

Statystyka i teoria obsługi masowej
Projekt - Statistica
Dawid Bitner

Na wybranym zbiorze (zbiorach) przeprowadź analizę statystyczną. Powinna ona zawierać następujące elementy:

- a) Weryfikacja hipotez statystycznych (E4)
- b) Regresja wieloraka (E5)

W części dotyczącej hipotez należy uwzględnić przynajmniej jeden test parametryczny i jeden nieparametryczny.

- a) Weryfikacja hipotez statystycznych

TESTY PARAMETRYCZNE - TESTY ISTOTNOŚCI

Poniższa tabela zawiera informacje dotyczące pewnej populacji zółwi. Ich długość, szerokość, wysokość z podziałem na płeć.

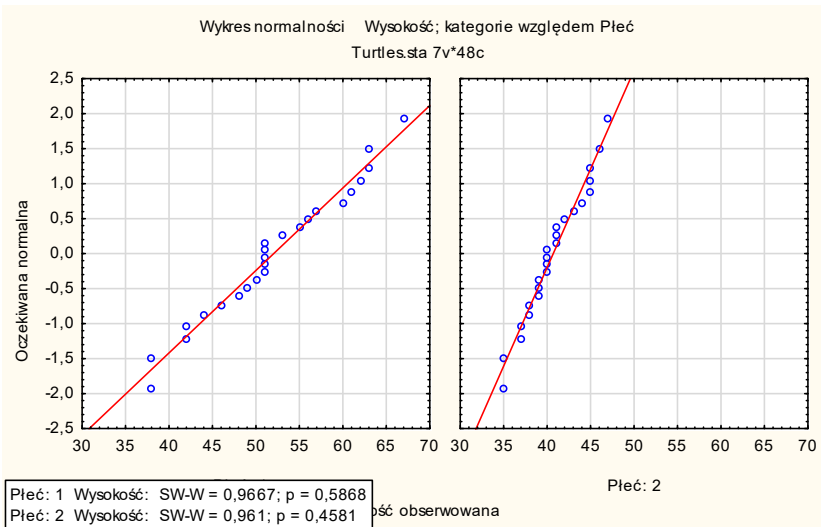
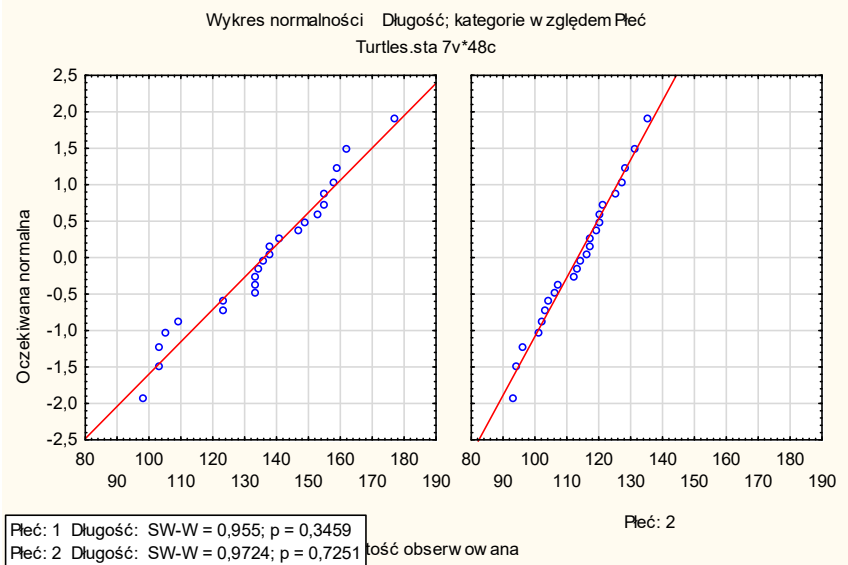
	1	2	3	4	5	6	7
	Długość	Szerokość	Wysokość	DługośćLo	SzerokośćLo	WysokośćLo	Płeć
1	98	81	38	4,58496748	4,39444915	3,63758616	1
2	103	84	38	4,63472899	4,4308168	3,63758616	1
3	103	86	42	4,63472899	4,4543473	3,73766962	1
4	105	86	42	4,65396035	4,4543473	3,73766962	1
5	109	88	44	4,69134788	4,47733681	3,78418963	1
6	123	92	50	4,81218436	4,52178858	3,91202301	1
7	123	95	46	4,81218436	4,55387689	3,8286414	1
8	133	99	51	4,89034913	4,59511985	3,93182563	1
9	133	102	51	4,89034913	4,62497281	3,93182563	1
10	133	102	51	4,89034913	4,62497281	3,93182563	1
11	134	100	48	4,8978398	4,60517019	3,87120101	1
12	136	102	49	4,91265489	4,62497281	3,8918203	1
13	138	98	51	4,92725369	4,58496748	3,93182563	1
14	138	99	51	4,92725369	4,59511985	3,93182563	1
15	141	105	53	4,94875989	4,65396035	3,97029191	1
16	147	108	57	4,99043259	4,68213123	4,04305127	1
17	149	107	55	5,00394631	4,67282883	4,00733319	1
18	153	107	56	5,03043792	4,67282883	4,02535169	1
19	155	115	63	5,04342512	4,74493213	4,14313473	1
20	155	117	60	5,04342512	4,76217393	4,09434456	1
21	158	115	62	5,06259503	4,74493213	4,12713439	1
22	159	118	63	5,0689042	4,77068462	4,14313473	1
23	162	124	61	5,08759634	4,82028157	4,11087386	1
24	177	132	67	5,17614973	4,88280192	4,20469262	1
25	93	74	37	4,53259949	4,30406509	3,61091791	2
26	94	78	35	4,54329478	4,35670883	3,55534806	2
27	96	80	35	4,56434819	4,38202663	3,55534806	2
28	101	84	39	4,61512052	4,4308168	3,66356165	2
29	102	85	38	4,62497281	4,44265126	3,63758616	2
30	103	81	37	4,63472899	4,39444915	3,61091791	2
31	104	83	39	4,6443909	4,41884061	3,66356165	2
32	106	83	39	4,66343909	4,41884061	3,66356165	2
33	107	82	38	4,67282883	4,40671925	3,63758616	2
34	112	89	40	4,71849887	4,48863637	3,68887945	2
35	113	88	40	4,72738782	4,47733681	3,68887945	2
36	114	86	40	4,73619845	4,4543473	3,68887945	2
37	116	90	43	4,75359019	4,49980967	3,76120012	2
38	117	90	41	4,76217393	4,49980967	3,71357207	2
39	117	91	41	4,76217393	4,51085951	3,71357207	2
40	119	93	41	4,77912349	4,53259949	3,71357207	2
41	120	89	40	4,78749174	4,48863637	3,68887945	2
42	120	93	44	4,78749174	4,53259949	3,78418963	2
43	121	95	42	4,79579055	4,55387689	3,73766962	2
44	125	93	45	4,82831374	4,53259949	3,80666249	2
45	127	96	45	4,84418709	4,56434819	3,80666249	2
46	128	95	45	4,85203026	4,55387689	3,80666249	2

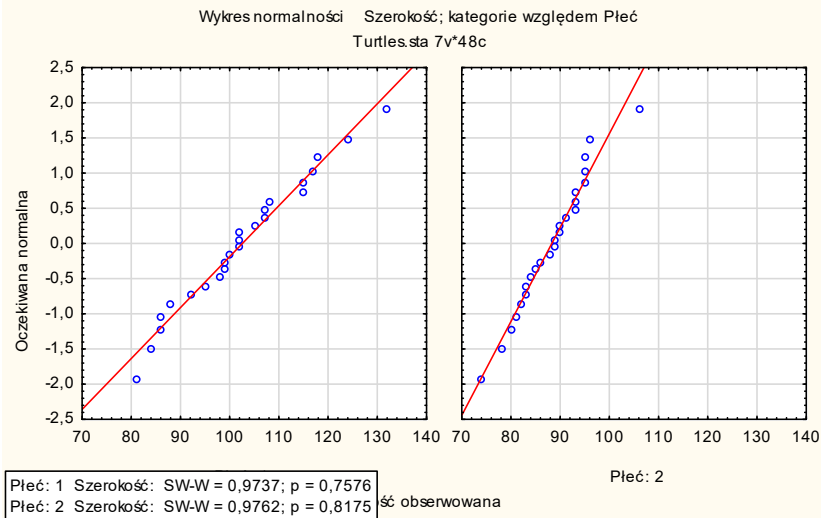
	1	2	3	4	5	6	7
	Długość	Szerokość	Wysokość	DługośćLo	SzerokośćLo	WysokośćLo	Płeć
47	131	95	46	4,87519732	4,55387689	3,8286414	2
48	135	106	47	4,90527478	4,66343909	3,8501476	2

Stawiamy hipotezę H_0 :

Na poziomie $\alpha = 0.05$ zweryfikuj hipotezę o równości średnich wartości zmiennych długość, szerokość i wysokość uwzględniając podział względem płci.

Testy parametryczne wymagają aby rozkład badanych zmiennych był rozkładem normalnym. Wobec tego, przed przystąpieniem do wykonania testu należy przeprowadzić test normalności.





Wykresy normalności zostały wykonane z testem Shapiro-Wilka. W porównaniu do α , p jest duże. Wobec tego możemy przyjąć, że rozkłady są normalne, a więc badając hipotezę możemy użyć testu parametrycznego.

Został przeprowadzony test t dla prób niezależnych w grupach z Testem Levena który bada jednorodność wariancji:

Zmienna	Testy t; Grupująca: Płeć (Turtles.sta) Grupa 1: 1 Grupa 2 2						
	Średnia 1	Średnia 2	t	df	p	Nważnych 1	Nważnych 2
Długość	136,0417	113,3750	4,570481	46	0,000037	24	24
Szerokość	102,5833	88,2917	4,701470	46	0,000024	24	24
Wysokość	52,0417	40,7083	6,368935	46	0,000000	24	24

Zmienna	Testy t; Grupująca: Płeć (Turtles.sta) Grupa 1: 1 Grupa 2 2					
	Odch.std 1	Odch.std 2	iloraz F Wariancje	p Wariancje	Levene'a F(1,df)	df Levene'a
Długość	21,24900	11,77991	3,253815	0,006449	5,67476	46
Szerokość	13,10465	7,07401	3,431778	0,004507	6,22315	46
Wysokość	8,04595	3,35545	5,749799	0,000087	12,63285	46

Zmienna	Testy t; Grupująca: Płeć (Turtles.sta) Grupa 1: 1 Grupa 2 2			
	p Levene'a	Średnia 1 - Średnia 2	Ufność -95,000%	Ufność +95,000%
Długość	0,021401	22,66667	12,68399	32,64934
Szerokość	0,016263	14,29167	8,17281	20,41053
Wysokość	0,000890	11,33333	7,75144	14,91522

Dla każdej ze zmiennych jak widać powyżej, założenie o jednorodności wariancji nie jest spełnione, więc musimy przeprowadzić test z niezależną estymacją wariancji.

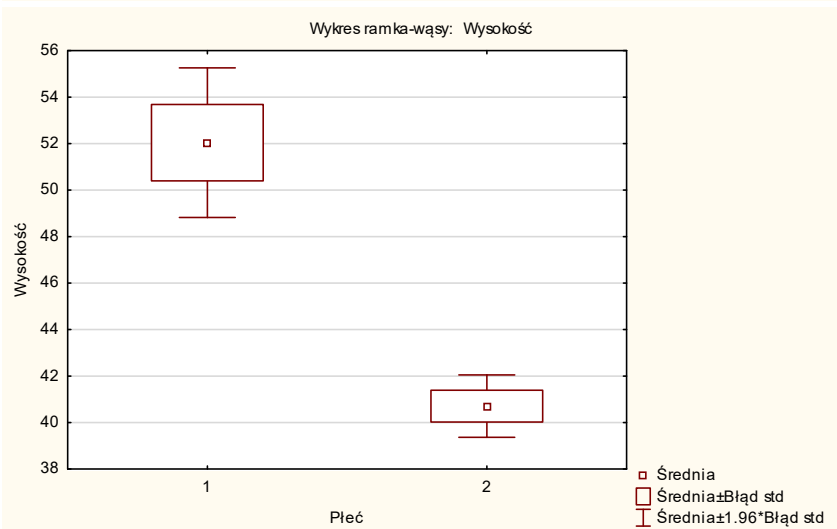
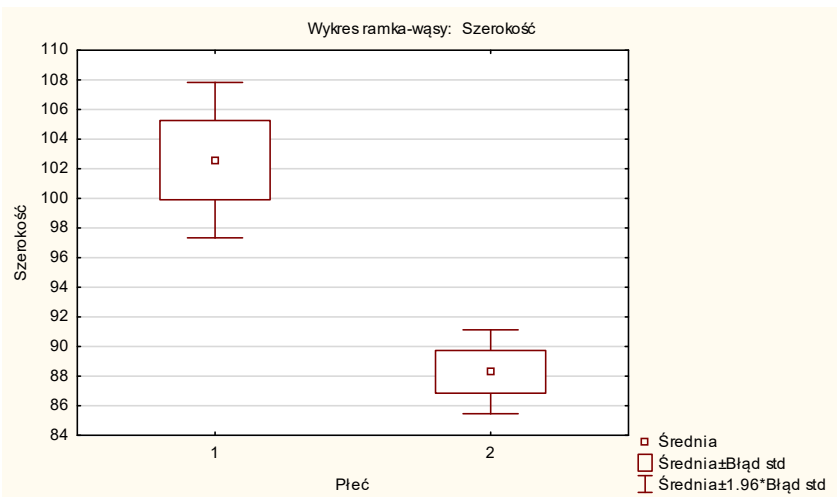
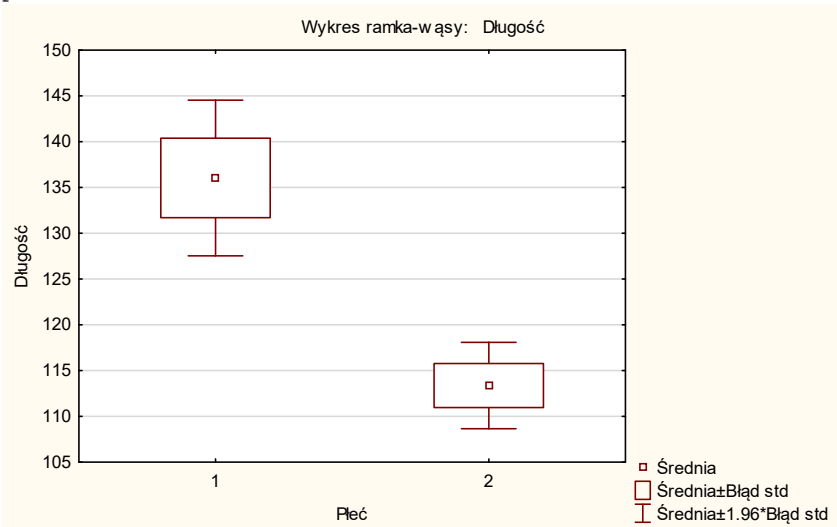
Zmienna	Testy t; Grupująca: Płeć (Turtles.sta) Grupa 1: 1 Grupa 2 2						
	Średnia 1	Średnia 2	t	df	p	t oddz. est.war.	df
Długość	136,0417	113,3750	4,570481	46	0,000037	4,570481	35,91719
Szerokość	102,5833	88,2917	4,701470	46	0,000024	4,701470	35,35506
Wysokość	52,0417	40,7083	6,368935	46	0,000000	6,368935	30,76539

Zmienna	Testy t; Grupująca: Płeć (Turtles.sta) Grupa 1: 1 Grupa 2 2					
	p dwustron	Nważnych 1	Nważnych 2	Odch.std 1	Odch.std 2	iloraz F Wariacje
Długość	0,000056	24	24	21,24900	11,77991	3,253815
Szerokość	0,000039	24	24	13,10465	7,07401	3,431778
Wysokość	0,000000	24	24	8,04595	3,35545	5,749799

Zmienna	Testy t; Grupująca: Płeć (Turtles.sta) Grupa 1: 1 Grupa 2 2					
	p Wariacje	Levene'a F(1,df)	df Levene'a	p Levene'a	Średnia 1 - Średnia 2	Ufność -95,000%
Długość	0,006449	5,67476	46	0,021401	22,66667	12,60781
Szerokość	0,004507	6,22315	46	0,016263	14,29167	8,12270
Wysokość	0,000087	12,63285	46	0,000890	11,33333	7,70296

Zmienna	Testy t; Grupująca: Płeć
	Ufność +95,000%
Długość	32,72552
Szerokość	20,46063
Wysokość	14,96371

Graficzna ilustracja różnic pomiędzy trzema średnimi z uwzględnieniem podziału na płcie:



Średnia długość osobników w tej populacji z podziałem na płcie nie jest równa, tak

samo jak wysokość i szerokość.

Hipoteza zerowa jest nieprawdziwa. Można zauważyć powyższych tabel i wykresów ramkawy, że osobniki Płci 1 są przeciętnie wyższe, szersze i dłuższe.

TESTY NIEPARAMETRYCZNE:

Poniższa tabela zawiera informacje dotyczące szkód wyrządzonych przez posiadaczy samochodów. Dane które zawiera tabela to: Wiek właściciela, typ pojazdu, wiek pojazdu i wartość wyrządzonej szkody.

	1 WiekWłaściciela	2 TypPojazdu	3 WiekPojazdu	4 WartośćSzkody
1	17-20	A	0-3	289
2	17-20	A	4-7	282
3	17-20	A	8-9	133
4	17-20	A	10+	160
5	17-20	B	0-3	372
6	17-20	B	4-7	249
7	17-20	B	8-9	288
8	17-20	B	10+	11
9	17-20	C	0-3	189
10	17-20	C	4-7	288
11	17-20	C	8-9	179
12	17-20	C	10+	0
13	17-20	D	0-3	763
14	17-20	D	4-7	850
15	17-20	D	8-9	0
16	17-20	D	10+	0
17	21-24	A	0-3	302
18	21-24	A	4-7	194
19	21-24	A	8-9	135
20	21-24	A	10+	166
21	21-24	B	0-3	420
22	21-24	B	4-7	243
23	21-24	B	8-9	196
24	21-24	B	10+	135
25	21-24	C	0-3	268
26	21-24	C	4-7	343
27	21-24	C	8-9	293
28	21-24	C	10+	104
29	21-24	D	0-3	407
30	21-24	D	4-7	320
31	21-24	D	8-9	205
32	21-24	D	10+	0
33	25-29	A	0-3	268
34	25-29	A	4-7	285
35	25-29	A	8-9	181
36	25-29	A	10+	110
37	25-29	B	0-3	275
38	25-29	B	4-7	243
39	25-29	B	8-9	179
40	25-29	B	10+	264
41	25-29	C	0-3	334
42	25-29	C	4-7	274
43	25-29	C	8-9	208
44	25-29	C	10+	150
45	25-29	D	0-3	383

	1 WiekWłaściciela	2 TypPojazdu	3 WiekPojazdu	4 WartośćSzkody
46	25-29	D	4-7	305
47	25-29	D	8-9	116
48	25-29	D	10+	636
49	30-34	A	0-3	236
50	30-34	A	4-7	270
51	30-34	A	8-9	160
52	30-34	A	10+	110
53	30-34	B	0-3	259
54	30-34	B	4-7	226
55	30-34	B	8-9	161
56	30-34	B	10+	107
57	30-34	C	0-3	340
58	30-34	C	4-7	260
59	30-34	C	8-9	189
60	30-34	C	10+	104
61	30-34	D	0-3	400
62	30-34	D	4-7	349
63	30-34	D	8-9	147
64	30-34	D	10+	65
65	35-39	A	0-3	207
66	35-39	A	4-7	129
67	35-39	A	8-9	157
68	35-39	A	10+	113
69	35-39	B	0-3	208
70	35-39	B	4-7	214
71	35-39	B	8-9	149
72	35-39	B	10+	137
73	35-39	C	0-3	251
74	35-39	C	4-7	232
75	35-39	C	8-9	204
76	35-39	C	10+	141
77	35-39	D	0-3	233
78	35-39	D	4-7	325
79	35-39	D	8-9	207
80	35-39	D	10+	0
81	40-49	A	0-3	254
82	40-49	A	4-7	213
83	40-49	A	8-9	149
84	40-49	A	10+	98
85	40-49	B	0-3	218
86	40-49	B	4-7	209
87	40-49	B	8-9	172
88	40-49	B	10+	110
89	40-49	C	0-3	239
90	40-49	C	4-7	250

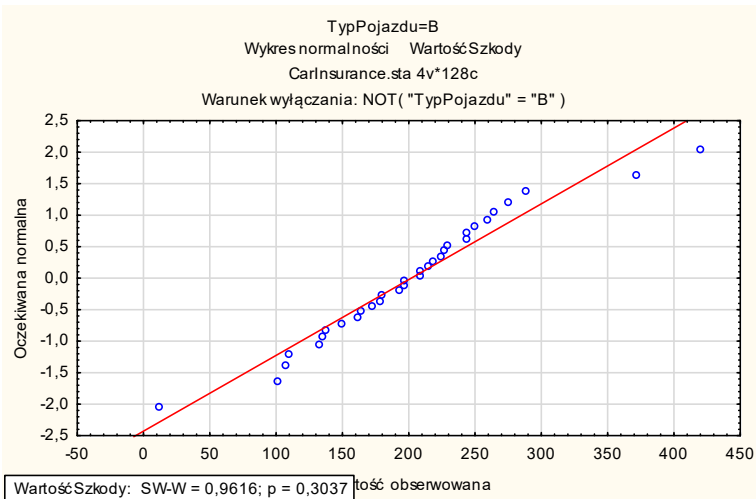
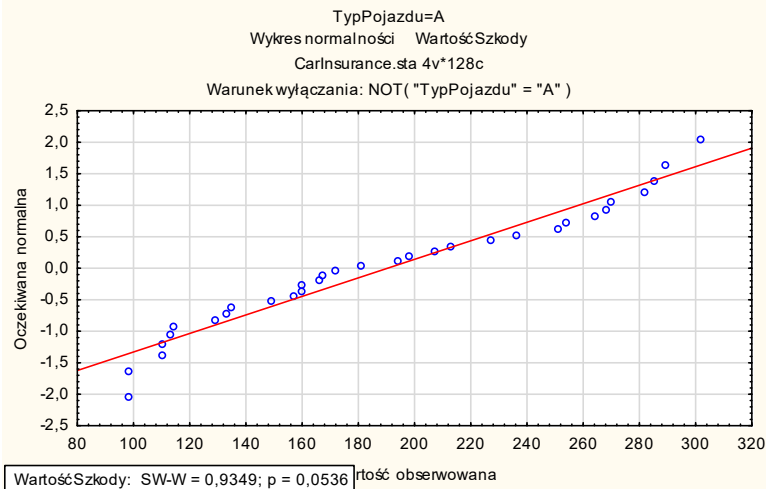
	1 WiekWłaściciela	2 TypPojazdu	3 WiekPojazdu	4 WartośćSzkody
91	40-49	C	8-9	174
92	40-49	C	10+	129
93	40-49	D	0-3	387
94	40-49	D	4-7	299
95	40-49	D	8-9	325
96	40-49	D	10+	137
97	50-59	A	0-3	251
98	50-59	A	4-7	227
99	50-59	A	8-9	172
100	50-59	A	10+	98
101	50-59	B	0-3	196
102	50-59	B	4-7	229
103	50-59	B	8-9	164
104	50-59	B	10+	132
105	50-59	C	0-3	268
106	50-59	C	4-7	250
107	50-59	C	8-9	175
108	50-59	C	10+	152
109	50-59	D	0-3	391
110	50-59	D	4-7	228
111	50-59	D	8-9	346
112	50-59	D	10+	167
113	60+	A	0-3	264
114	60+	A	4-7	198
115	60+	A	8-9	167
116	60+	A	10+	114
117	60+	B	0-3	224
118	60+	B	4-7	193
119	60+	B	8-9	178
120	60+	B	10+	101
121	60+	C	0-3	269
122	60+	C	4-7	258
123	60+	C	8-9	227
124	60+	C	10+	119
125	60+	D	0-3	385
126	60+	D	4-7	324
127	60+	D	8-9	192
128	60+	D	10+	123

Stawiamy hipotezę H_0 :

Pomiędzy typem pojazdu A a B nie występuje znaczna różnica w wyrządzonej na nim szkodzie. Poziom istotności: $\alpha = 0.05$.

Musimy przeprowadzić testy dla prób niezależnych. Zanim przejdziemy do wykonywania testu nieparametrycznego, sprawdzimy najpierw, czy przypadkiem nie zostały spełnione założenia testu parametrycznego.

Poniżej przedstawiony został skategoryzowany wykres normalności dla dwóch grup wraz z testem Shapiro-Wilka:



Nie możemy w tym przypadku mówić o normalności rozkładów, ponieważ w obu przypadkach $p > \alpha$. Musimy więc przeprowadzić testy nieparametryczne.

Przeprowadzamy kolejno test Walda-Wolfowitza, test Kołomogorowa-Smirnowa i test U Manna-Withneya:

Zmienna	Test serii Walda-Wolfowitza (CarInsurance.sta) Względem zmiennej: TypPojazdu Zaznaczone wyniki są istotne z $p < 0,05000$					
	N ważn. A	N ważn. B	Średnia A	Średnia B	Z	p
WartośćSzkody	32	32	190,3750	201,9375	1,008032	0,313440

Zmienna	Test serii Walda-Wolfowitza (CarInsurance.sta) Względem zmiennej: TypPojazdu Zaznaczone wyniki są istotne z $p < 0,05000$			
	Z skoryg.	p	Liczba serii	Liczba wiązanych
WartośćSzkody	0,882028	0,377762	37	5

Przy założonym poziomie istotności nie mamy podstaw do odrzucenia hipotezy zerowej.

Zmienna	Test Kołmogorowa-Smirnowa (CarInsurance.sta) Względem zmiennej: TypPojazdu Zaznaczone wyniki są istotne z $p < ,05000$					
	Maks.uj. Różnica	Maks.dod Różnica	p	Średnia A	Średnia B	Odch.std A
WartośćSzkody	-0,156250	0,093750	$p > .10$	190,3750	201,9375	64,13129

Zmienna	Test Kołmogorowa-Smirnowa (CarInsurance.sta) Względem zmiennej: TypPojazdu		
	Odch.std B	N ważn. A	N ważn. B
WartośćSzkody	78,28069	32	32

$p > 0.1$, więc i większe od 0.05 - nie mamy podstaw do odrzucenia hipotezy zerowej, która mówi o tym, że wartość szkód w obu typach pojazdów jest jednakowa.

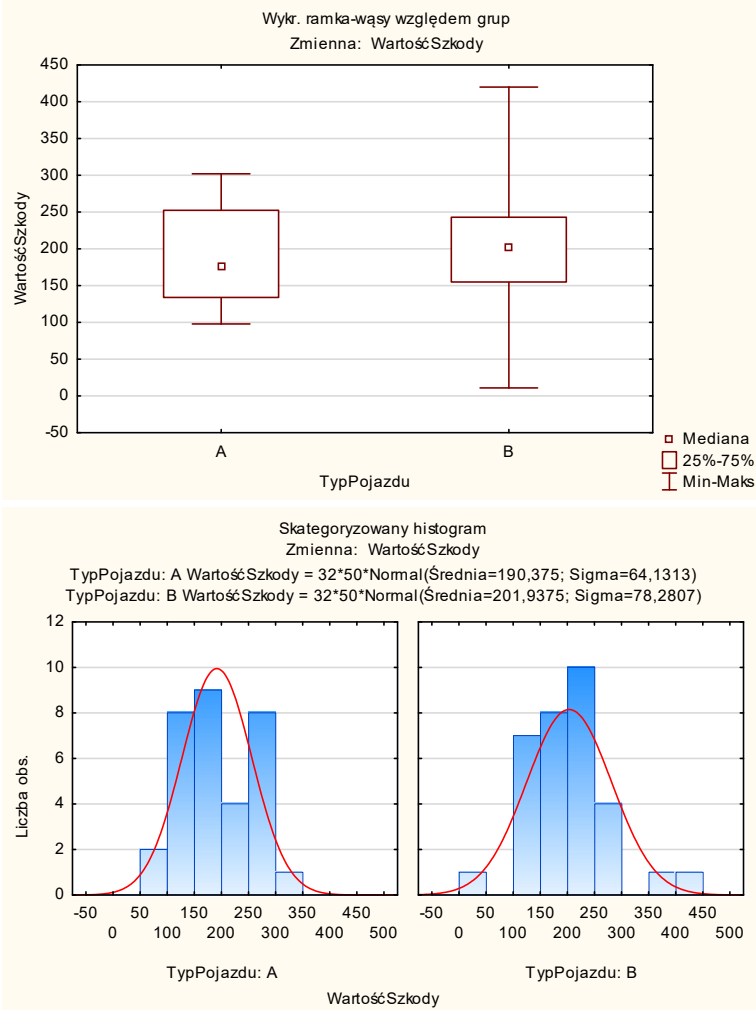
Zmienna	Test U Manna-Whitneya (CarInsurance.sta) Względem zmiennej: TypPojazdu Zaznaczone wyniki są istotne z $p < 0,05000$					
	Sum.rang A	Sum.rang B	U	Z	p	Z popraw.
WartośćSzkody	997,0000	1083,000	469,0000	-0,577368	0,563692	-0,577447

Zmienna	Test U Manna-Whitneya (CarInsurance.sta) Względem zmiennej: TypPojazdu Zaznaczone wyniki są istotne z $p < 0,05000$			
	p	N ważn. A	N ważn. B	2*1str. dokł. p
WartośćSzkody	0,563638	32	32	0,570687

Dla prób zawierających więcej niż 20 pomiarów możemy wykorzystać statystykę Z - podstawową wersję testu. Rangi się nie powtarzają więc Z poprawione jest takie samo.

Tutaj również $p > \alpha$ więc nie mamy podstaw do odrzucenia naszej hipotezy zerowej.

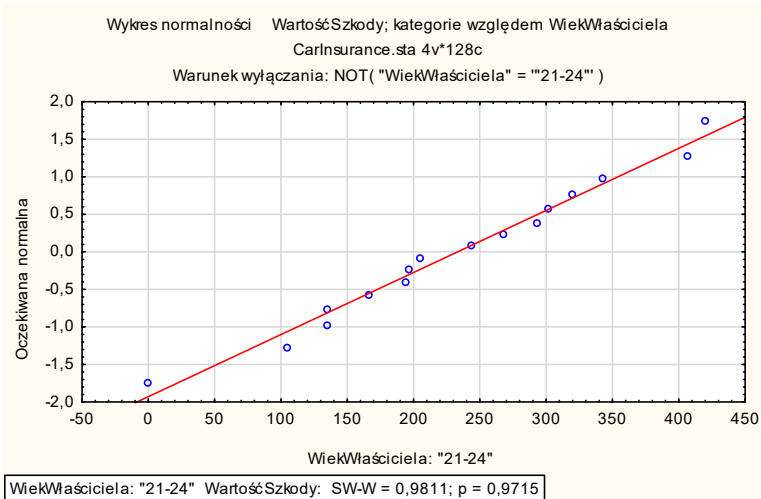
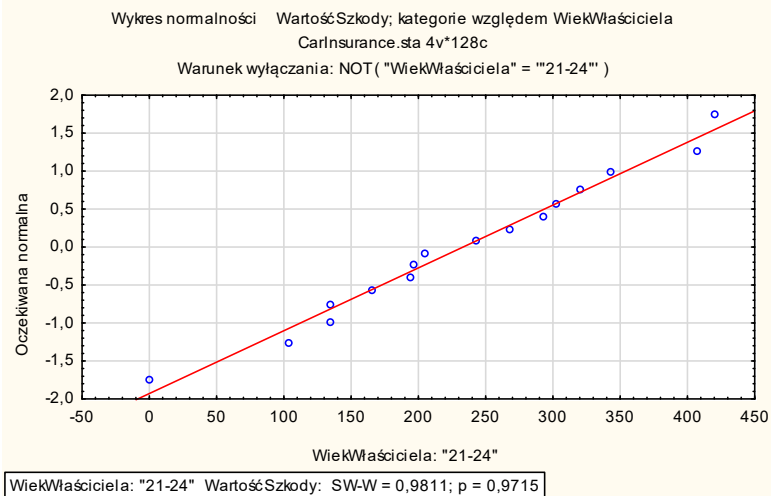
Poniżej przedstawiono wykres ramka-wąsy, oraz skategoryzowany histogram potwierdzające prawdziwość hipotezy zerowej:



Wykonujemy drugi test nieparametryczny i stawiamy nową hipotezę:

Hipoteza zerowa: Nie istnieje znacząca różnica w wartości szkód przez sprawców w wieku: 21-24 i w wieku 35-39. Poziom istotności: $\alpha = 0.05$.

Ponownie tworzymy wykresy normalności, dla nowych zmiennych:



Nie możemy w tym przypadku mówić o normalności rozkładów, ponieważ w obu przypadkach $p > \alpha$. Musimy więc przeprowadzić testy nieparametryczne.

Przeprowadzamy kolejno test Walda-Wolfowitza, test Kołomogorowa-Smirnowa i test U Manna-Withneya:

Zmienna	Test serii Walda-Wolfowitza (CarlInsurance.sta) Względem zmiennej: WiekWłaściciela Zaznaczone wyniki są istotne z $p < 0,05000$					
	N ważn. "21-24"	N ważn. "35-39"	Średnia "21-24"	Średnia "35-39"	Z	p
WartośćSzkody	16	16	233,1875	181,6875	-1,07819	0,280948

Zmienna	Test serii Walda-Wolfowitza (CarlInsurance.sta) Względem zmiennej: WiekWłaściciela Zaznaczone wyniki są istotne z $p < 0,05000$			
	Z skoryg.	p	Liczba serii	Liczba wiązaných
WartośćSzkody	0,898494	0,368923	14	1

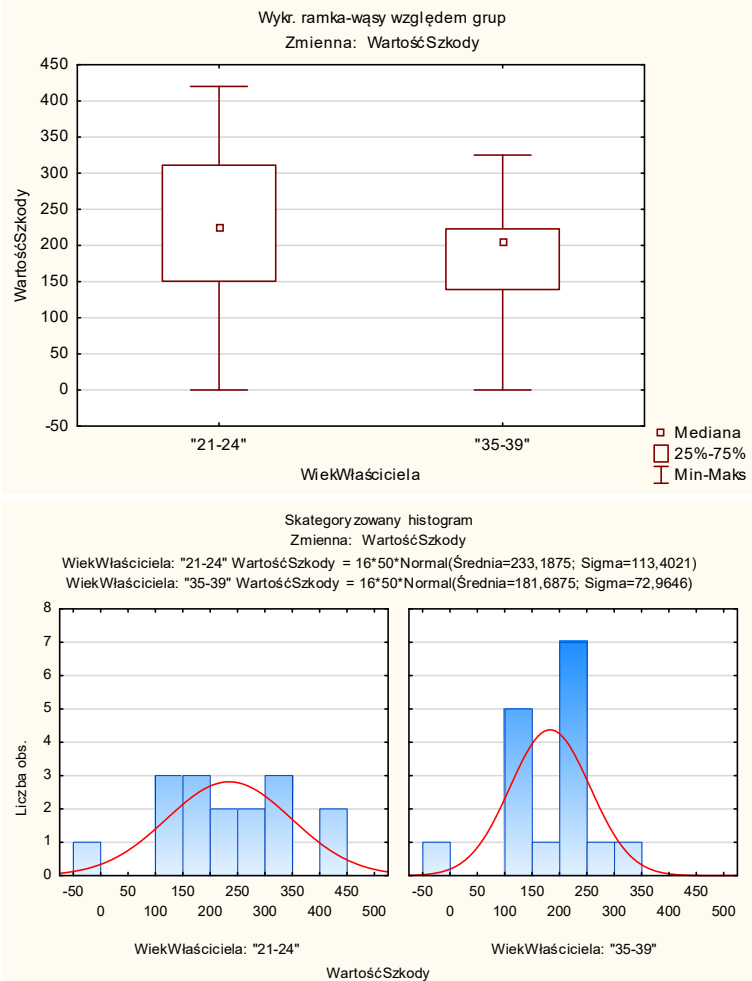
Zmienna	Test Kołmogorowa-Smirnowa (CarInsurance.sta) Względem zmiennej: WiekWłaściciela Zaznaczone wyniki są istotne z $p < ,05000$					
	Maks.uj. Różnica	Maks.dod Różnica	p	Średnia "21-24"	Średnia "35-39"	Odch.std "21-24"
WartośćSzkody	-0,062500	0,375000	$p > .10$	233,1875	181,6875	113,4021

Zmienna	Test Kołmogorowa-Smirnowa (CarInsurance.sta) Względem zmiennej:		
	Odch.std "35-39"	N ważn. "21-24"	N ważn. "35-39"
WartośćSzkody	72,96457	16	16

Zmienna	Test U Manna-Whitneya (CarInsurance.sta) Względem zmiennej: WiekWłaściciela Zaznaczone wyniki są istotne z $p < 0,05000$					
	Sum.rang "21-24"	Sum.rang "35-39"	U	Z	p	Z popraw.
WartośćSzkody	294,5000	233,5000	97,50000	1,149512	0,250346	1,149828

Zmienna	Test U Manna-Whitneya (CarInsurance.sta) Względem zmiennej: WiekWłaściciela Zaznaczone wyniki są istotne z $p < 0,05000$			
	p	N ważn. "21-24"	N ważn. "35-39"	2*1str. dokł. p
WartośćSzkody	0,250216	16	16	0,254227

Tutaj również $p > \alpha$ więc nie mamy podstaw do odrzucenia naszej hipotezy zerowej. Poniżej przedstawiono wykres ramka-wąsy, oraz skategoryzowany histogram potwierdzające prawdziwość hipotezy zerowej:



KOLERACJA

Poniższa tabela zawiera informacje dotyczące zwycięstw w meczach baseballowych i danych dotyczących meczu. Przeprowadzone zostaną badania dotyczące koleracji pomiędzy poszczególnymi współczynnikami występującymi w tej grze.

Zwycięstwa w meczach bejsbolowych (Reichler, 1985)							
	1 ROK	2 ZWYCIĘSTWA A	3 RUNS	4 BA	5 DP	6 WALK	7 SO
1	1965	0,599	608	0,245	135	425	1079
2	1965	0,586	682	0,252	124	408	1060
3	1965	0,556	675	0,265	189	469	882
4	1965	0,549	825	0,273	142	587	1113
5	1965	0,531	708	0,256	145	541	996
6	1965	0,528	654	0,25	153	466	1071
7	1965	0,497	707	0,254	152	467	916
8	1965	0,444	635	0,238	166	481	855
9	1965	0,401	569	0,237	130	388	931
10	1965	0,309	495	0,221	153	498	776
11	1966	0,586	606	0,256	128	356	1064
12	1966	0,578	675	0,248	131	359	973
13	1966	0,568	759	0,279	215	463	898
14	1966	0,537	696	0,258	147	412	928
15	1966	0,525	782	0,263	139	485	884
16	1966	0,512	571	0,251	166	448	892
17	1966	0,475	692	0,26	133	490	1043
18	1966	0,444	612	0,255	126	391	929
19	1966	0,41	587	0,239	171	521	773
20	1966	0,364	644	0,254	132	479	908
21	1967	0,627	695	0,263	127	431	956
22	1967	0,562	652	0,245	149	453	990
23	1967	0,54	702	0,251	143	463	888
24	1967	0,537	604	0,248	124	498	1065
25	1967	0,506	612	0,242	174	403	967
26	1967	0,5	679	0,277	186	561	820
27	1967	0,475	631	0,24	148	449	862
28	1967	0,451	519	0,236	144	393	967
29	1967	0,426	626	0,249	120	485	1060
30	1967	0,377	498	0,238	147	536	893
31	1968	0,599	583	0,249	135	375	971
32	1968	0,543	599	0,239	125	344	942
33	1968	0,519	612	0,242	149	392	894
34	1968	0,512	690	0,273	144	573	963
35	1968	0,5	514	0,252	139	362	871
36	1968	0,494	583	0,252	162	485	897
37	1968	0,469	470	0,23	144	414	994
38	1968	0,469	543	0,233	163	421	935
39	1968	0,451	473	0,228	142	430	1014
40	1968	0,444	510	0,231	129	479	1021

Poniżej została przedstawiona macierz korelacji wszystkich zmiennych dostępnych w arkuszu. Współczynnik istotności pozostał niezmieniony tzn. posiada domyślną wartość $p = 0.05$. Można założyć, na podstawie powyższej tabeli (na podstawie czerwonego koloru czcionki współczynników korelacji), że można doszukać się tutaj paru zmiennych które możliwe, że są ze sobą powiązane.

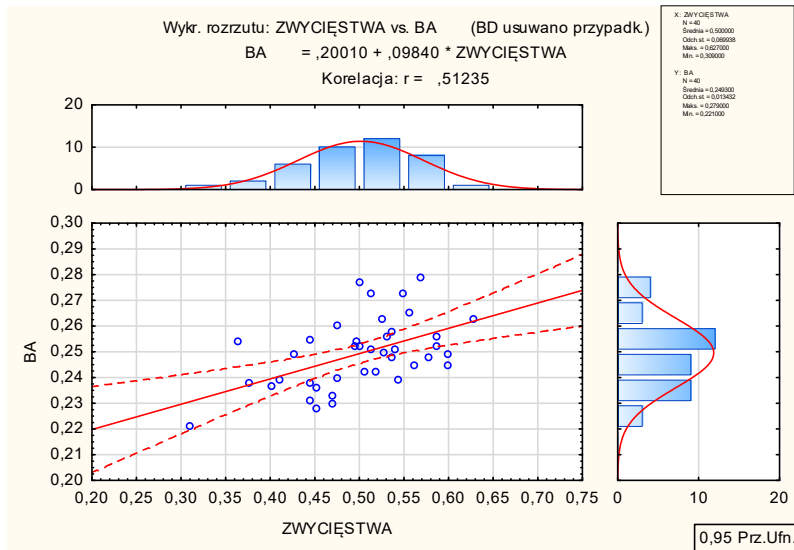
Np.:

a) ZWYCIĘSTWA i BA - wsp. korelacji: ~ 0.51

b) RUNS i BA - wsp. korelacji: ~ 0.80

Zmienna	Korelacje (Baseball.sta) Oznaczone wsp. korelacji są istotne z $p < ,05000$ N=40 (Braki danych usuwano przypadkami)					
	Średnia	Odch.std	ZWYCIĘSTWA	RUNS	BA	DP
ZWYCIĘSTWA	0,5000	0,06994	1,000000	0,510492	0,512347	-0,015229
RUNS	624,4250	83,35647	0,510492	1,000000	0,804049	0,141743
BA	0,2493	0,01343	0,512347	0,804049	1,000000	0,253885
DP	146,7750	20,12587	-0,015229	0,141743	0,253885	1,000000
WALK	452,0250	61,00798	-0,262924	0,341235	0,346465	0,288107
SO	948,5250	82,81459	0,423432	0,136215	0,087632	-0,514820

Zmienna	Korelacje (Baseball.sta) Oznaczone wsp. korelacji są istotne z $p < ,05000$	
	WALK	SO
ZWYCIĘSTWA	-0,262924	0,423432
RUNS	0,341235	0,136215
BA	0,346465	0,087632
DP	0,288107	-0,514820
WALK	1,000000	-0,094018
SO	-0,094018	1,000000



Powyżej widoczny mamy model regresji, oraz współczynnik r korelacji pomiędzy dwoma zmiennymi ZWYCIĘSTWA i BA.

Wyniki regresji wielorakiej dla a):

Zmn. zależ. BA	Wielor. R = ,51234670	F = 13,52536
	R^2 = ,26249914	df = 1,38
Liczba przyp. 40	Popraw. R^2 = ,24309122	p = ,000725
	Błąd standardowy estymacji: ,011685969	
Wyr. wolny ,200100465	Błąd std.: ,0135049	t(38) = 14,817 p = ,0000

ZWYCIĘSTWA b* = ,512

(istotne b* są podświetlone na czerwono)

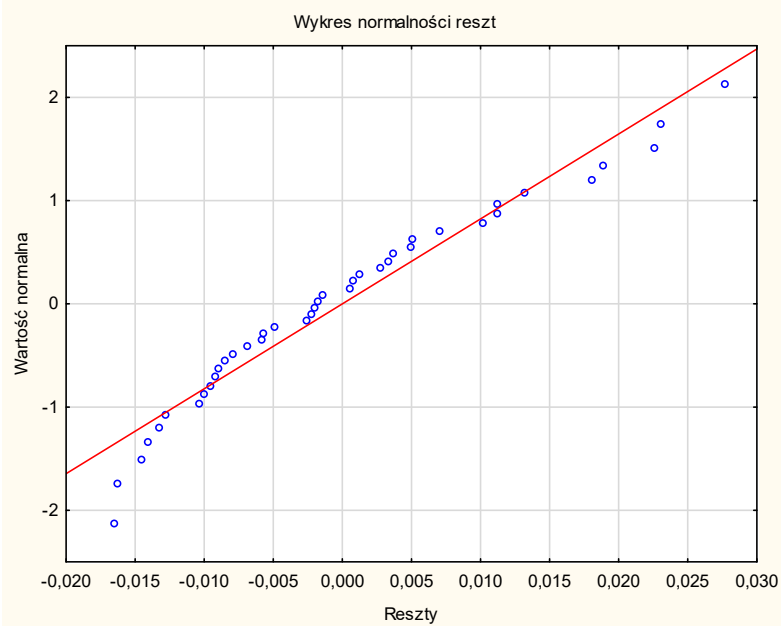
Analiza reszt dla a):

Zmn. zależ. BA	Wielor. R : ,51234670	F = 13,52536
	R^2: ,26249914	df = 1,38
Liczba przyp. 40	Popraw. R^2: ,24309122	p = ,000725
	Błąd standardowy estymacji: ,011685969	
Wyr. wolny ,200100465	Błąd std.: ,0135049	t(38) = 14,817 p < ,0000

	Wartości przewidywane i reszty BA					
	Obserw. Wartość	Przewidyw. Wartość	Reszta	Standard. Przewid.	Standard. Reszta	Bł. std. W.przew.
1	0,245000	0,259042	-0,014042	1,415531	-1,201570	0,003230
2	0,252000	0,257762	-0,005762	1,229653	-0,493096	0,002951
3	0,265000	0,254810	0,010190	0,800704	0,871956	0,002379
4	0,273000	0,254122	0,018878	0,700616	1,615481	0,002266
5	0,256000	0,252350	0,003650	0,443247	0,312311	0,002025
6	0,250000	0,252055	-0,002055	0,400352	-0,175866	0,001994
7	0,254000	0,249005	0,004995	-0,042895	0,427454	0,001849
8	0,238000	0,243790	-0,005790	-0,800704	-0,495436	0,002379
9	0,237000	0,239558	-0,002558	-1,415531	-0,218936	0,003230
10	0,221000	0,230506	-0,009506	-2,730973	-0,813435	0,005434
11	0,256000	0,257762	-0,001762	1,229653	-0,150805	0,002951
12	0,248000	0,256975	-0,008975	1,115267	-0,768025	0,002787
13	0,279000	0,255991	0,023009	0,972284	1,968932	0,002593
14	0,258000	0,252941	0,005059	0,529037	0,432930	0,002096
15	0,263000	0,251760	0,011240	0,357457	0,961840	0,001965
16	0,251000	0,250481	0,000519	0,171579	0,044428	0,001875
17	0,260000	0,246840	0,013160	-0,357457	1,126133	0,001965
18	0,255000	0,243790	0,011210	-0,800704	0,959299	0,002379
19	0,239000	0,240444	-0,001444	-1,286846	-0,123574	0,003035
20	0,254000	0,235918	0,018082	-1,944567	1,547349	0,004081
21	0,263000	0,261797	0,001203	1,815883	0,102972	0,003868
22	0,245000	0,255401	-0,010401	0,886494	-0,890020	0,002483
23	0,251000	0,253236	-0,002236	0,571932	-0,191339	0,002135
24	0,248000	0,252941	-0,004941	0,529037	-0,422796	0,002096
25	0,242000	0,249890	-0,007890	0,085790	-0,675202	0,001855
26	0,277000	0,249300	0,027700	0,000000	2,370365	0,001848
27	0,240000	0,246840	-0,006840	-0,357457	-0,585320	0,001965
28	0,236000	0,244478	-0,008478	-0,700616	-0,725524	0,002266
29	0,249000	0,242018	0,006982	-1,058073	0,597428	0,002708
30	0,238000	0,237197	0,000803	-1,758690	0,068723	0,003774
31	0,249000	0,259042	-0,010042	1,415531	-0,859280	0,003230
32	0,239000	0,253531	-0,014531	0,614826	-1,243471	0,002177
33	0,242000	0,251170	-0,009170	0,271668	-0,784667	0,001916
34	0,273000	0,250481	0,022519	0,171579	1,927029	0,001875
35	0,252000	0,249300	0,002700	0,000000	0,231046	0,001848
36	0,252000	0,248710	0,003290	-0,085790	0,281568	0,001855
37	0,230000	0,246250	-0,016250	-0,443247	-1,390525	0,002025
38	0,233000	0,246250	-0,013250	-0,443247	-1,133807	0,002025
39	0,228000	0,244478	-0,016478	-0,700616	-1,410105	0,002266
40	0,231000	0,243790	-0,012790	-0,800704	-1,094445	0,002379

	Wartości przewidywane i reszty BA		
	Mahaln. Odległ.	Usunięte Reszta	Cooka Odległ.
1	2,003727	-0,015203	0,064632
2	1,512046	-0,006155	0,008845
3	0,641127	0,010630	0,017145
4	0,490863	0,019616	0,052952
5	0,196468	0,003763	0,001557
6	0,160282	-0,002117	0,000478
7	0,001840	0,005124	0,002407
8	0,641127	-0,006040	0,005535
9	2,003727	-0,002770	0,002146
10	7,458215	-0,012128	0,116459
11	1,512046	-0,001882	0,000827
12	1,243820	-0,009517	0,018865
13	0,945336	0,024201	0,105585
14	0,279880	0,005227	0,003219
15	0,127776	0,011567	0,013852
16	0,029440	0,000533	0,000027
17	0,127776	0,013543	0,018988
18	0,641127	0,011695	0,020751
19	1,655973	-0,001549	0,000592
20	3,781343	0,020594	0,189376
21	3,297430	0,001351	0,000732
22	0,785872	-0,010893	0,019614
23	0,327106	-0,002313	0,000654
24	0,279880	-0,005105	0,003070
25	0,007360	-0,008094	0,006042
26	0,000000	0,028410	0,073881
27	0,127776	-0,007039	0,005130
28	0,490863	-0,008810	0,010680
29	1,119519	0,007378	0,010703
30	3,092989	0,000897	0,000307
31	2,003727	-0,010872	0,033053
32	0,378012	-0,015053	0,028784
33	0,073803	-0,009423	0,008743
34	0,029440	0,023115	0,050381
35	0,000000	0,002769	0,000702
36	0,007360	0,003375	0,001051
37	0,196468	-0,016753	0,030866
38	0,196468	-0,013660	0,020521
39	0,490863	-0,017122	0,040344
40	0,641127	-0,013343	0,027010

Średnia otrzymanych reszt wynosi $2,710505 \times 10^{-19}$
Można stwierdzić, że jest ona zerowa i świadczy to o nieobciążoności reszt. Możemy stwierdzić, że model regresji jest użyteczny. Wykres normalności dla reszty przedstawia się następująco:

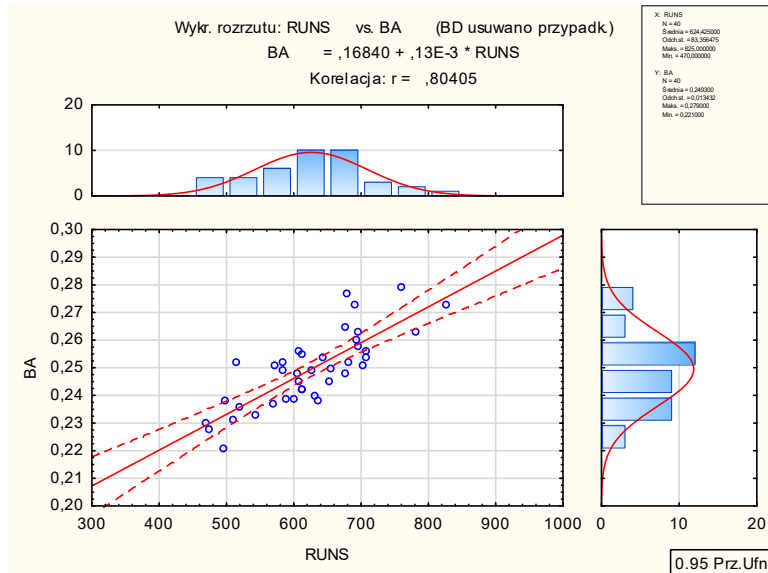


Otrzymany wykres potwierdza warunek normalności reszt..
Reszty odstające dla a):

								Reszty standaryz.: BA (Baseball.sta) Odstające		
Reszty standaryz.								Obserw. Wartość	Przewidyw. Wartość	Reszta
Przyp.	-5.	-4.	-3.	±2.	3.	4.	5.			
26	.	.	.		*	.	.	0,277000	0,249300	0,027700
Minimum	.	.	.		*	.	.	0,277000	0,249300	0,027700
Maksim.	.	.	.		*	.	.	0,277000	0,249300	0,027700
Średnia	.	.	.		*	.	.	0,277000	0,249300	0,027700
Mediana	.	.	.		*	.	.	0,277000	0,249300	0,027700

								Reszty standaryz.: BA (Baseball.sta) Odstające		
Reszty standaryz.								Standard Przewid.	Standard Reszta	Bł. std. W.przew.
Przyp.	-5.	-4.	-3.	±2.	3.	4.	5.			
26	.	.	.		*	.	.	0,00	2,370365	0,001848
Minimum	.	.	.		*	.	.	0,00	2,370365	0,001848
Maksim.	.	.	.		*	.	.	0,00	2,370365	0,001848
Średnia	.	.	.		*	.	.	0,00	2,370365	0,001848
Mediana	.	.	.		*	.	.	0,00	2,370365	0,001848

Wykres rozrzutu i model regresji b):



Powyżej widoczny mamy model regresji, oraz współczynnik r korelacji pomiędzy dwoma zmiennymi RUNS i BA.

Wyniki regresji wielorakiej dla b):

Zmn. zależ. BA	Wielor. R = ,80404932	F = 69,49503
	R^2 = ,64649531	df = 1,38
Liczba przyp. 40	Popraw. R^2 = ,63719255	p = ,000000
	Błąd standardowy estymacji: ,008090606	
Wyr. wolny ,168396635	Błąd std.: ,0097888	t(38) = 17,203 p = 0,0000

RUNS b* = ,804

(istotne b* są podświetlone na czerwono)

Analiza reszt dla b):

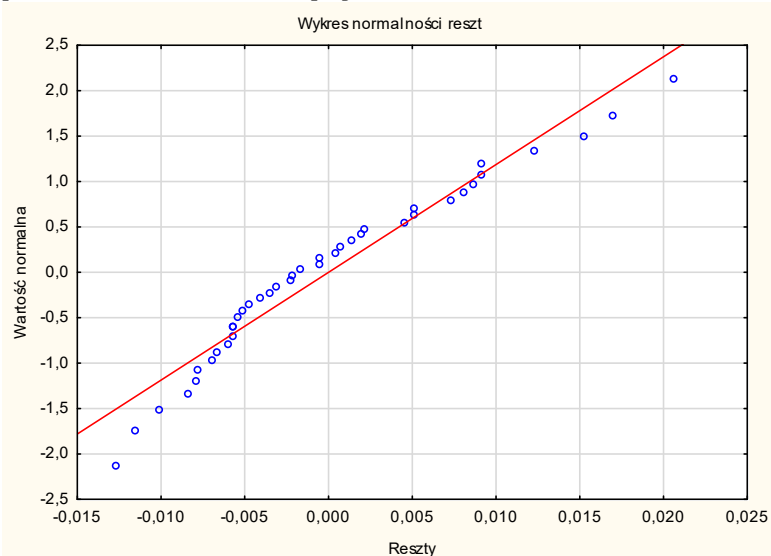
Zmn. zależ. BA	Wielor. R : ,80404932	F = 69,49503
	R^2 : ,64649531	df = 1,38
Liczba przyp. 40	Popraw. R^2 : ,63719255	p = ,000000
	Błąd standardowy estymacji: ,008090606	
Wyr. wolny ,168396635	Błąd std.: ,0097888	t(38) = 17,203 p < 0,0000

	Wartości przewidywane i reszty BA					
	Obserw. Wartość	Przewidyw. Wartość	Reszta	Standard. Przewid.	Standard. Reszta	Bł. std. W.przew.
1	0,245000	0,247172	-0,002172	-0,197045	-0,268448	0,001304
2	0,252000	0,256760	-0,004760	0,690708	-0,588296	0,001561
3	0,265000	0,255853	0,009147	0,606732	1,130602	0,001501
4	0,273000	0,275287	-0,002287	2,406232	-0,282725	0,003370
5	0,256000	0,260128	-0,004128	1,002622	-0,510263	0,001823
6	0,250000	0,253132	-0,003132	0,354801	-0,387099	0,001359
7	0,254000	0,259999	-0,005999	0,990625	-0,741451	0,001812
8	0,238000	0,250670	-0,012670	0,126865	-1,566030	0,001290
9	0,237000	0,242119	-0,005119	-0,664915	-0,632694	0,001542
10	0,221000	0,232531	-0,011531	-1,552669	-1,425246	0,002384
11	0,256000	0,246913	0,009087	-0,221039	1,123183	0,001311
12	0,248000	0,255853	-0,007853	0,606732	-0,970599	0,001501
13	0,279000	0,266736	0,012264	1,614452	1,515815	0,002452
14	0,258000	0,258574	-0,000574	0,858662	-0,070898	0,001695
15	0,263000	0,269716	-0,006716	1,890375	-0,830115	0,002763
16	0,251000	0,242378	0,008622	-0,640922	1,065677	0,001525
17	0,260000	0,258055	0,001945	0,810675	0,240360	0,001655
18	0,255000	0,247690	0,007310	-0,149059	0,903497	0,001294
19	0,239000	0,244451	-0,005451	-0,448975	-0,673751	0,001405
20	0,254000	0,251836	0,002164	0,234835	0,267442	0,001315
21	0,263000	0,258444	0,004556	0,846665	0,563122	0,001685
22	0,245000	0,252873	-0,007873	0,330808	-0,973071	0,001349
23	0,251000	0,259351	-0,008351	0,930642	-1,032184	0,001758
24	0,248000	0,246654	0,001346	-0,245032	0,166409	0,001318
25	0,242000	0,247690	-0,005690	-0,149059	-0,703304	0,001294
26	0,277000	0,256371	0,020629	0,654718	2,549749	0,001535
27	0,240000	0,250152	-0,010152	0,078878	-1,254773	0,001283
28	0,236000	0,235641	0,000359	-1,264749	0,044415	0,002079
29	0,249000	0,249504	-0,000504	0,018895	-0,062302	0,001279
30	0,238000	0,232920	0,005080	-1,516679	0,627914	0,002345
31	0,249000	0,243933	0,005067	-0,496962	0,626308	0,001432
32	0,239000	0,246006	-0,007006	-0,305015	-0,865921	0,001339
33	0,242000	0,247690	-0,005690	-0,149059	-0,703304	0,001294
34	0,273000	0,257796	0,015204	0,786682	1,879192	0,001636
35	0,252000	0,234993	0,017007	-1,324732	2,102089	0,002141
36	0,252000	0,243933	0,008067	-0,496962	0,997110	0,001432
37	0,230000	0,229292	0,000708	-1,852586	0,087511	0,002720
38	0,233000	0,238750	-0,005750	-0,976829	-0,710727	0,001799
39	0,228000	0,229681	-0,001681	-1,816596	-0,207733	0,002679
40	0,231000	0,234475	-0,003475	-1,372719	-0,429457	0,002191

	Wartości przewidywane i reszty BA		
	Mahaln. Odległ.	Usunięte Reszta	Cooka Odległ.
1	0,038827	-0,002230	0,000987
2	0,477078	-0,004944	0,006951
3	0,368123	0,009474	0,023609
4	5,789951	-0,002767	0,010148
5	1,005250	-0,004349	0,007336
6	0,125884	-0,003223	0,002240
7	0,981338	-0,006316	0,015283
8	0,016095	-0,013001	0,032808
9	0,442112	-0,005312	0,007832
10	2,410780	-0,012627	0,105737
11	0,048858	0,009332	0,017464
12	0,368123	-0,008133	0,017400
13	2,606454	0,013504	0,127916
14	0,737300	-0,000600	0,000121
15	3,573518	-0,007603	0,051495
16	0,410781	0,008940	0,021691
17	0,657194	0,002030	0,001317
18	0,022218	0,007502	0,010991
19	0,201579	-0,005621	0,007280
20	0,055147	0,002222	0,000997
21	0,716841	0,004763	0,007516
22	0,109434	-0,008098	0,013928
23	0,866094	-0,008765	0,027701
24	0,060041	0,001383	0,000388
25	0,022218	-0,005839	0,006660
26	0,428656	0,021399	0,125892
27	0,006222	-0,010414	0,020842
28	1,599589	0,000385	0,000075
29	0,000357	-0,000517	0,000051
30	2,300315	0,005546	0,019731
31	0,246971	0,005231	0,006549
32	0,093034	-0,007203	0,010853
33	0,022218	-0,005839	0,006660
34	0,618868	0,015852	0,078441
35	1,754915	0,018287	0,178809
36	0,246971	0,008328	0,016600
37	3,432073	0,000798	0,000550
38	0,954194	-0,006049	0,013828
39	3,300019	-0,001888	0,002983
40	1,884357	-0,003749	0,007873

Średnia otrzymanych reszt wynosi **0,00000000149011611949861**

Można stwierdzić, że jest ona zerowa i świadczy to o nieobciążoności reszt. Możemy stwierdzić, że model regresji jest użyteczny. Wykres normalności dla reszty przedstawia się następująco:



Otrzymany wykres potwierdza warunek normalności reszt. Reszty odstające dla b):

								Reszty standaryz.: BA (Baseball.sta)		
								Odstające		
Reszty standaryz.								Obserw.	Przewidyw.	Reszta
Przyp.	-5.	-4.	-3.	±2.	3.	4.	5.	Wartość	Wartość	
26	.	.	.		*	.	.	0,277000	0,256371	0,020629
35	.	.	.		*	.	.	0,252000	0,234993	0,017007
Minimum	.	.	.		*	.	.	0,252000	0,234993	0,017007
Maksim.	.	.	.		*	.	.	0,277000	0,256371	0,020629
Średnia	.	.	.		*	.	.	0,264500	0,245682	0,018818
Mediana	.	.	.		*	.	.	0,264500	0,245682	0,018818

								Reszty standaryz.: BA (Baseball.sta)		
								Odstające		
Reszty standaryz.								Standard	Standard	Bł. std.
Przyp.	-5.	-4.	-3.	±2.	3.	4.	5.	Przewid.	Reszta	W.przew.
26	.	.	.		*	.	.	0,65472	2,549749	0,001535
35	.	.	.		*	.	.	-1,32473	2,102089	0,002141
Minimum	.	.	.		*	.	.	-1,32473	2,102089	0,001535
Maksim.	.	.	.		*	.	.	0,65472	2,549749	0,002141
Średnia	.	.	.		*	.	.	-0,33501	2,325919	0,001838
Mediana	.	.	.		*	.	.	-0,33501	2,325919	0,001838

Podsumowanie:

Badane zmienne tj.:

- a) ZWYCIĘSTWA i BA - wsp. korelacji: ~ 0.51
- b) RUNS i BA - wsp. korelacji: ~ 0.80

mają wysokie współczynniki korelacji (szczególnie b)). Zależność zmiennych potwierdzają nam powyżej przeprowadzone badania, opierające się na modelu regresji, wykresach rozrzutu i analizie reszt. Nie ma przesłanek do tego, by z tych badań wnioskować inaczej. Nie mam niestety wiedzy nt. rozgrywek baseballowych i nie mogę potwierdzić, że te współczynniki mogą mieć ze sobą powiązanie. W tym przypadku pozostaje wierzyć przeprowadzonej analizie.

b) Regresja wieloraka

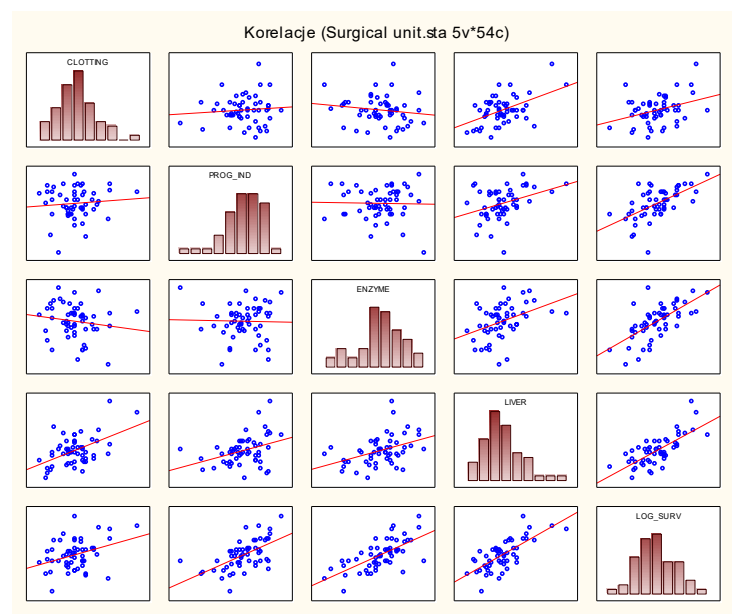
Przeprowadzono analizę regresji wielorakiej dla danych zawartych w arkuszu poniżej. Badany jest wpływ zmiennych niezależnych na zmienną LOG_SURV.

	Example taken from Neter/Wasserman/Kutner (1990): Applied Linear Stat. Models				
	1 CLOTTING	2 PROG_IN D	3 ENZYM E	4 LIVER	5 LOG_SUR V
1	6,7	62	81	2,59	2,301
2	5,1	59	66	1,70	2,004
3	7,4	57	83	2,16	2,310
4	6,5	73	41	2,01	2,004
5	7,8	65	115	4,30	2,707
6	5,8	38	72	1,42	1,903
7	5,7	46	63	1,91	1,903
8	3,7	68	81	2,57	2,104
9	6,0	67	93	2,50	2,305
10	3,7	76	94	2,40	2,307
11	6,3	84	83	4,13	2,517
12	6,7	51	43	1,86	1,813
13	5,8	96	114	3,95	2,919
14	5,8	83	88	3,95	2,519
15	7,7	62	67	3,40	2,225
16	7,4	74	68	2,40	2,336
17	6,0	85	28	2,98	1,940
18	3,7	51	41	1,55	1,531
19	7,3	68	74	3,56	2,332
20	5,6	57	87	3,02	2,236
21	5,2	52	76	2,85	2,037
22	3,4	83	53	1,12	2,134
23	6,7	26	68	2,10	1,845
24	5,8	67	86	3,40	2,342
25	6,3	59	100	2,95	2,441
26	5,8	61	73	3,50	2,158
27	5,2	52	86	2,45	2,258
28	11,2	76	90	5,59	2,759
29	5,2	54	56	2,71	1,857
30	5,8	76	59	2,58	2,250
31	3,2	64	65	0,74	1,851
32	8,7	45	23	2,52	1,763
33	5,0	59	73	3,50	2,064
34	5,8	72	93	3,30	2,470
35	5,4	58	70	2,64	2,061
36	5,3	51	99	2,60	2,265
37	2,6	74	86	2,05	2,072
38	4,3	8	119	2,85	2,079
39	4,8	61	76	2,45	2,179
40	5,4	52	88	1,81	2,170
41	5,2	49	72	1,84	1,978
42	3,6	28	99	1,30	1,875
43	8,8	86	88	6,40	2,684
44	6,5	56	77	2,85	2,185
45	3,4	77	93	1,48	2,281
46	6,5	40	84	3,00	2,090

	Example taken from Neter/Wasserman/Kutner (1990): Applied Linear Stat. Models				
	1 CLOTTING	2 PROG_IN D	3 ENZYM E	4 LIVER	5 LOG_SUR V
47	4,5	73	106	3,05	2,493
48	4,8	86	101	4,10	2,600
49	5,1	67	77	2,86	2,199
50	3,9	82	103	4,55	2,491
51	6,6	77	46	1,96	2,093
52	6,4	85	40	1,21	2,097
53	6,4	59	85	2,33	2,297
54	8,8	78	72	3,20	2,496

Wybrana zmienna zależna: LOG SURV;
 Wybrane zmienne niezależne: CLOTTING, PROG_IND, ENZYME, LIVER

Macierzowy wykres korelacji prezentuje się następująco:



Wyniki regresji wielorakiej

Zmn. zależ.	LOG_SURV	Wielor. R =	,98608421	F =	430,9814
		R^2=	,97236206	df =	4,49
Liczba przyp.	54	Popraw. R^2=	,97010591	p =	0,000000
		Błąd standardowy estymacji:	,047336835		
Wyr. wolny	,488738390	Błąd std.:	,0502405	t(49) =	9,7280
				p =	,0000

CLOTTING b* =	,401	PROG_IND b* =	,571	ENZYME b* =	,735
LIVER b* =	,007				

(istotne b* są podświetlone na czerwono)

Sprawdzamy wyniki regresji.
 Korelacje cząstkowe:

Zmienna	Aktualnie w równaniu są zmienne ; DV: LOG_SURV (Surgical unit.sta)					
	b* w	Cząstk. Korelac.	Semicz. Korelac.	Tolerancja	R-kwadr.	t(49)
CLOTTING	0,401230	0,874058	0,299100	0,555709	0,444291	12,59395
PROG_IND	0,571332	0,949505	0,503112	0,775446	0,224554	21,18409
ENZYME	0,735497	0,959714	0,567836	0,596051	0,403949	23,90936
LIVER	0,007500	0,028206	0,004691	0,391248	0,608752	0,19752

Zmienna	Aktualnie w
	p
CLOTTING	0,000000
PROG_IND	0,000000
ENZYME	0,000000
LIVER	0,844239

Współczynnik równania regresji i ich istotność:

N=54	Podsumowanie regresji zmiennej zależnej: LOG_SURV (Surgical unit.sta) R= ,98608421 R^2= ,97236206 Popraw. R2= ,97010591					
	b*	Bł. std. z b*	b	Bł. std. z b	t(49)	p
W. wolny			0,488738	0,050240	9,72798	0,000000
CLOTTING	0,401230	0,031859	0,068526	0,005441	12,59395	0,000000
PROG_IND	0,571332	0,026970	0,009254	0,000437	21,18409	0,000000
ENZYME	0,735497	0,030762	0,009474	0,000396	23,90936	0,000000
LIVER	0,007500	0,037969	0,001919	0,009713	0,19752	0,844239

Współczynnik regresji wielorakiej i istotność modelu:

	Stat.podsum.; Zmn. zal.:LOG_SURV (Surgical unit.sta)
statystyka	Wartość
R wielorakie	0,986084207
Wielorakie R2	0,972362063
Skorygowane R2	0,970105905
F(4,49)	430,981351
p	1,62831924E-37
Błąd std. estymacji	0,0473368345

Możemy zauważyć, że prawie wszystkie zmienne są mocno skorelowane ze zmienną LOG SURV (dodatnie). Niski współczynnik korelacji, który odstaje do reszty występuje w przypadku zmiennych LOG_SURV i LIVER. Z dotychczasowych analiz wynika, że zmienna LIVER powinna zostać wyeliminowana z modelu (okazała się nieistotna). Poniżej przedstawiono wynik analizy bez uwzględnienia tej zmiennej:

Wyniki regresji wielorakiej

Zmn. zależ.	LOG_SURV	Wielor. R =	,98607305	F =	585,8894
		R^2=	,97234006	df =	3,50
Liczba przyp.	54	Popraw. R^2=	,97068046	p =	0,000000
		Błąd standardowy estymacji:	,046879727		
Wyr. wolny	,483622571	Błąd std.:	,0426339	t(50) =	11,344 p = ,0000

CLOTTING b* = ,405
PROG_IND b* = ,574
ENZYME b* = ,739

(istotne b* są podświetlone na czerwono)

Możemy zauważyć, że poprzez usunięcie zmiennej LIVER zmniejszył nam się błąd standardowy wyrazu wolnego:

Zmienna	Aktualnie w równaniu są zmienne ; DV: LOG_SURV (Surgical unit.sta)					
	b* w	Cząstk. Korelac.	Semicz. Korelac.	Tolerancja	R-kwadr.	t(50)
CLOTTING	0,405342	0,923107	0,399238	0,970108	0,029892	16,97428
PROG_IND	0,573820	0,960163	0,571455	0,991774	0,008226	24,29638
ENZYME	0,739293	0,975077	0,730931	0,977506	0,022494	31,07675

Zmienna	Aktualnie w
	p
CLOTTING	0,000000
PROG_IND	0,000000
ENZYME	0,000000

Koleracje w tej chwili prezentują się następująco:

N=54	Podsumowanie regresji zmiennej zależnej: LOG_SURV (Surgical unit.sta) R= ,98607305 R^2= ,97234006 Popraw. R2= ,97068046						
	b*	Bł. std. z b*	b	Bł. std. z b	t(50)	p	
	W. wolny		0,483623	0,042634	11,34361	0,000000	
	CLOTTING	0,405342	0,023880	0,069229	0,004078	16,97428	0,000000
	PROG_IND	0,573820	0,023618	0,009295	0,000383	24,29638	0,000000
	ENZYME	0,739293	0,023789	0,009523	0,000306	31,07675	0,000000

Wzór modelu regresji możemy odczytać z powyższego podsumowania:
 $\text{LOG_SURV} = 0,48 - 0,07 \text{ CLOTTING} + 0,01 \text{ PROG_IND} + 0,01 \text{ ENZYME}$
(w przybliżeniu).

Wartości przewidywane i reszty:

	Wartości przewidywane i reszty LOG SURV					
	Obserw. Wartość	Przewidyw. Wartość	Reszta	Standard. Przewid.	Standard. Reszta	Bł. std. W.przew.
1	2,301030	2,295106	0,005924	0,329527	0,126371	0,007612
2	2,004321	2,013607	-0,009285	-0,713177	-0,198070	0,008051
3	2,309630	2,316139	-0,006509	0,407438	-0,138846	0,009966
4	2,004321	2,002569	0,001752	-0,754062	0,037375	0,013164
5	2,706718	2,722933	-0,016215	1,914247	-0,345882	0,016465
6	1,903090	1,924020	-0,020930	-1,045017	-0,446460	0,011696
7	1,903090	1,905744	-0,002654	-1,112712	-0,056620	0,010180
8	2,103804	2,143187	-0,039384	-0,233196	-0,840100	0,010834
9	2,305351	2,407398	-0,102047	0,745471	-2,176777	0,008274
10	2,307496	2,341347	-0,033851	0,500809	-0,722075	0,012580
11	2,517196	2,490942	0,026254	1,054927	0,560028	0,010486
12	1,812913	1,830980	-0,018067	-1,389647	-0,385391	0,013355
13	2,919078	2,863085	0,055994	2,433386	1,194409	0,018126
14	2,518514	2,494649	0,023865	1,068660	0,509058	0,010463
15	2,225309	2,231008	-0,005699	0,092103	-0,121565	0,010253
16	2,336460	2,331298	0,005161	0,463588	0,110096	0,009923
17	1,939519	1,955687	-0,016168	-0,927717	-0,344885	0,018151
18	1,531479	1,604248	-0,072769	-2,229490	-1,552247	0,016710
19	2,332438	2,325747	0,006691	0,443028	0,142726	0,008912
20	2,235528	2,229621	0,005908	0,086964	0,126015	0,007417
21	2,037426	2,050700	-0,013274	-0,575779	-0,283144	0,007965
22	2,133539	1,995186	0,138353	-0,781411	2,951238	0,016731
23	1,845098	1,836697	0,008401	-1,368470	0,179194	0,016507
24	2,342423	2,326889	0,015533	0,447257	0,331347	0,007093
25	2,440909	2,420473	0,020436	0,793901	0,435930	0,010085
26	2,158362	2,147319	0,011043	-0,217892	0,235567	0,006560
27	2,257679	2,145933	0,111745	-0,223025	2,383661	0,008265
28	2,758912	2,822469	-0,063557	2,282939	-1,355739	0,023990
29	1,857332	1,878824	-0,021491	-1,212429	-0,458437	0,010211
30	2,250420	2,153412	0,097008	-0,195323	2,069295	0,009726
31	1,851258	1,919022	-0,067764	-1,063530	-1,445480	0,013329
32	1,763428	1,723204	0,040224	-1,788861	0,858017	0,021548
33	2,064458	2,073347	-0,008889	-0,491893	-0,189612	0,007447
34	2,469822	2,440025	0,029797	0,866326	0,635597	0,008720
35	2,060698	2,063174	-0,002476	-0,529575	-0,052821	0,007243
36	2,264818	2,267364	-0,002546	0,226769	-0,054316	0,010263
37	2,071882	2,170420	-0,098538	-0,132324	-2,101925	0,015256
38	2,079181	1,988933	0,090248	-0,804570	1,925094	0,025252
39	2,178977	2,106660	0,072317	-0,368497	1,542605	0,007579
40	2,170262	2,178825	-0,008564	-0,101188	-0,182675	0,008352
41	1,977724	1,984723	-0,007000	-0,820165	-0,149313	0,008803
42	1,875061	1,935900	-0,060838	-1,001014	-1,297754	0,017497
43	2,683947	2,730219	-0,046272	1,941237	-0,987037	0,016523
44	2,184691	2,187399	-0,002708	-0,069430	-0,057759	0,007636
45	2,281033	2,320349	-0,039316	0,423033	-0,838661	0,013493

	Wartości przewidywane i reszty LOG SURV		
	Mahaln. Odległ.	Usunięte Reszta	Cooka Odległ.
1	0,415804	0,006085	0,000111
2	0,581813	-0,009568	0,000307
3	1,413806	-0,006817	0,000239
4	3,197638	0,001902	0,000032
5	5,556316	-0,018496	0,004801
6	2,317256	-0,022319	0,003527
7	1,517955	-0,002786	0,000042
8	1,849323	-0,041606	0,010518
9	0,669289	-0,105327	0,039306
10	2,835099	-0,036477	0,010900
11	1,670117	0,027637	0,004347
12	3,319974	-0,019663	0,003569
13	6,941484	0,065835	0,073705
14	1,658507	0,025116	0,003574
15	1,553922	-0,005985	0,000195
16	1,393067	0,005403	0,000149
17	6,963806	-0,019019	0,006169
18	5,752492	-0,083360	0,100435
19	0,934008	0,006942	0,000198
20	0,345081	0,006059	0,000105
21	0,548522	-0,013668	0,000614
22	5,769017	0,158547	0,364205
23	5,589862	0,009590	0,001297
24	0,231777	0,015897	0,000658
25	1,471439	0,021428	0,002417
26	0,056277	0,011264	0,000283
27	0,666036	0,115330	0,047034
28	12,897763	-0,086105	0,220861
29	1,532973	-0,022562	0,002747
30	1,299989	0,101372	0,050320
31	3,303070	-0,073724	0,049982
32	10,216350	0,050999	0,062509
33	0,355764	-0,009119	0,000239
34	0,852118	0,030864	0,003749
35	0,283665	-0,002537	0,000017
36	1,558530	-0,002674	0,000039
37	4,631103	-0,110209	0,146314
38	14,396567	0,127137	0,533505
39	0,403773	0,074258	0,016395
40	0,700607	-0,008844	0,000282
41	0,887486	-0,007256	0,000211
42	6,401465	-0,070685	0,079173
43	5,602268	-0,052835	0,039447
44	0,424633	-0,002782	0,000023
45	3,409146	-0,042867	0,017317

	Wartości przewidywane i reszty LOG_SURV					
	Obserw. Wartość	Przewidyw. Wartość	Reszta	Standard. Przewid.	Standard. Reszta	Bł. std. W.przew.
46	2,089905	2,105349	-0,015444	-0,373355	-0,329429	0,011779
47	2,492760	2,483125	0,009635	1,025973	0,205530	0,012279
48	2,599883	2,577107	0,022776	1,374093	0,485837	0,013598
49	2,198657	2,192720	0,005937	-0,049722	0,126650	0,007160
50	2,491362	2,496670	-0,005308	1,076143	-0,113229	0,014384
51	2,093422	2,094287	-0,000865	-0,414329	-0,018456	0,012525
52	2,096910	2,097658	-0,000748	-0,401841	-0,015959	0,015217
53	2,296665	2,284546	0,012119	0,290414	0,258509	0,007612
54	2,495544	2,503490	-0,007946	1,101406	-0,169488	0,014434

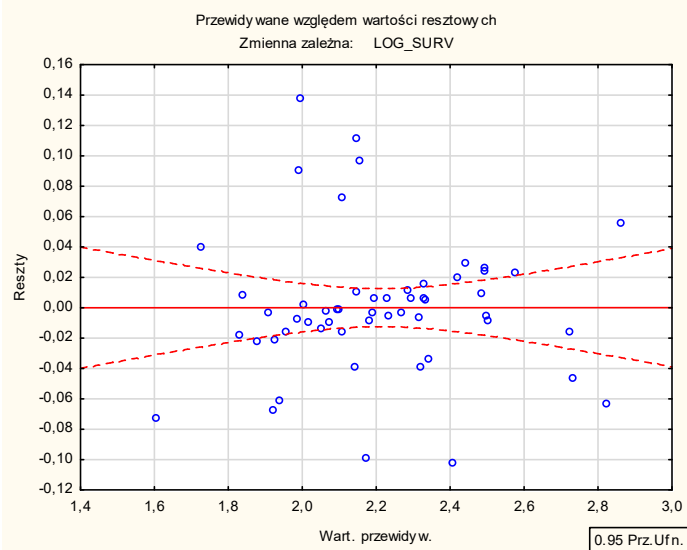
	Wartości przewidywane i reszty LOG_SURV		
	Mahaln. Odległ.	Usunięte Reszta	Cooka Odległ.
46	2,364353	-0,016484	0,001951
47	2,654672	0,010345	0,000835
48	3,477721	0,024868	0,005919
49	0,254800	0,006079	0,000098
50	4,008347	-0,005860	0,000368
51	2,801826	-0,000932	0,000007
52	4,602772	-0,000836	0,000008
53	0,415696	0,012447	0,000465
54	4,042855	-0,008778	0,000831

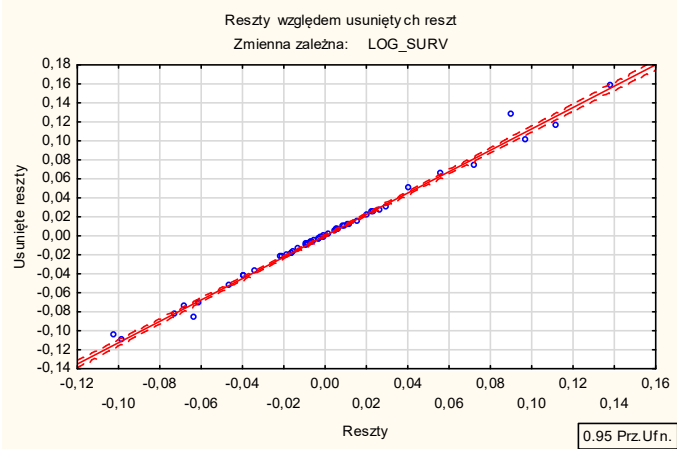
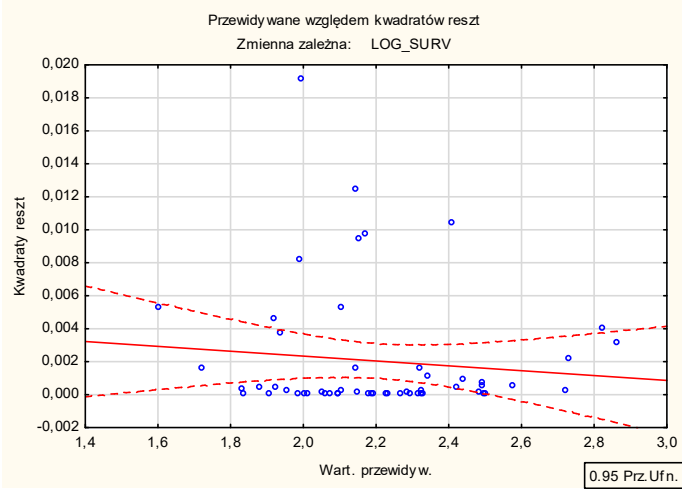
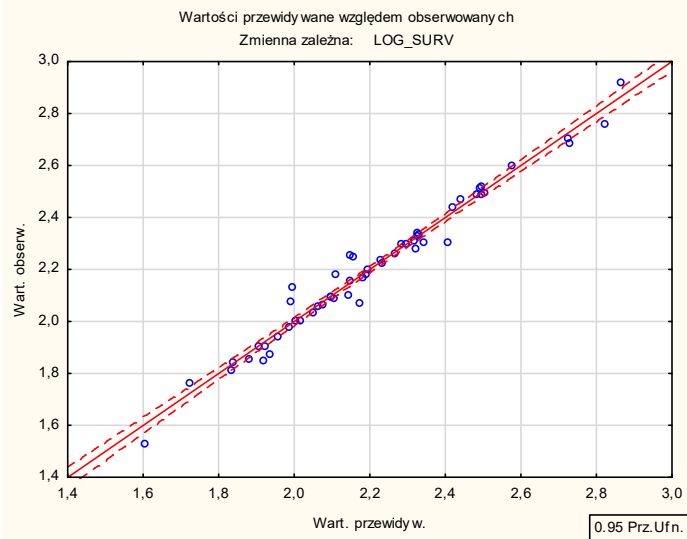
Średnia reszt w tym przypadku wynosi: 0.00000000220757943560556
Jest ona bliska zero, więc reszty są nieobciążone, nasz model regresji wielorakiej jest użyteczny.

Wykres normalności reszt prezentuje się tak:



Wykresy rozrzutu:





Przeprowadzamy dokładniejszą analizę skoleowaną na poszukiwanie wartości odstających.

Przyp. -2,2 Standaryz. przewidywane 2,43					Standaryz. przewidywane (Surgical unit.sta)		
					Obserw. Wartość	Przewidyw. Wartość	Reszta
13	*
28	*.
18	.*
43	*.
5*
32	.	*
12	.	.	.*		.	.	.
48	*	.
23	.	.	.*		.	.	.
29	.	.	*		.	.	.
7	.	.	*		.	.	.
54	*	.
50	*	.
14	*	.
31	.	.	*		.	.	.
11	*	.
6	.	.	*		.	.	.
47	*	.
42	.	.	*		.	.	.
17	.	.	*		.	.	.
34	.	.	.		*	.	.
41	.	.	*		.	.	.
38	.	.	*		.	.	.
25	.	.	.		*	.	.
22	.	.	*		.	.	.
4	.	.	*		.	.	.
9	.	.	.		*	.	.
2	.	.	*		.	.	.
21	.	.	.*		.	.	.
35	.	.	.*		.	.	.
10	.	.	.		*	.	.
33	.	.	.*		.	.	.
16	.	.	.		*	.	.
24	.	.	.		*	.	.
19	.	.	.		*	.	.
45	.	.	.		*	.	.
51	.	.	.		*	.	.
3	.	.	.		*	.	.
52	.	.	.		*	.	.
46	.	.	.		*	.	.
39	.	.	.		*	.	.
1	.	.	.		*	.	.
53	.	.	.		*	.	.
8	.	.	.		*	.	.
36	.	.	.		*	.	.

Standaryz. przewidywane					Standaryz. przewidywane (Surgical unit.sta)		
					Standard Przewid.	Standard Reszta	Bł. std. W.przew.
Przyp. -2,2	.	.	.	2,43			
13	.	.	.	*	2,43339	1,19441	0,018126
28	.	.	.	*	2,28294	-1,35574	0,023990
18	*	.	.	.	-2,22949	-1,55225	0,016710
43	.	.	.	*	1,94124	-0,98704	0,016523
5	.	.	.	*	1,91425	-0,34588	0,016465
32	*	.	.	.	-1,78886	0,85802	0,021548
12	.	*	.	.	-1,38965	-0,38539	0,013355
48	.	.	.	*	1,37409	0,48584	0,013598
23	.	*	.	.	-1,36847	0,17919	0,016507
29	.	.	*	.	-1,21243	-0,45844	0,010211
7	.	.	*	.	-1,11271	-0,05662	0,010180
54	.	.	.	*	1,10141	-0,16949	0,014434
50	.	.	.	*	1,07614	-0,11323	0,014384
14	.	.	.	*	1,06866	0,50906	0,010463
31	.	.	*	.	-1,06353	-1,44548	0,013329
11	.	.	.	*	1,05493	0,56003	0,010486
6	.	.	*	.	-1,04502	-0,44646	0,011696
47	.	.	.	*	1,02597	0,20553	0,012279
42	.	.	*	.	-1,00101	-1,29775	0,017497
17	.	.	*	.	-0,92772	-0,34489	0,018151
34	.	.	.	*	0,86633	0,63560	0,008720
41	.	.	*	.	-0,82016	-0,14931	0,008803
38	.	.	*	.	-0,80457	1,92509	0,025252
25	.	.	.	*	0,79390	0,43593	0,010085
22	.	.	*	.	-0,78141	2,95124	0,016731
4	.	.	*	.	-0,75406	0,03738	0,013164
9	.	.	.	*	0,74547	-2,17678	0,008274
2	.	.	*	.	-0,71318	-0,19807	0,008051
21	.	.	*	.	-0,57578	-0,28314	0,007965
35	.	.	*	.	-0,52957	-0,05282	0,007243
10	.	.	.	*	0,50081	-0,72207	0,012580
33	.	.	*	.	-0,49189	-0,18961	0,007447
16	.	.	.	*	0,46359	0,11010	0,009923
24	.	.	.	*	0,44726	0,33135	0,007093
19	.	.	.	*	0,44303	0,14273	0,008912
45	.	.	.	*	0,42303	-0,83866	0,013493
51	.	.	.	*	-0,41433	-0,01846	0,012525
3	.	.	.	*	0,40744	-0,13885	0,009966
52	.	.	.	*	-0,40184	-0,01596	0,015217
46	.	.	.	*	-0,37335	-0,32943	0,011779
39	.	.	.	*	-0,36850	1,54261	0,007579
1	.	.	.	*	0,32953	0,12637	0,007612
53	.	.	.	*	0,29041	0,25851	0,007612
8	.	.	.	*	-0,23320	-0,84010	0,010834
36	.	.	.	*	0,22677	-0,05432	0,010263

Standaryz. przewidywane Przyp. -2,2 . . . 2,43					Standaryz. przewidywane (Surgical unit.sta)		
					Mahaln. Odległ.	Usunięte Reszta	Cooka Odległ.
13 *	6,94148	0,065835	0,073705				
28 *.	12,89776	-0,086105	0,220861				
18 .*	5,75249	-0,083360	0,100435				
43 *.	5,60227	-0,052835	0,039447				
5 *.	5,55632	-0,018496	0,004801				
32 . *	10,21635	0,050999	0,062509				
12 . .*	3,31997	-0,019663	0,003569				
48 * . .	3,47772	0,024868	0,005919				
23 . .*	5,58986	0,009590	0,001297				
29 . . *	1,53297	-0,022562	0,002747				
7 . . *	1,51796	-0,002786	0,000042				
54* . .	4,04286	-0,008778	0,000831				
50* . .	4,00835	-0,005860	0,000368				
14* . .	1,65851	0,025116	0,003574				
31 . . *	3,30307	-0,073724	0,049982				
11* . .	1,67012	0,027637	0,004347				
6 . . *	2,31726	-0,022319	0,003527				
47* . .	2,65467	0,010345	0,000835				
42 . . *	6,40147	-0,070685	0,079173				
17 . . *	6,96381	-0,019019	0,006169				
34 . . . * . . .	0,85212	0,030864	0,003749				
41 . . *	0,88749	-0,007256	0,000211				
38 . . *	14,39657	0,127137	0,533505				
25 . . . *. . . .	1,47144	0,021428	0,002417				
22 . . *	5,76902	0,158547	0,364205				
4 . . *	3,19764	0,001902	0,000032				
9 . . . *. . . .	0,66929	-0,105327	0,039306				
2 . . *	0,58181	-0,009568	0,000307				
21 . . .*	0,54852	-0,013668	0,000614				
35 . . .*	0,28367	-0,002537	0,000017				
10 . . . *	2,83510	-0,036477	0,010900				
33 . . .*	0,35576	-0,009119	0,000239				
16 . . . *	1,39307	0,005403	0,000149				
24 . . . *	0,23178	0,015897	0,000658				
19 . . . *	0,93401	0,006942	0,000198				
45 . . . *	3,40915	-0,042867	0,017317				
51 . . . *	2,80183	-0,000932	0,000007				
3 . . . *	1,41381	-0,006817	0,000239				
52 . . . *	4,60277	-0,000836	0,000008				
46 . . . *	2,36435	-0,016484	0,001951				
39 . . . *	0,40377	0,074258	0,016395				
1 . . . *	0,41580	0,006085	0,000111				
53 . . . *	0,41570	0,012447	0,000465				
8 . . . *	1,84932	-0,041606	0,010518				
36 . . . *	1,55853	-0,002674	0,000039				

					Standaryz. przewidywane (Surgical unit.sta)		
Standaryz. przewidywane					Obserw. Wartość	Przewidyw. Wartość	Reszta
Przyp. -2,2	.	.	.	2,43			
27	.	.	*	. . .	2,257679	2,145933	0,111745
26	.	.	*	. . .	2,158362	2,147319	0,011043
30	.	.	*	. . .	2,250420	2,153412	0,097008
37	.	.	*	. . .	2,071882	2,170420	-0,098538
40	.	.	*	. . .	2,170262	2,178825	-0,008564
15	.	.	*	. . .	2,225309	2,231008	-0,005699
20	.	.	*	. . .	2,235528	2,229621	0,005908
44	.	.	*	. . .	2,184691	2,187399	-0,002708
49	.	.	*	. . .	2,198657	2,192720	0,005937
Minimum	.*	1,531479	1,604248	-0,102047
Maksim. *	2,919078	2,863085	0,138353
Średnia	.	.	*	. . .	2,206143	2,206143	-0,000000
Mediana	.	.	*	. . .	2,191674	2,174623	-0,002600

					Standaryz. przewidywane (Surgical unit.sta)		
Standaryz. przewidywane					Standard Przewid.	Standard Reszta	Bł. std. W.przew.
Przyp. -2,2	.	.	.	2,43			
27	.	.	*	. . .	-0,22303	2,38366	0,008265
26	.	.	*	. . .	-0,21789	0,23557	0,006560
30	.	.	*	. . .	-0,19532	2,06930	0,009726
37	.	.	*	. . .	-0,13232	-2,10193	0,015256
40	.	.	*	. . .	-0,10119	-0,18268	0,008352
15	.	.	*	. . .	0,09210	-0,12156	0,010253
20	.	.	*	. . .	0,08696	0,12601	0,007417
44	.	.	*	. . .	-0,06943	-0,05776	0,007636
49	.	.	*	. . .	-0,04972	0,12665	0,007160
Minimum	.*	-2,22949	-2,17678	0,006560
Maksim. *	2,43339	2,95124	0,025252
Średnia	.	.	*	. . .	0,00000	-0,00000	0,011994
Mediana	.	.	*	. . .	-0,11676	-0,05547	0,010474

					Standaryz. przewidywane (Surgical unit.sta)		
Standaryz. przewidywane					Mahaln. Odległ.	Usunięte Reszta	Cooka Odległ.
Przyp. -2,2	.	.	.	2,43			
27	.	.	*		0,66604	0,115330	0,047034
26	.	.	*		0,05628	0,011264	0,000283
30	.	.	*		1,29999	0,101372	0,050320
37	.	.	*		4,63110	-0,110209	0,146315
40	.	.	*		0,70061	-0,008844	0,000282
15	.	.	*		1,55392	-0,005985	0,000195
20	.	.	*		0,34508	0,006059	0,000105
44	.	.	*		0,42463	-0,002782	0,000023
49	.	.	*		0,25480	0,006079	0,000098
Minimum	.	*	.		0,05628	-0,110209	0,000007
Maksim.	.	.	.		14,39657	0,158547	0,533505
Średnia	.	.	*		2,94444	0,000097	0,035414
Mediana	.	.	*		1,66431	-0,002728	0,002184

Predykcja dla dwóch wybranych zestawów wartości zmiennych niezależnych:

Obliczanie wartości (Surgical unit.sta) zmiennej: LOG_SURV			
Zmienna	Wagi b	Wartość	Wagi b * Wartość
CLOTTING	0,069229	6,00000	0,415372
PROG_IND	0,009295	70,00000	0,650622
ENZYME	0,009523	60,00000	0,571397
W. wolny			0,483623
Przewidyw.			2,121013
-95,0%GU			2,103781
+95,0%GU			2,138246

Określ wartości zmiennych niezależnych

CLOTTING

PROG_IND

ENZYME

Wspólna wartość

OK Anuluj Zastosuj

Obliczanie wartości (Surgical unit.sta) zmiennej: LOG_SURV			
Zmienna	Wagi b	Wartość	Wagi b * Wartość
CLOTTING	0,069229	8,00000	0,553829
PROG_IND	0,009295	90,00000	0,836514
ENZYME	0,009523	70,00000	0,666630
W. wolny			0,483623
Przewidyw.			2,540596
-95,0%GU			2,511526
+95,0%GU			2,569665

Określ wartości zmiennych niezależnych

CLOTTING

PROG_IND

ENZYME

Wspólna wartość

OK Anuluj Zastosuj

Rozrzut w 3D badanych zmiennych prezentuje się następująco:

Rozrzutu 3W ENZYME w zględem CLOTTING i PROG_IND
Surgical unit.sta 5v*54c

