Q-test TEST FOR OUTLIERS (non-parametric)

	ug/l	sorted	Q
1	14.27		
2	13.43		
3	14.25		
4	14.83		
5	14.64		
6	14.09		
7	15.19		
8	12.93		
9	13.94		
10	11.20		
	alpha=		
	range=		
	Qcrit=		

H0…..

$$Q_{exp} = \frac{X_2 - X_1}{X_N - X_1} \qquad Q_{exp} = \frac{X_N - X_{N-1}}{X_N - X_1}$$

N	Q_{crit} (CL:90%)	Q_{crit} (CL:95%)	Q_{crit} (CL:99%)
3	0.941	0.970	0.994
4	0.765	0.829	0.926
5	0.642	0.710	0.821
6	0.560	0.625	0.740
7	0.507	0.568	0.680
8	0.468	0.526	0.634
9	0.437	0.493	0.598
10	0.412	0.466	0.568

Confidence interval
for MEDIAN
(non-parametric distributions)

https://www.youtube.com/watch?v=cxUWQCwxQgk

Find 95%-confidence interval of medians for both samples A and B:

species A	species B
16	34
32	36
37	38
39	45
40	50
41	54
42	56
50	59
82	69
	91

n	j	
$n \leq 5$: no confidence inte		
6	1	
7	1	
8	1	
9	2	
10	2	
11	2	
12	3	
13	3	
14	3	
15	4	
16	4	
17	5	
18	5	
19	5	
20	6	

k	p
erval possible.	
6	0.969
7	0.984
7	0.961
8	0.961
9	0.979
10	0.988
10	0.961
11	0.978
11	0.965
12	0.965
12	0.951
13	0.951
14	0.969
15	0.981
15	0.959

Wilcoxon signed-rank test
FOR PAIRS (DEPENDENT SAMPLES)

A new drug for blood pressure correction was tested on 9 patients.
The pressure was measured before and after application.
Decide, if there is an effect of the drug!

$$W = \min (W^+ ; W^-$$

after	84	75	88	91	85	65	71
before	97	72	93	110	95	78	69

differences

abs

rank

)

90	75
115	75

alpha=0.05		
n	one-tailed	double-tailed
6	2	0
7	3	2
8	5	3
9	8	5
10	10	8
11	13	10
12	17	13
13	21	17
14	25	21
15	30	25
16	35	30
17	41	35
18	47	40
19	53	46
20	60	52
21	67	59
22	75	66
23	83	73
24	91	81
25	100	89

Mann-Whitney U-test

There were selected 11 fields of similar quality. On 5 randomly selected fields a new fertilizer was teste
Yields of wheat obtained in August were: 5.1, 6.7, 5.6, 6.3, 5.9 t/ha and 4.5, 5.4, 4.8, 4.4, 5.3, 5.0 t/ha, r
Find, if the fertilizer has an effect.

fertilizer	no fertil.
5.1	4.5
6.7	5.4
5.6	4.8
6.3	4.4
5.9	5.3
	5.0

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

n1\n2	2
2	
3	
4	
5	0
6	0
7	0
8	1
9	1
10	1
11	1
12	2
13	2
14	2
15	3
16	3
17	3
18	4
19	4
20	4

d, the rest was kept without the fertilizer.
esp.

3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
		0	0	0	1	1	1	1	2	2	2	3	3	3	4	4	4
	0	1	2	2	3	3	4	5	5	6	7	7	8	9	9	10	11
0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

Kruskal-Wallis test

In 3 samples of crude oils, Ni was determined by AAS. Decide (applying Kruskal-Wallis test), if there is a significant difference among the samples.

Sample						
1	14.2	16.8	19.1	15.5	16	15.9
2	14.5	20	18	15.4	16.1	17.7
3	18.3	20.1	17.7	17.9	19.3	16.9

$$H = \frac{12}{n(n+1)}\left(\frac{R_1^2}{n_1}\right.$$

alpha=0.05

$$H = \frac{12}{n(n+1)}\left(\frac{R_1^2}{n_1} + \frac{R_2^2}{n_2} + \ldots + \frac{R_k^2}{n_k}\right) - 3(n+1)$$

6	4	3	5.610
6	4	4	5.681
6	5	1	4.990
6	5	2	5.338
6	5	3	5.602
6	5	4	5.661
6	5	5	5.729
6	6	1	4.945
6	6	2	5.410
6	6	3	5.625
6	6	4	5.725
6	6	5	5.765
6	6	6	5.801
7	7	7	5.819
8	8	8	5.805

NON-PARAMETRIC

	data	sorted	Q	
1	32.56	30.10	0.35	< 0,526
2	35.14	32.56		
3	34.10	33.33		
4	33.33	34.10		
5	37.10	34.45		
6	34.45	35.12		
7	35.12	35.14		
8	30.10	37.10	0.28	< 0,526

$$Q_{exp} = \frac{X_2 - X_1}{X_N - X_1} \qquad Q_{exp} = \frac{X_N - X_{N-1}}{X_N - X_1}$$

N	Q_{crit} (CL:90%)	Q_{crit} (CL:95%)	Q_{crit} (CL:99%)
3	0.941	0.970	0.994
4	0.765	0.829	0.926
5	0.642	0.710	0.821
6	0.560	0.625	0.740
7	0.507	0.568	0.680
8	0.468	0.526	0.634
9	0.437	0.493	0.598
10	0.412	0.466	0.568

Q-test

alpha: 0.05
range: 7.00
Qcritical: 0.526 OK

H0: Accepted. There is no outlier present

Confidence interval estimation

	data	sorted
1	32.56	30.10
2	35.14	32.56
3	34.10	33.33
4	33.33	34.10
5	37.10	34.45
6	34.45	35.12
7	35.12	35.14
8	30.10	37.10

count: 8
95% CI = <30,10;35,14>

n	j	k	p
$n \leq 5$: no confidence interval possible.			
6	1	6	0.969
7	1	7	0.984
8	1	7	0.961
9	2	8	0.961
10	2	9	0.979
11	2	10	0.988
12	3	10	0.961
13	3	11	0.978
14	3	11	0.965
15	4	12	0.965
16	4	12	0.951
17	5	13	0.951
18	5	14	0.969
19	5	15	0.981
20	6	15	0.959

Conclusion: In my opinion in this case is better Student's CI because it is more acuracy than CI estimatic

PARAMETRIC

	data	sorted	G	
1	32.56	30.10	1.8745	< 2,1266
2	35.14	32.56		
3	34.10	33.33		
4	33.33	34.10		
5	37.10	34.45		
6	34.45	35.12		
7	35.12	35.14		
8	30.10	37.10	1.5008	< 2,1266

n	g_{cr} α=(
3	1.1
4	1.4
5	1.7
6	1.8
7	2.0
8	2.1
9	2.2
10	2.2
11	2.3
12	2.4
13	2.4
14	2.5

Grubbs test

mean: 33.9875
stand. dev. (sample): 2.0739
Gcritical: 2.1266 OK

H0: Accepted. There is no outlier present

Student's confident interval

	data	sorted
1	32.56	30.10
2	35.14	32.56
3	34.10	33.33
4	33.33	34.10
5	37.10	34.45
6	34.45	35.12
7	35.12	35.14
8	30.10	37.10

count = 8
mean = 33.99
stand. dev. (sample) = 2.0739
v = n - 1
v = 8 - 1
v = 7
alpha/2 = 0.025
tv;alpha/2 = 2.3646

https://www.youtube.com/watch?v

95% CI = mean +/- tv;alpha/2 * (stand. dev. (sample) / SQUARE(count))
95% CI = 33,99 +/- 1.7338
32.254
35.721
95% CI = <32,254;35,721>

ɔn. Non-parametric is much more suitable when we don't know distribution data.

it
0.05
543
812
150
871
200
266
150
900
547
116
620
073

There are no outliers in the data set

There is exactly one outlier in the data set

The Grubbs' test statistic is defined as:

$$G = \frac{\max |Y_i - \bar{Y}|}{s}$$

=MUD390jtgQs