	393	392,7	395,6	395,3	399,4	398,4	400,3	396,9	397,4	395,6	400,8	403,2	402,7	401,7
397,8	401,2	397,5	396,8	396,7	393,3	394,8	396,3	399,1	399,6	398,5	400,4	396,9	397,9	395,7	400,7	403,8	402,6	402,3
398,1	400,6	398,1	397,2	396,1	393,3	394,3	395,9	399,4	398,6	398	400,9	396,3	397,7	395,4	400,8	403,5	402,6	401,6
398,3	402,1	397,2	394,7	394,8	393,6	394,3	396,1	399	398,8	398,8	401,1	397,9	395,5	394,8	399,5	402,5	403	402
398,6	403	397,2	397,1	397,7	394,4	396,6	396,6	395,8	399,4	398,8	400,8	398,2	395,7	396,3	399,6	402,9	403,5	401,6
399,1	399,7	397,1	397,6	396,8	392,8	396,7	396,6	393,9	398,8	399,9	400,9	398,4	395,7	398,4	400	403,9	402,8	401,9
399,2	400,8	396,3	398,3	397,4	393,9	396,1	396,5	393,8	398,9	399,6	400,7	397,8	396	398	400	403,1	403,1	402,1
399	400,6	397,5	397,9	399,3	393,6	396,1	397,5	393,5	400,3	399	400,8	397,5	395,9	399,5	400,1	402,8	400,5	401,8
403,3	400,2	397,3	398,2	398,5	393,2	396,2	398	394,6	400,9	399,8	401,1	397,7	395,8	397,5	400	403,4	401,4	401,4
399,8	401	397,1	396,3	398	393,6	396	397,3	393,7	401,3	399,3	401,6	396,4	395,5	397,8	400,1	402,5	401,5	401,4
399,9	400,6	397,9	395,7	397,5	393,4	395,8	397	398,6	398,5	399,8	401,3	398,6	394,9	399,8	400,6	403,4	401,3	401,6
401,3	400,3	397,2	396	397,4	393,5	396,1	396,5	398,4	398,7	398,7	401,5	398,4	397,6	396,7	400,5	403,5	401	402
401,2	400,9	395,7	396,8	399	393,9	395,8	397,3	396,1	398,3	399,2	401,4	397,7	396,6	398	400,6	403,8	400,1	402,4
399,5	399,7	396,4	396,9	396,6	392,4	396	398,1	390,8	398,2	398,5	402,5	397,8	396,4	398,5	400,6	403,9	400	402,8
400,2	400,6	393	397,2	396	392,4	396,1	398,4	390,5	398,2	398,6	401,4	397,6	396,2	398,5	400,5	404,1	399,4	403,2
397,3	399,3	394,9	397	396	392,3	396,1	398,5	390,5	398,7	398,7	401,5	398	396,2	399,1	401	403,6	401	403
399,9	399,9	394,8	396	397,5	392,1	396,2	399,2	397,6	398,5	399,2	401,2	398,5	396,7	398,3	400,2	404,2	402,9	403,1
399,5	400,5	395	396,5	395,9	392,5	396,5	399,4	398,9	399,2	403,2	401	398,4	396,3	398	400,2	404,2	402,7	402
399,7	400,4	395,1	396,6	392,5	392,3	396,5	398,1	399,1	397,9	397,4	399	399	396,8	398,1	400,8	403	402,2	403
399,5	399,6	395,3	396	393,8	392,3	396,7	397,4	399	398,8	398,7	399,3	398,5	395,9	398,1	401,9	401,6	402	403,1
400,9	401,8	395,8	398	393,5	392,4	397	397,6	396,2	398	398,8	399,1	398,5	396	398,6	401,6	403	402,7	403,6
400,7	400,2	395,6	395	393,3	392,5	397,1	397,9	398,6	398,4	398,8	399,3	398,7	396,1	398,7	400,6	402,7	401,8	403,1
401,4	400,3	395,7	395,1	393,6	392,4	397,1	398,3	398,7	397,8	398,8	399,3	398,7	396,4	398,2	401	403,7	402,6	402
401,4	401,5	395,7	394,9	393,6	391,5	396,9	398,5	399,2	397,1	399	398,1	398,1	396,2	398,3	402,9	403,7	402,2	401,4
400,9	400,2	395,7	395,5	393,7	392,2	397,1	398,7	398,4	397,4	399,1	400,4	398,8	396,5	398,1	403,2	402,6	401,7	401,6
401,5	400,3	395,2	395,2	393,7	391,3	396,6	398,5	398,6	396,7	399,5	400,1	399,1	396,4	398,3	403,8	403,5	402	402,1
402,3	401,6	395,6	395,6	393,4	390,1	396,1	398,8	398,4	396,9	399,6	400,3	398,9	396,5	398,3	405,7	404,1	402,6	403,7
401,8	401,1	394,4	395,9	393,5	391,5	396,1	398,7	398,1	398,2	399	399,9	398,4	395,9	398,4	404	402,9	402,3	403
400,8	402,4	394,6	396,1	393,6	392,3	396,2	399,1	397,2	394,8	399,1	400	398,5	396,4	397,9	404,7	402,3	402,3	403,9
401	399,6	394,5	396,8	394,2	391,2	396,1	398,8	399,4	395,8	398,9	399,9	398,3	397,1	401,3	405,5	402	402,6	403,2
401,6	401	394,5	396,5	394	392,5	396,4	398,8	399,5	398,9	399,6	400,3	399,1	394,9	400,4	404,6	402,2	401,9	403,5
401	401	394,8	395,9	394,8	391,6	395,5	398,6	399,8	398,2	399,9	400,3	399,3	395,3	401,4	404,4	401,5	402,4	402,9
401	402,1	394,3	396,9	394,6	389,9	395,6	398,9	399,1	399,2	400,2	400,3	398,9	396,1	400,7	405,3	402,2	401,9	
401,9	401,5	395	396,4	394,7	393,1	396,4	398,6	398,9	399,5	400,5	400	398,7	395,9	398,4	404,9	402,1	401,7	
400,4	399,8	395,1	396,7	394,9	395,2	396,8	398,3	399,1	399,7	399,7	400,2	398,8	395,6	399,1	404,3	402,4	402,7	
400,3	401,5	395	396,7	395	393,2	396,4	397,8	399,4	399	399,5	400,1	398,8	396,9	399,3	404,2	402,6	402	
401,3	400,7	395,5	396,6	395,1	392,8	396,3	399,2	399,1	397,9	399,4	399,8	398,8	395,9	398,9	404,6	402,8	401,6	
400,2	400,6	395,5	397,1	394,4	391,8	393,4	398,6	399,3	398	399	400	399,1	395,6	398,9	405,2	403,1	402,9	
400,3	401,5	395,4	397,5	395,2	391,5	391,3	398,9	399,2	399	398,6	399,9	399,2	396,7	398,3	404,3	402,9	400,7	

Rys1. Przedstawia zestaw danych nr.6

2. Histogramy

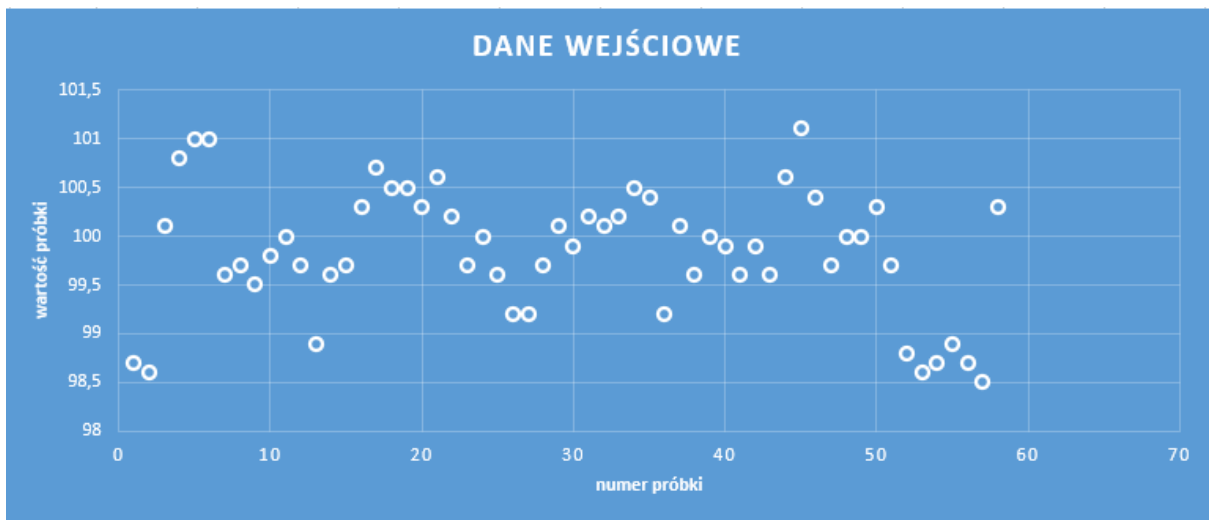
W tym ćwiczeniu należało wyznaczyć histogramy dla 5, 10 oraz 15 klas do rozwiązania tego zadania uprzednio wyznaczyłem wartość średnią, wartość minimalną oraz wartość maksymalną.

Rys.2.

Lp.	dane wejściowe	wartość średnia	wartość min	wartość max
1	98,7	99,84	98,5	101,1
2	98,6			
3	100,1			
4	100,8			
5	101		max - min	2,6
6	101			
7	99,6			
8	99,7			
9	99,5			
10	99,8			

Rys2. Przedstawia fragment wyznaczonych danych.

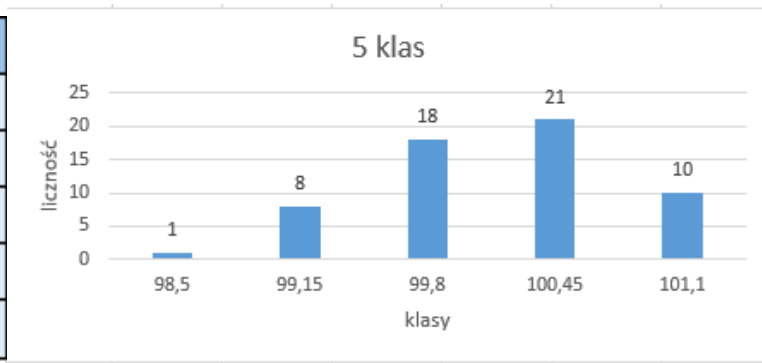
Rys.3.



Rys.3. przedstawia wykres danych wejściowych.

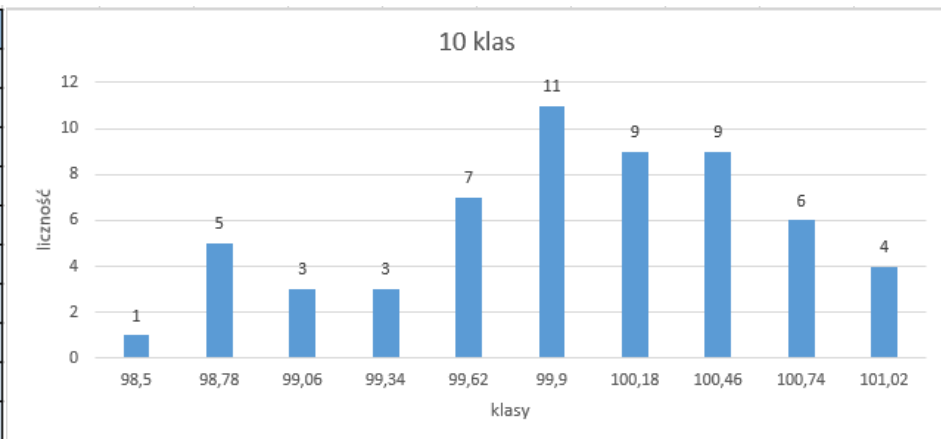
Wyznaczenie dla 5 klas:

klasy	liczność
98,5	1
99,15	8
99,8	18
100,45	21
101,1	10



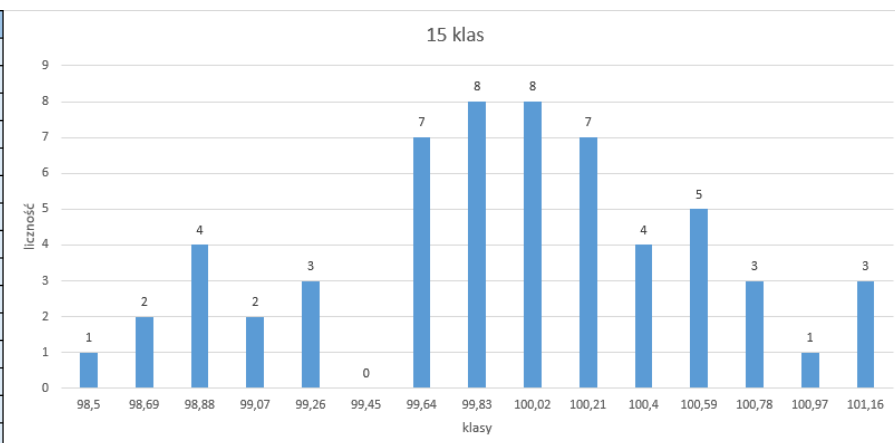
Wyznaczenie dla 10 klas:

klasy	liczność
98,5	1
98,78	5
99,06	3
99,34	3
99,62	7
99,9	11
100,18	9
100,46	9
100,74	6
101,02	4



Wyznaczenie dla 15 klas:

klasy	liczność
98,5	1
98,69	2
98,88	4
99,07	2
99,26	3
99,45	0
99,64	7
99,83	8
100,02	8
100,21	7
100,4	4
100,59	5
100,78	3
100,97	1
101,16	3



3. Test Serii

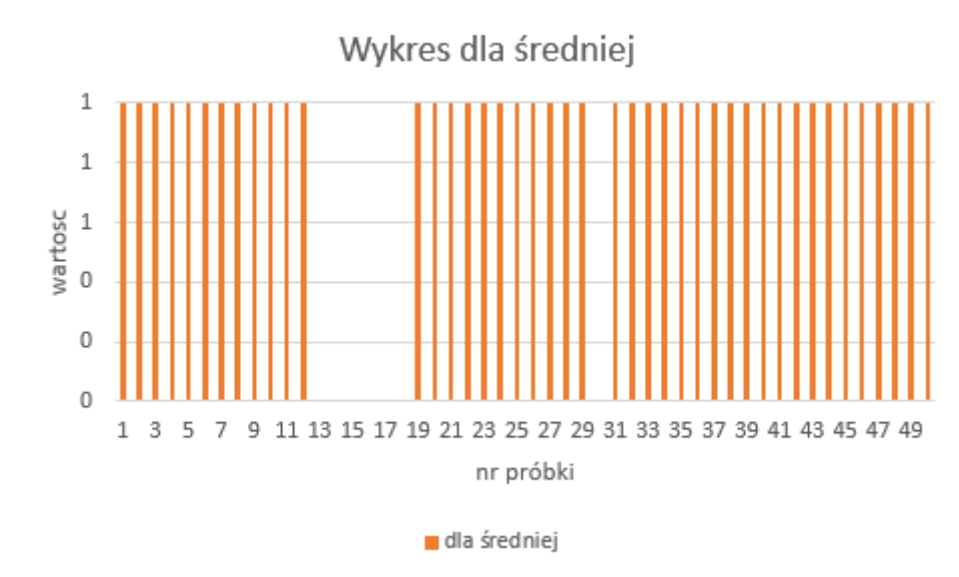
W tym ćwiczeniu należało wyznaczyć 2 wykresy dla wartości średniej oraz mediany do rozwiązania tego zadania uprzednio wyznaczyłem wartość średnią, medianę. Następnie wygenerowałem dane dla średniej oraz dla mediany oraz sporządziłem nowe wykresy.

LP	dane wejściowe	dla średniej	dla mediany	test dla średniej	test dla mediany
1	399,3	1	1	1	1
2	399,3	1	1	1	1
3	399,4	1	1	1	1
4	400,1	1	1	1	1
5	399,8	1	1	1	1
6	400,5	1	1	1	0
7	398,6	1	0	1	1
8	398,5	1	0	1	0
9	398,9	1	1	1	1
10	399,6	1	1	1	1
11	402,9	1	1	1	1
12	399,2	1	1	0	0
13	397,7	0	0	1	1
14	391,5	0	0	1	1
15	396,4	0	0	1	1
16	397,8	0	0	1	1
17	398,1	0	0	1	1
18	398,3	0	0	0	1
19	398,6	1	0	1	0
20	399,1	1	1	1	1
21	399,2	1	1	1	1
22	399	1	1	1	1
23	403,3	1	1	1	1
24	399,8	1	1	1	1
25	399,9	1	1	1	1
26	401,3	1	1	1	1
27	401,2	1	1	1	1
28	399,5	1	1	1	1
29	400,2	1	1	0	0
30	397,3	0	0	0	0
31	399,9	1	1	1	1

Następnie wyznaczyłem kilka wartości które pomogły mi potem wyznaczyć wartości krytyczne k_1 i k_2 , takich jak liczba zer, jedynek, liczność próbek itp.

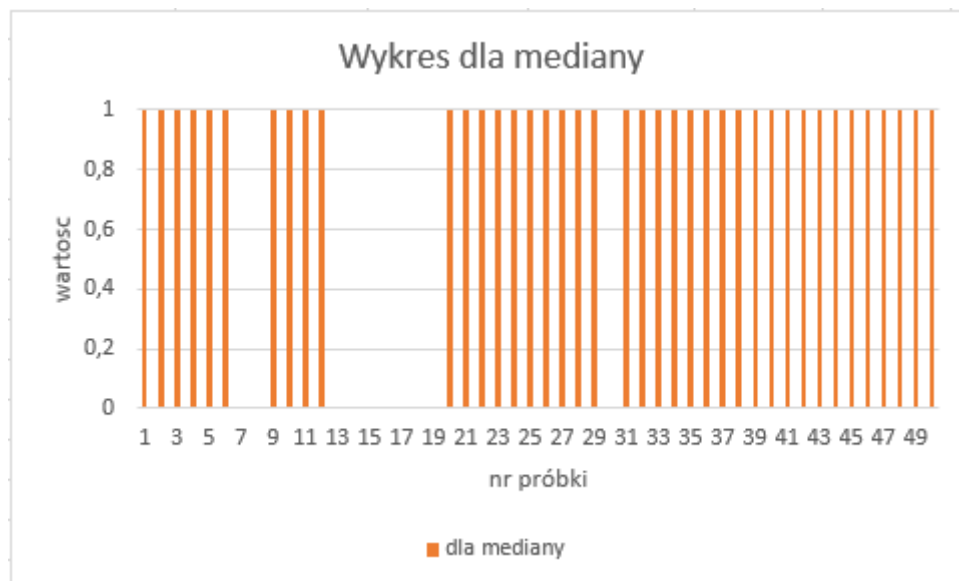
wartość średnia	mediana		
398,39	398,6		
		dla średniej	dla mediany
	liczba zer	460	509
	liczba jedynek	540	491
	liczność próbek	1000	1000
	u	497,8	500,838
	ro	15,70	15,80
	K1	472,0	474,9
	K2	523,6	526,8
	liczba serii	919	903

Rys.3.



Rys.3. przedstawia wykres dla wartości średniej

Rys.4.



Rys.4. przedstawia wykres dla mediany

Pod koniec sprawdziłem czy uzyskane dane liczby serii dla wartości średniej oraz mediany zawierały się w przedziale losowości. ($K1 < \text{liczba serii} < K2$). W przypadku moich danych nie zawierały się w tym przedziale więc postulat losowości **nie** został spełniony.

4. Eliminacja błędów grubych

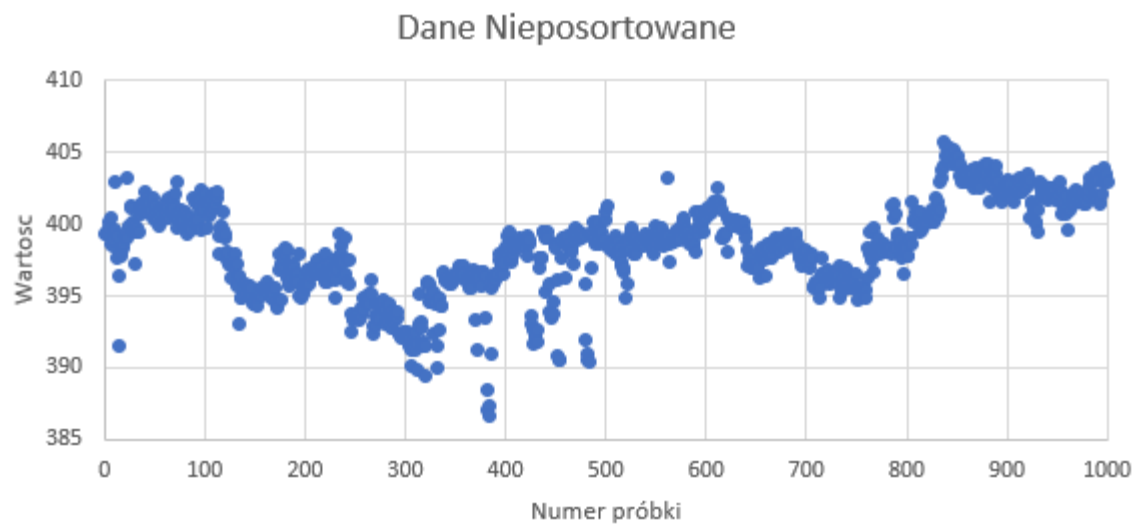
W tym ćwiczeniu należało sprawdzić czy nasze dane są obarczone błędem grubym. W tym celu uporządkowałem dane i posortowałem je rosnąco oraz wyznaczyłem wykres.

NR próbki	dane wejściowe	dane posort.
1	399,3	386,6
2	399,3	387,1
3	399,4	387,4
4	400,1	388,5
5	399,8	389,5
6	400,5	389,9
7	398,6	390
8	398,5	390,1
9	398,9	390,4
10	399,6	390,5
11	402,9	390,5
12	399,2	390,6
13	397,7	390,8
14	391,5	390,9
15	396,4	390,9
16	397,8	391,2
17	398,1	391,3
18	398,3	391,3
19	398,6	391,5
20	399,1	391,5

Następnie Wyznaczyłem charakterystyczne wartości dla celów wykonania zadania (wartość max, wartość min, średnia, odchylenie standardowe. W dalszej części zadania wyznaczyłem B4 plus i B4 minus czyli statystyki testowe B4. Następnie obliczyłem wartość krytyczną b4.

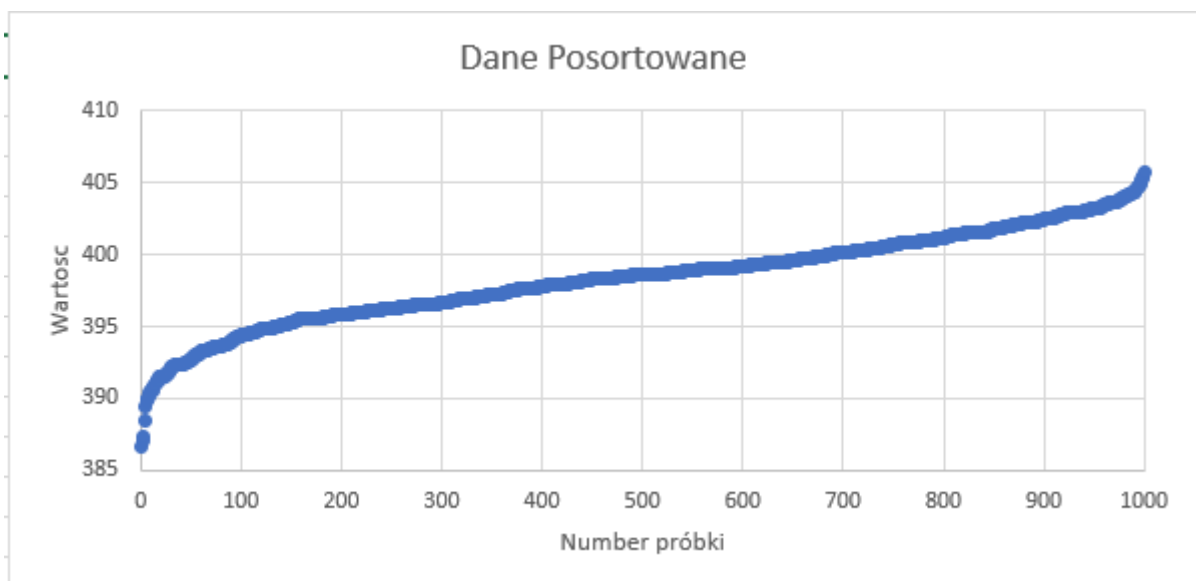
Wartość Max	405,7	B4plus	2,32
Wartość Min	386,6	B4 minus	3,75
średnia	398,39		
Odchylenie standardowe	3,15	a=	0,05
		b=	0,99995
		y=	3,89
		b4	3,88

Rys.5.



Rys.5. Przedstawia wykres danych nieposortowanych.

Rys.6.



Rys.6. przedstawia wykres danych posortowanych.

Wnioskiem końcowym tego zadania jest że wyniki nie są obciążone błędem standardowym ponieważ wartości **B4** są **mniejsze** od **b4**.

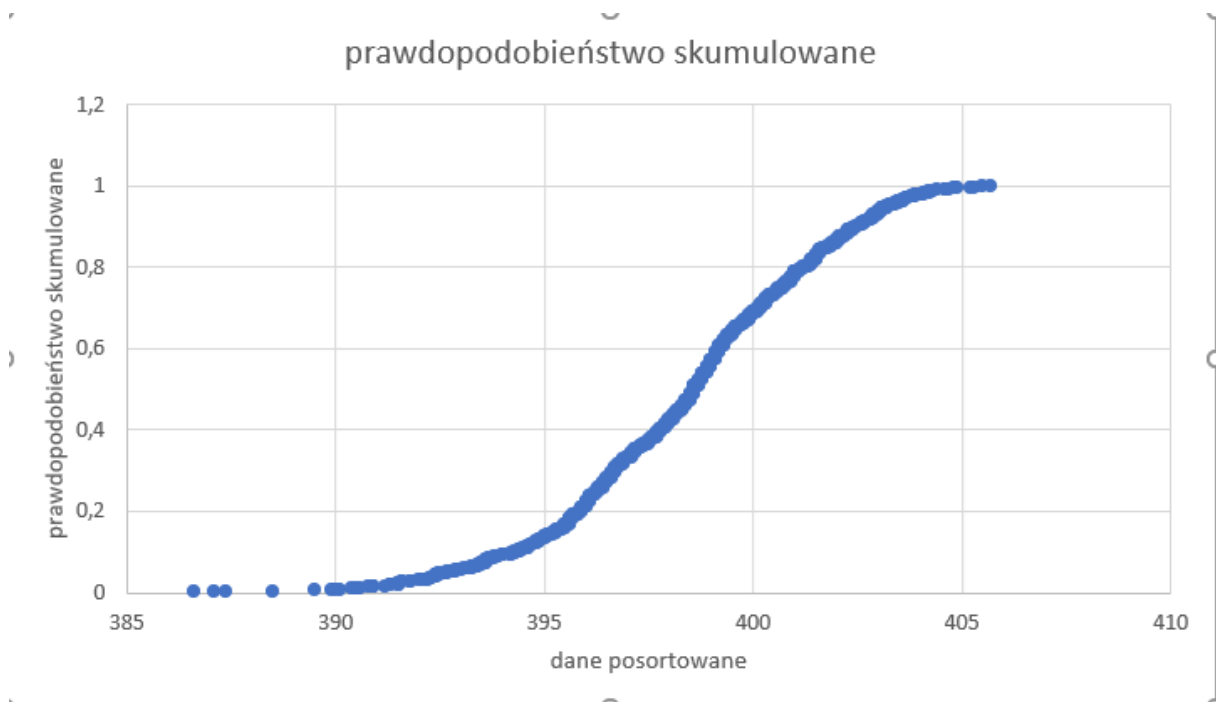
5. Rozkład normalny

W tym ćwiczeniu należało wyznaczyć wykres który przedstawia dystrybuantę empiryczną na tle teoretycznej rozkładu normalnego. W tym celu zacząłem od posortowanie danych rosnąco oraz wyznaczeniu kilku kluczowych danych dla zadania (średnia, mediana, max, min, odchylenie standardowe, ilość próbek).

Następnie wyznaczyłem prawdopodobieństwo skumulowane jak i wykres dla niego.

id	dane wejściowe	dane posortowane	średnia	mediana	max	min	odchylenie standardowe	ilość próbek	prawdopodobieństwo skumulowane	wartość standaryzowana	Dystrybuanta teoretyczna rozkładu normalnego
1	399,3	386,6	398,386	398,6	406	386,6	3,146194662	1000	0,000999001	-3,090529138	8,97638E-05
2	399,3	387,1							0,001998002	-2,878477015	0,000167069
3	399,4	387,4							0,002997003	-2,748109103	0,00023977
4	400,1	388,5							0,003996004	-2,652407258	0,00083812
5	399,8	389,5							0,004995005	-2,5761749	0,00236802
6	400,5	389,9							0,005994006	-2,512497013	0,003494982
7	398,6	390							0,006993007	-2,457622404	0,003843267
8	398,5	390,1							0,007992008	-2,409280312	0,004222343
9	398,9	390,4							0,008991009	-2,365988191	0,005568034
10	399,6	390,5							0,00999001	-2,326722867	0,006094663
11	402,9	390,5							0,010989011	-2,290747494	0,006094663
12	399,2	390,6							0,011988012	-2,257513226	0,006664962
13	397,7	390,8							0,012987013	-2,226599895	0,007948709
14	391,5	390,9							0,013986014	-2,197678455	0,008668596
15	396,4	390,9							0,014985015	-2,170486239	0,008668596
16	397,8	391,2							0,015984016	-2,144810117	0,01118211
17	398,1	391,3							0,016983017	-2,120474688	0,012150398
18	398,3	391,3							0,017982018	-2,097333809	0,012150398
19	398,6	391,5							0,018981019	-2,075264389	0,01430674
20	399,1	391,5							0,01998002	-2,054161741	0,01430674
21	399,2	391,5							0,020979021	-2,033936067	0,01430674
22	399	391,5							0,021978022	-2,014509734	0,01430674
23	403,3	391,5							0,022977023	-1,995815161	0,01430674
24	399,8	391,5							0,023976024	-1,977793138	0,01430674
25	399,9	391,6							0,024975025	-1,960391488	0,015503351
26	401,3	391,6							0,025974026	-1,943563989	0,015503351

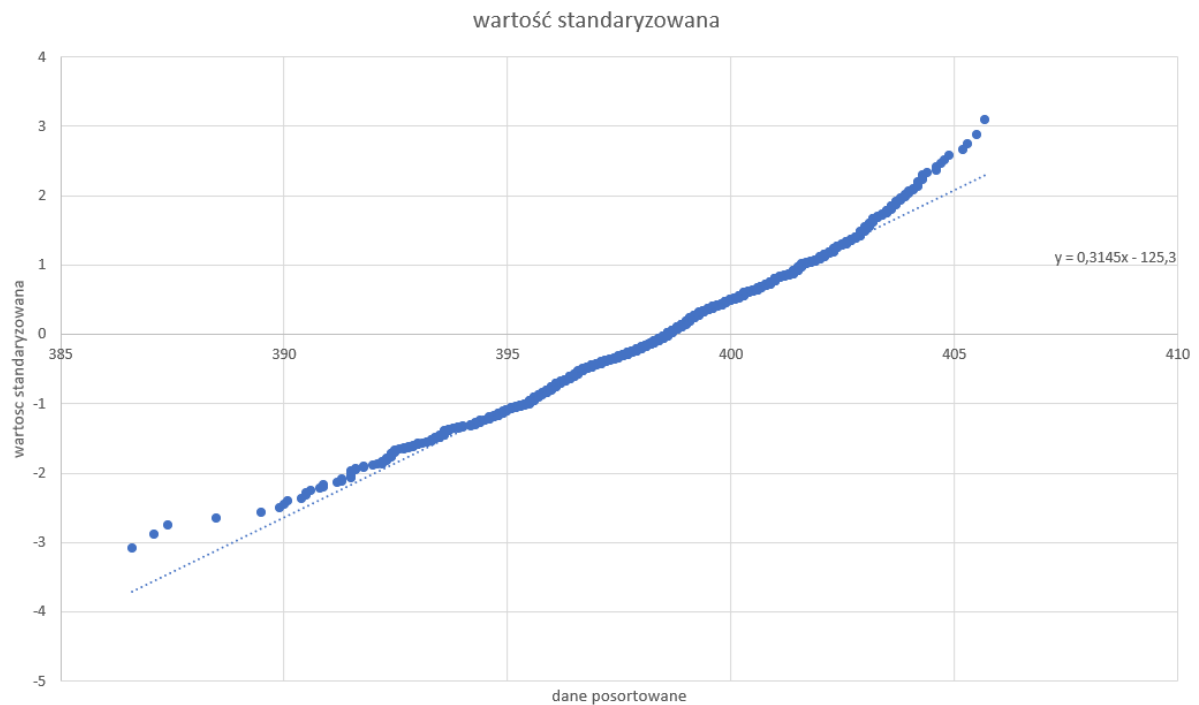
Rys.7.



Rys.7. Przedstawia wykres prawdopodobieństwa skumulowanego.

Następnie Wyznaczyłem wartości standaryzowane dla wartości wektora zgodnie z procedurą przedstawioną na laboratoriach. Wyznaczyłem również wykres dla wartości standaryzowanych jak i prostą regresji która była później potrzeba do wyznaczenia parametry rozkładu Gaussa.

Rys.8.

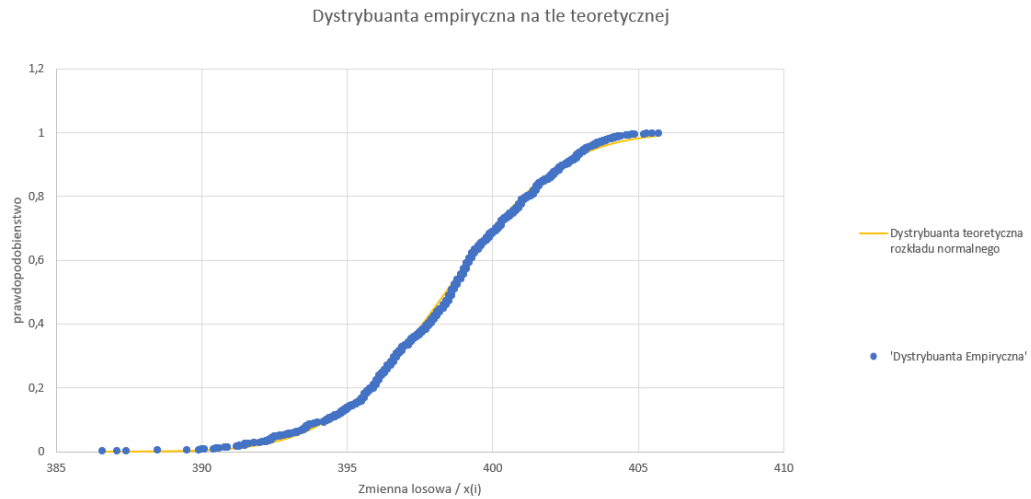


Rys.8. przedstawia wykres wartości standaryzowanych.

	metoda graficzna	metoda punktowa
μ	0,0025	3,15
σ	-0,0080	398,39

Pod koniec wyznaczyłem dystrybuantę teoretyczną jak i przedstawiłem te dystrybuanty na wykresie.

Rys.9.



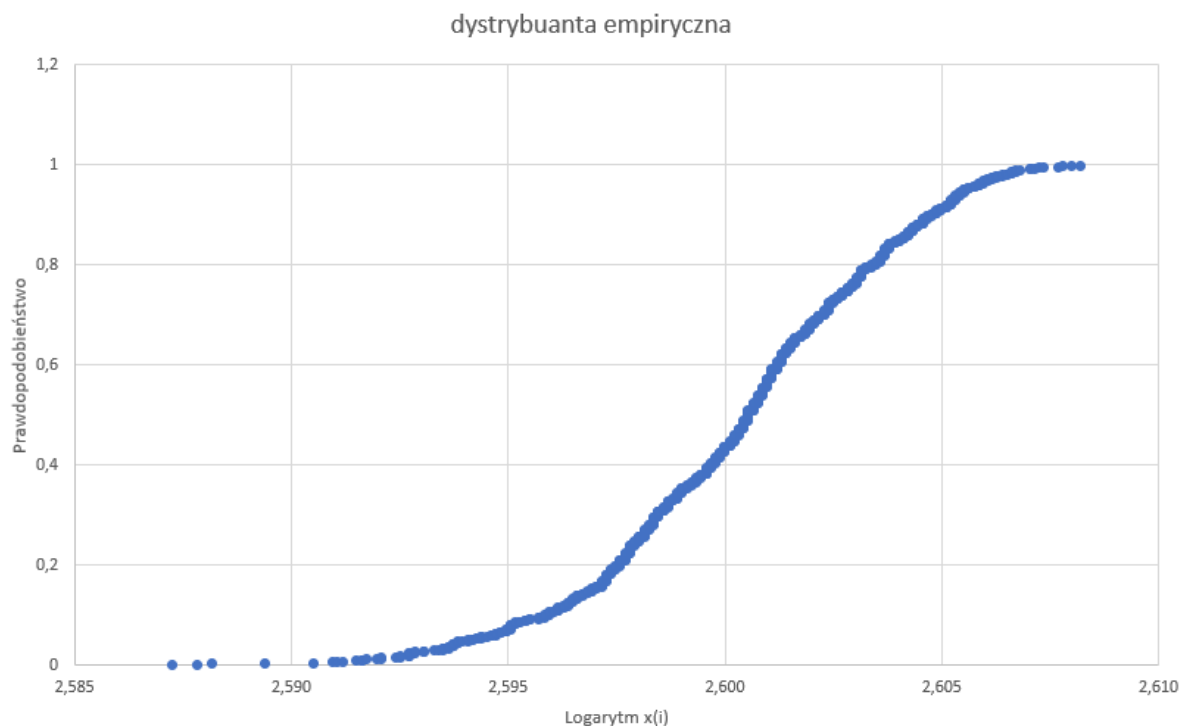
Rys.9. Przedstawia wykres dystrybuanty empirycznej na tle teoretycznej.

6. Rozkład logarytm normalny

Początkowe kroki w rozwiązaniu tego zadania były tożsame jak kroki w zadaniu **nr.5 Rozkład Normalny**. Tym razem celem zadania było wyznaczenie i przedstawienie na wykresie dystrybuanty empirycznej na tle teoretycznej logarytm.

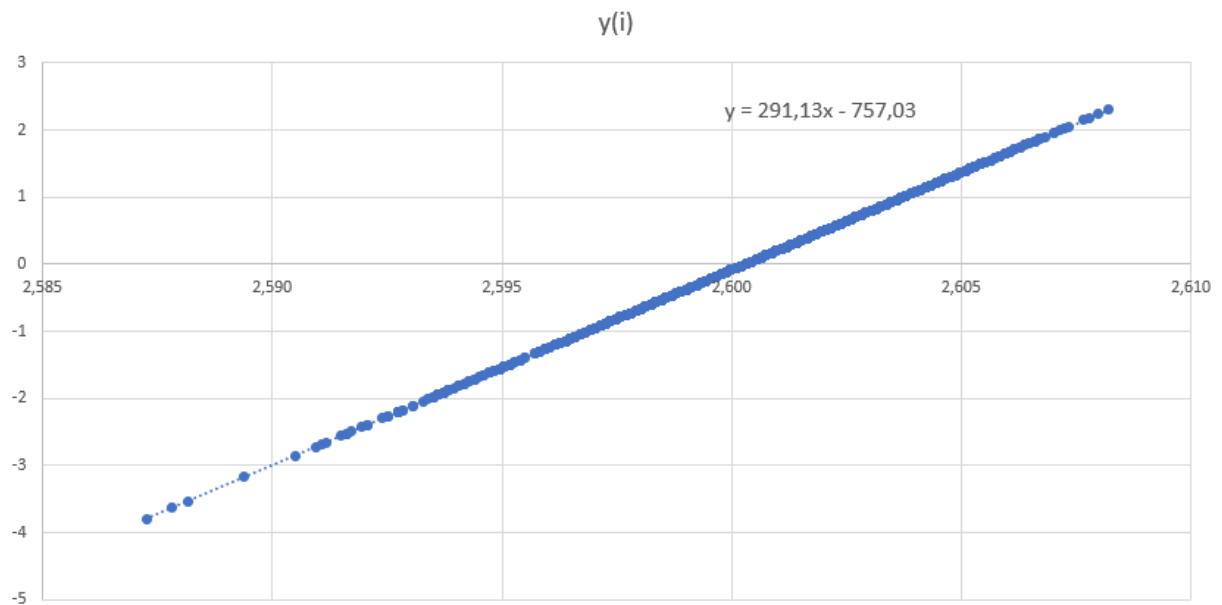
I.p	dane wejściowe	dane posortowane	logarytm naturalny	wart standaryzowana	DANE	y(i)	dystrybuanta teoretyczna
1	399,3	386,6	2,587	0,000999001		-3,79317	0,0000743691
2	399,3	387,1	2,588	0,001998002	Średnia	-3,62975	0,0001418492
3	399,4	387,4	2,588	0,002997003	2,60029084	-3,5318	0,0002063726
4	400,1	388,5	2,589	0,003996004	Mediana	-3,17329	0,0007535976
5	399,8	389,5	2,591	0,004995005	2,601	-2,84826	0,0021979324
6	400,5	389,9	2,591	0,005994006	Odch.stand	-2,71848	0,0032790983
7	398,6	390	2,591	0,006993007	0,003434858	-2,68606	0,0036150128
8	398,5	390,1	2,591	0,007992008	MIN	-2,65364	0,0039813939
9	398,9	390,4	2,592	0,008991009	2,587	-2,55645	0,0052873684
10	399,6	390,5	2,592	0,00999001	MAX	-2,52406	0,0058003420
11	402,9	390,5	2,592	0,010989011	2,608	-2,52406	0,0058003420
12	399,2	390,6	2,592	0,011988012	ilość próbek	-2,49169	0,0063568526
13	397,7	390,8	2,592	0,012987013	1000	-2,42697	0,0076128374
14	391,5	390,9	2,592	0,013986014		-2,39462	0,0083188706
15	396,4	390,9	2,592	0,014985015	σ punktowa	-2,39462	0,0083188706
16	397,8	391,2	2,592	0,015984016	0,003435	-2,29762	0,0107917571
17	398,1	391,3	2,593	0,016983017	μ punktowa	-2,2653	0,0117470708
18	398,3	391,3	2,593	0,017982018	2,6003	-2,2653	0,0117470708
19	398,6	391,5	2,593	0,018981019	σ graficzna	-2,20069	0,0138788318
20	399,1	391,5	2,593	0,01998002	0,003435	-2,20069	0,0138788318
21	399,2	391,5	2,593	0,020979021	μ graficzna	-2,20069	0,0138788318
22	399	391,5	2,593	0,021978022	2,6003	-2,20069	0,0138788318
23	403,3	391,5	2,593	0,022977023		-2,20069	0,0138788318
24	399,8	391,5	2,593	0,023976024		-2,20069	0,0138788318
25	399,9	391,6	2,593	0,024975025		-2,1684	0,0150640219
26	401,3	391,6	2,593	0,025974026		-2,1684	0,0150640219
27	401,2	391,8	2,593	0,026973027		-2,10384	0,0176960008
28	399,5	391,8	2,593	0,027972028		-2,10384	0,0176960008
29	400,2	392	2,593	0,028971029		-2,03932	0,0207090853
30	397,3	392,1	2,593	0,02997003		-2,00707	0,0223711595

Rys.9.



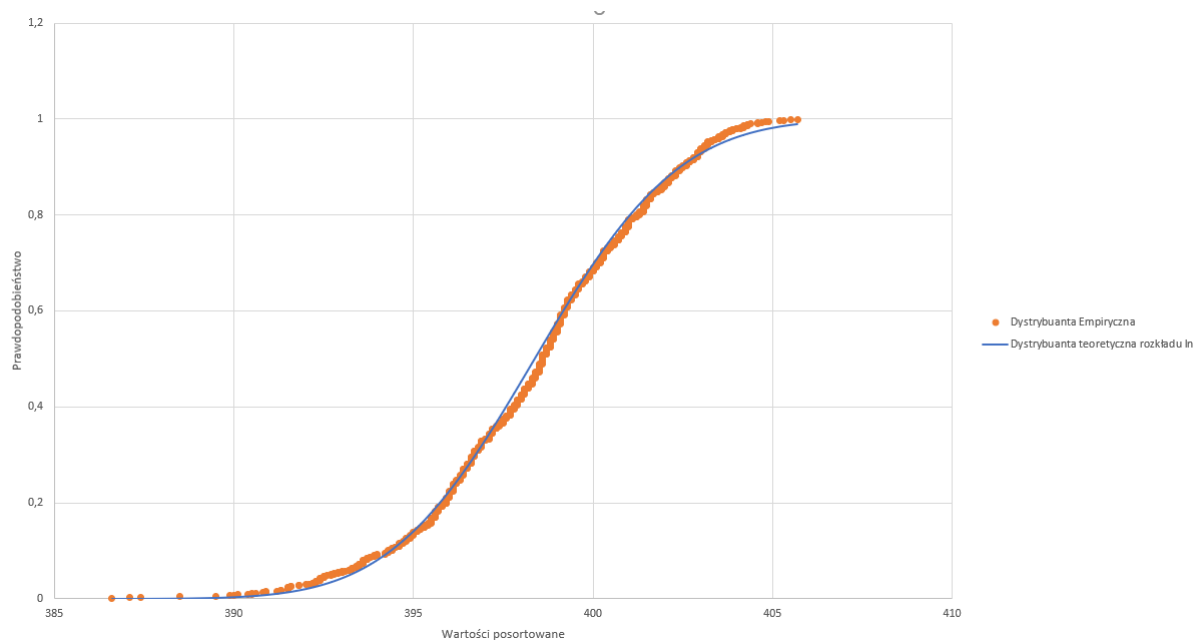
Rys.9. Przedstawia wykres dystrybuanty empirycznej.

Rys.10.



Rys.10. Przedstawia wykres $y(i)$

Rys.11.



Rys.11. Przedstawia dystrybuantę empiryczną na tle dystrybuanty teoretycznej rozkładu ln.