Criteria of periodicity and their statistical properties

S. Repin

Space Research Institute of RAS

Pushchino, April 22, 2008

Periodicity criteria

- Criteria of periodicity are the functions of measured values and trial period T, which take large value when the data contain the component with period T and take small value if the data set does not include such a component.
- Criteria of periodicity can be applied to pulsars, variable stars, cepheids, motion of planets and satellites, quasi-periodic oscillations in accretion disks and many other phenomena.

Periodicity criteria

Assume that we have so-called rare pulses:

t_i $i = 1, 2, \ldots, M$

First criterion (Rayleigh criterion) is

$$K_1(T) = \frac{2}{M}\left[\left(\sum_{i=1}^{M} \sin\frac{2\pi t_i}{T}\right)^2 + \left(\sum_{i=1}^{M} \cos\frac{2\pi t_i}{T}\right)^2\right],$$

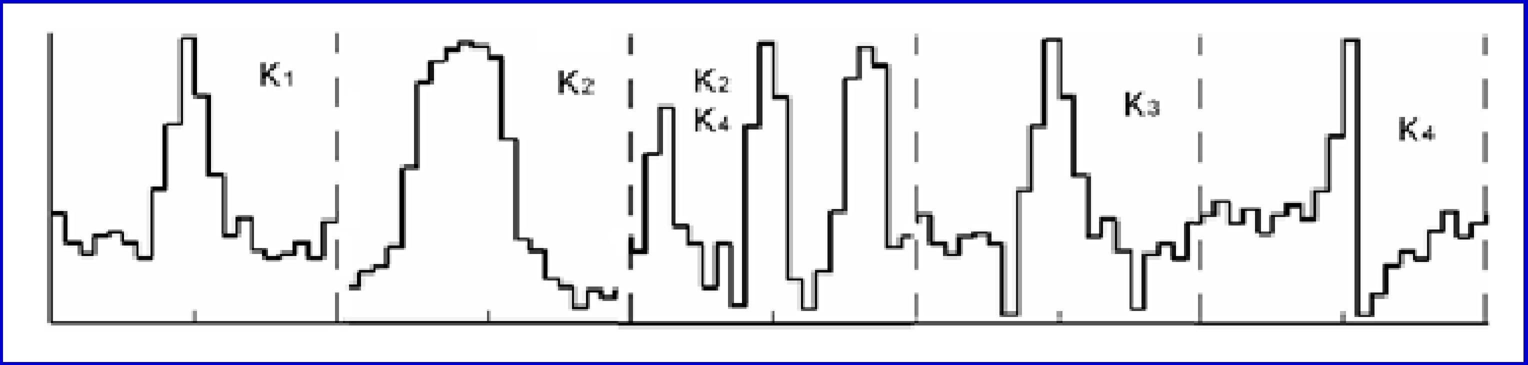
Where t_i are the time moment of events,
T is a trial period

Periodicity criteria

Define the phase of event as

$$\varphi_i = t_i - T \cdot \left[\frac{t_i}{T}\right], \qquad (0 \le \varphi_i \le T),$$

And build an m-bin light curve



Periodicity criteria

Criteria K_2, K_3, K_4 are analyze the structure of the light curve and have the form:

$$K_2(T) = \frac{m}{M} \sum_{j=1}^{m} \left(n_j - \frac{M}{m} \right)^2 ,$$

$$K_3(T) = 1 - \frac{\min n_j}{\max n_j},$$

$$K_4(T) = \frac{m}{M} \left[\sum_{j=1}^{m-1} (n_{j+1} - n_j)^2 + (n_1 - n_m)^2 \right]$$

Properties of the criteria with constant T

Mean values, covariation and correlation matrices

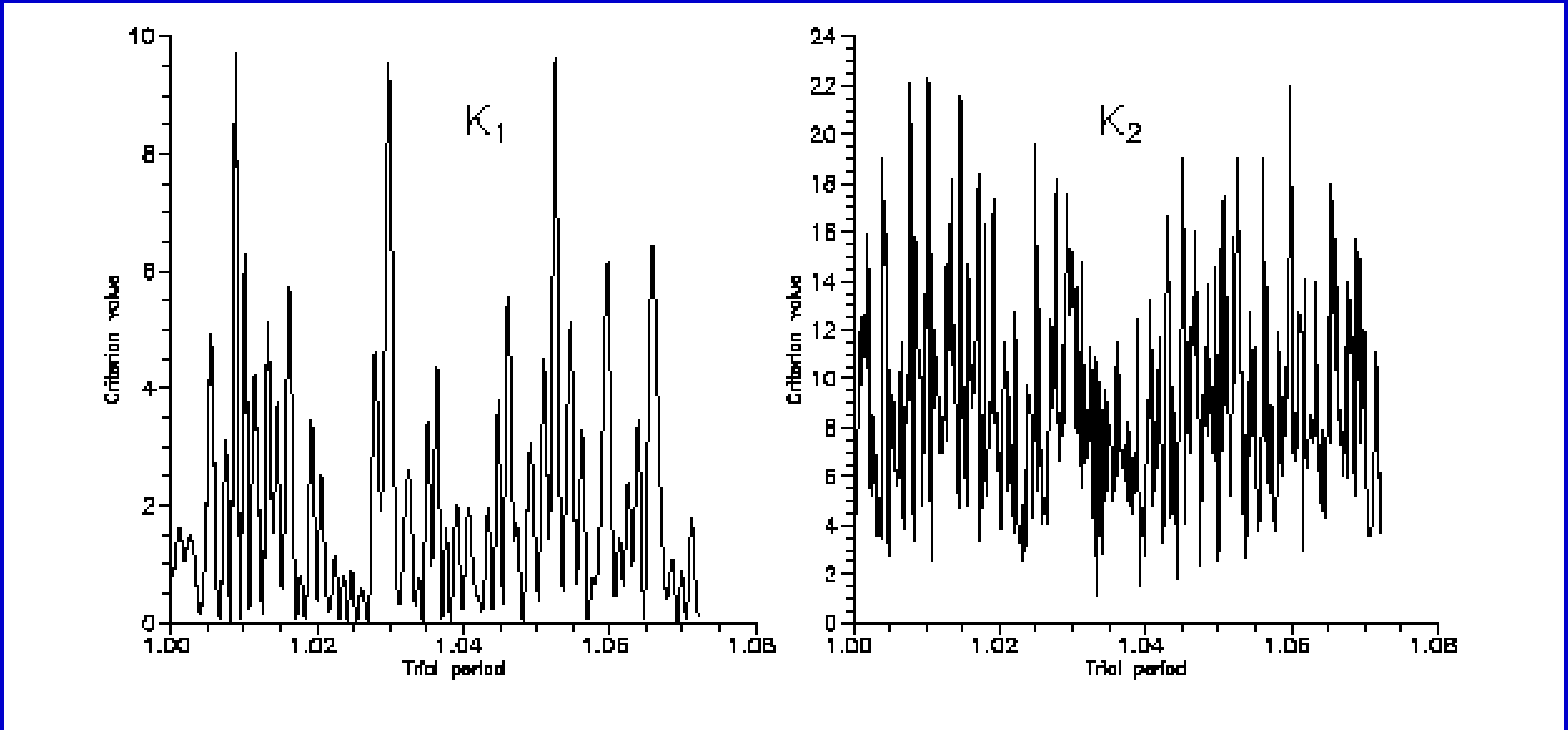
$$MK = (2.033, \quad 8.961, \quad 0.2633, \quad 20.0004)$$

$$DK = \begin{pmatrix} 4.151 & 4.030 & 0.0499 & 1.563 \\ 4.030 & 17.98 & 0.2247 & 40.28 \\ 0.0499 & 0.2247 & 0.0035 & 0.5030 \\ 1.563 & 40.28 & 0.5030 & 121.4 \end{pmatrix} \qquad \rho = \begin{pmatrix} 1 & 0.466 & 0.414 & 0.069 \\ 0.466 & 1 & 0.897 & 0.862 \\ 0.414 & 0.897 & 1 & 0.773 \\ 0.069 & 0.862 & 0.773 & 1 \end{pmatrix}$$

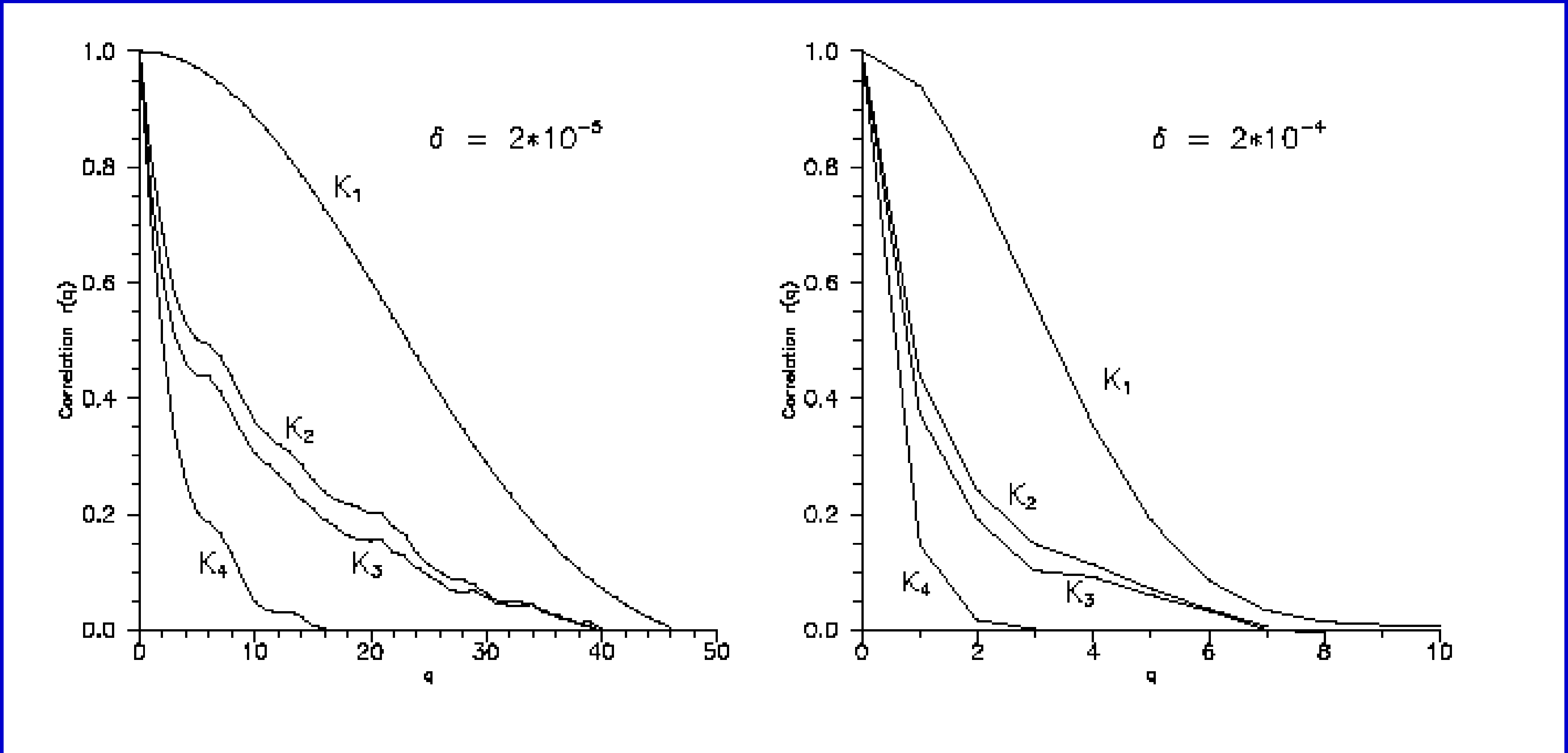
N = 100000 realizations, M = 1000 event in a series

Correlation function

Typical structure of the criteria



Correlation function



Correlation function

q	K_1	K_2	K_3	K_4
1	0.9987	0.7952	0.7398	0.7026
2	0.9948	0.6354	0.5670	0.4708
3	0.9885	0.5199	0.4507	0.3029
5	0.9683	0.4204	0.3552	0.1574
10	0.8782	0.2948	0.2490	0.0658
15	0.7427	0.1886	0.1582	0.0175
20	0.5807	0.1246	0.1042	0.0081
30	0.2598	0.0541	0.0444	0.0029
40	0.0478	0.0153	0.0128	0.0004

Averaging over N = 1000 realizations

Confidence level

To define the confidence level we should build the maxima distribution

N	$\delta = 2 \cdot 10^{-5}$				$\delta = 2 \cdot 10^{-4}$			
	K_1	K_2	K_3	K_4	K_1	K_2	K_3	K_4
9499	7.4319	24.92	0.4375	67.04	11.1117	25.94	0.4463	68.40
9500	7.4338	24.94	0.4375	67.12	11.1136	25.94	0.4463	68.40
9501	7.4352	24.94	0.4375	67.18	11.1225	25.96	0.4463	68.42
...	...	...	...	...	...	...	...	...
9829	9.7274	27.76	0.4615	76.50	13.6808	28.76	0.4667	77.40
9830	9.7275	27.80	0.4615	76.54	13.6883	28.78	0.4672	77.48
9831	9.7327	27.84	0.4615	76.56	13.6996	28.80	0.4672	77.52
9832	9.7591	27.86	0.4615	76.66	13.7476	28.82	0.4672	77.70
...	...	...	...	...	...	...	...	...
9990	16.8694	36.10	0.5242	100.88	19.6401	36.10	0.5159	99.08
9991	17.3255	37.02	0.5245	103.04	19.9599	36.18	0.5185	99.98
9992	17.4510	37.42	0.5246	103.54	20.7694	36.20	0.5200	100.42
9993	17.4652	38.16	0.5306	103.64	20.7842	36.28	0.5246	102.98
9994	17.5183	38.22	0.5308	105.92	21.3265	37.02	0.5276	103.04
9995	17.5626	39.46	0.5333	106.06	21.9388	37.28	0.5282	104.32
9996	17.5744	39.70	0.5338	107.42	22.1278	37.60	0.5299	105.72
9997	18.5172	42.50	0.5447	107.84	23.0042	37.70	0.5303	108.16
9998	19.3895	43.26	0.5504	110.86	23.5767	38.24	0.5333	112.24
9999	20.4752	48.00	0.5515	116.18	23.6343	39.16	0.5593	113.20
10000	20.7974	48.72	0.5547	129.68	26.0905	39.30	0.5736	114.68

Confidence level

δ	n	$\alpha = 0.05$				$\alpha = 0.01$			
		K_1	K_2	K_3	K_4	K_1	K_2	K_3	K_4
$2 \cdot 10^{-4}$	3	7.110	19.96	0.3952	49.72	10.330	24.92	0.4351	64.32
	10	9.063	23.16	0.4237	59.20	12.770	27.60	0.4597	73.20
	30	11.154	26.16	0.4463	68.50	14.704	30.88	0.4779	82.38
	100	13.526	29.04	0.4681	78.12	17.098	33.14	0.4960	90.88
	500	15.650	32.46	0.4925	91.00	18.997	36.26	0.5194	103.72
	2000	18.307	36.48	0.5188	102.24	21.505	40.62	0.5426	114.76
	5000	22.445	39.60	0.5347	109.60	26.117	43.86	0.5574	122.62
$2 \cdot 10^{-5}$	3	6.144	18.90	0.3871	47.62	9.315	24.24	0.4286	62.20
	10	6.654	21.72	0.4123	57.20	9.854	26.50	0.4511	71.76
	30	7.414	24.44	0.4344	66.20	11.214	29.14	0.4706	80.18
	100	9.172	28.20	0.4632	76.72	12.579	32.80	0.4926	91.36
	500	12.496	32.10	0.4921	89.52	15.922	36.20	0.5203	102.72
	2000	14.839	35.22	0.5124	99.38	17.814	39.32	0.5390	111.62
	5000	15.930	37.52	0.5252	108.60	19.371	41.56	0.5496	121.82

$$T_p = T_1 \left(1+\delta\right)^{p-1}; \qquad p = 1, 2, \ldots, P.$$

Confidence level

δ	b	K_1	K_2	K_3	K_4
$2 \cdot 10^{-4}$	1	232237	21806	14625	10208
	2	562896	343073	469900	262592
	3	795021	767079	932489	690109
	5	964563	988073	999955	974477
	7	994635	999629	1000000	998325
	10	999709	999996	1000000	999980
	14	999997	1000000	1000000	999999
$2 \cdot 10^{-5}$	1	751558	236968	188712	124993
	2	898623	631071	694636	492181
	3	960112	877169	959399	795163
	5	993910	992634	999962	980925
	7	999147	999685	1000000	998636
	10	999958	999997	1000000	999974
	14	999999	1000000	1000000	1000000

N = 1000000 realization, n = 30 trial periods

Generalization to continuous signal

Rayleigh criterion

$$K_1(T) = \frac{\left(\sum_{i=1}^{M} x_i \sin\frac{2\pi t_i}{T}\right)^2 + \left(\sum_{i=1}^{M} x_i \cos\frac{2\pi t_i}{T}\right)^2}{\left(\sum_{i=1}^{M} x_i\right)^2},$$

Other criteria analyze the light curve and have the same form

What do the criteria really extract?

We do not know what is a periodicity exactly because the standard mathematical definition

$$f(T) = f(T+t)$$

is not good and cannot be applied to real objects.

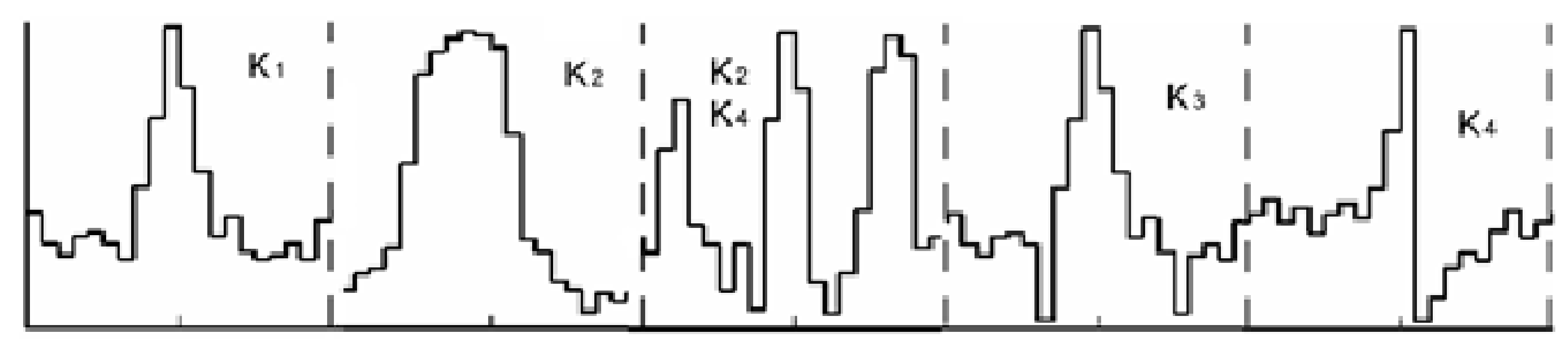
The criteria reveal not the periodicity itself, but its external manifestations, such as non-smooth light curve.

If the phase is not a linear function of a time we should use the time transformation

$$t = g(t'),$$

so that in new time variable t' the process would be pure periodic.

What do the criteria really extract?



The light curves which are good detected by different periodicity criteria.

Conclusions

- We know statistical properties of at least four criteria of periodicity.
- We can apply these criteria to real objects and correctly define the confidence levels.
- We know the individual properties of the criteria
- We can apply the criteria to both discrete and continuous signal