Statistics

Experiments details:

In this section we compare performance of some of our statistical routines with matlab functions. In the following table you can see what methods were tested and what matlab's functions were chosen as competitor:

FinMath	Matlab
StandardDeviationSample(Vector)	std(v), with default parameters.
GetTop(Vector, count)	w = sort(v); w = w(1:count) this typical way to do basketing in matlab.
Quantile(Vector, pvalue)	quantile(v, pvalue) , with default parameters.
Correlation(Matrix)	corr(X) , with default parameters.

1. Test calculation of sample standard deviation of the data vector. Complexity of this operation depends only on number of samples in data vector.
2. Test getting K smallest elements from data vector. Performance of this operation depends on number of samples in data vector and value of K.
3. Test getting quantile corresponding to pvalue. In order to choose some reasonable pvalue, we select integer number from 0 to number of samples in data vector and divide it by vector size. So this operation performance depends on integer selected and data vector length.
4. Test computation of data set covariance matrix. Depends on number of objects or series in data set (N) and number of factors or observation in single series.

Performance metrics:

In this test we use the following performance metrics:

1. Total working time – time which was needed to complete specified numbers of computations.
2. FLOPS – approximated number of floating point operations number per second. Typically we use mega flops (MFLOPS which) is $FLOPS \times 10^6$. In statistics test we calculate this metric only for correlation test. Denote repeats count as R. So we can write complexity of calculation correlation matrix in asymptotic form as $O(N^2MR)$.. We will use this expression as approximation of floating point calculations count (FLOPS). So for example if we have 250 objects, 10 factors and we make 10000 runs in 15 seconds we will say that method performance is:
$$\frac{N^2MR}{computation\ time} = \frac{250^2 * 10 * 10000}{15} = 416666666 \approx 416\ MFLOPS$$
3. Time of single computation – time which was needed to solve one problem with selected parameters.

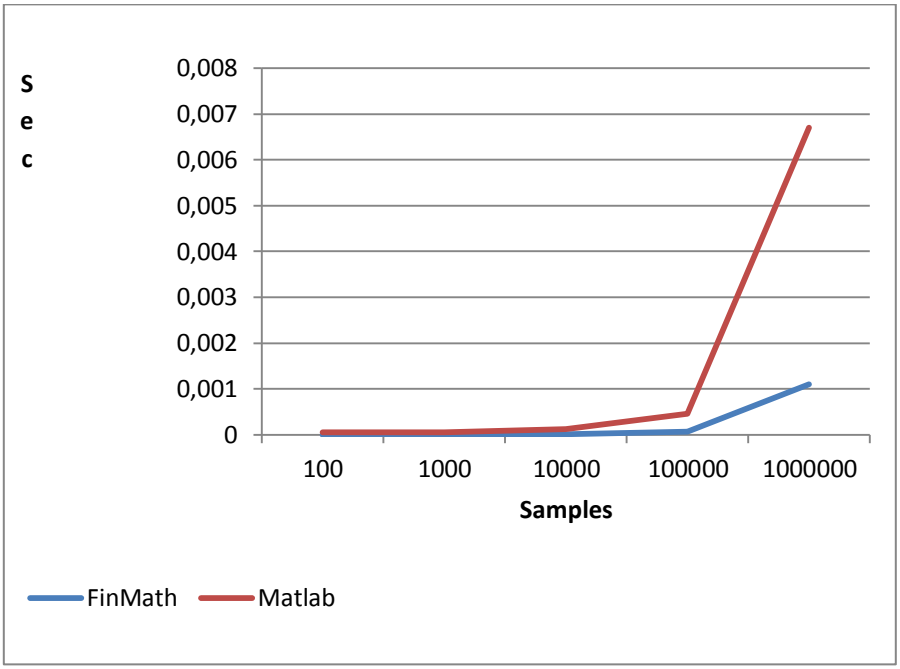
Testing system details:

All tests were performed on typical medium class desktop machine.

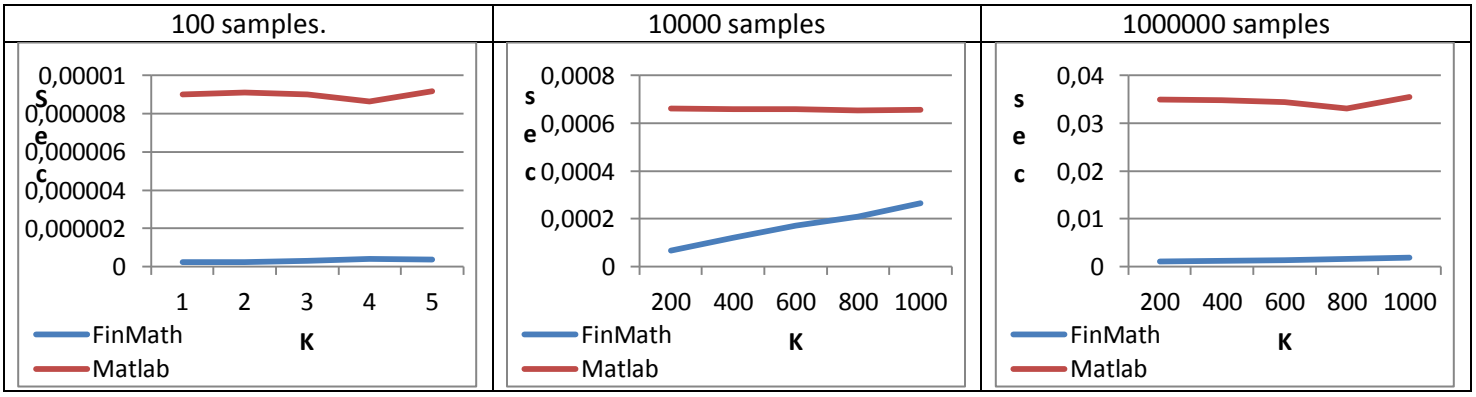
CPU: Core i7 870 (2.93GHz)
Memory: 8GB RAM
OS: Windows 7 x64

All source code of this test are available, one can get performance results on his own system. Changing test parameters is also possible.

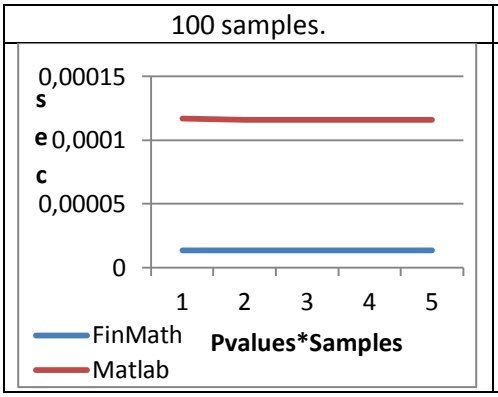
Standard Deviation					
Samples	Repeats	FinMath		Matlab	
		Total Time	Single Time	Total Time	Single Time
100	100000	0.021	2.10E-07	5.17	5.17E-05
1000	10000	0.006	6.00E-07	0.59	5.90E-05
10000	1000	0.006	0.000006	0.125	0.000125
100000	100	0.007	0.00007	0.046	0.000464
1000000	10	0.011	0.0011	0.067	0.006702

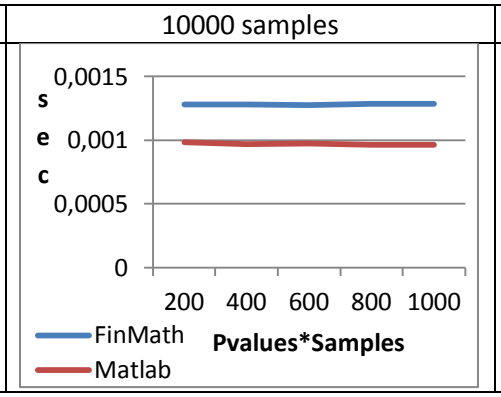


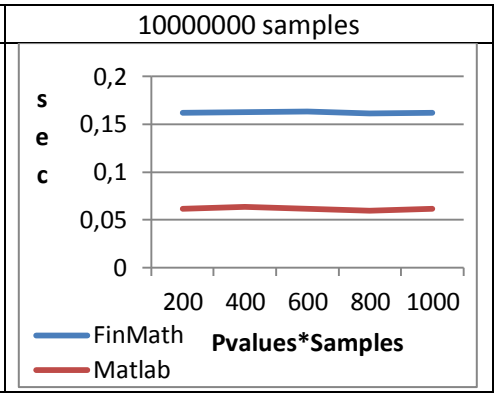
Get k smallest elements						
Samples	K	Repeats	FinMath		Matlab	
			Total Time	Single Time	Total Time	Single Time
100	1	100000	0.024	2.40E-07	0.901	9.01E-06
100	2	100000	0.022	2.20E-07	0.911	9.11E-06
100	3	100000	0.03	3.00E-07	0.9	0.000009
100	4	100000	0.039	3.90E-07	0.864	8.64E-06
100	5	100000	0.036	3.60E-07	0.918	9.18E-06
10000	200	1000	0.068	0.000068	0.663	0.000663
10000	400	1000	0.12	0.00012	0.658	0.000658
10000	600	1000	0.17	0.00017	0.658	0.000658
10000	800	1000	0.21	0.00021	0.654	0.000654
10000	1000	1000	0.265	0.000265	0.656	0.000656
1000000	200	10	0.011	0.0011	0.349	0.034921
1000000	400	10	0.012	0.0012	0.349	0.034852
1000000	600	10	0.013	0.0013	0.345	0.03447
1000000	800	10	0.016	0.0016	0.33	0.033048
1000000	1000	10	0.019	0.0019	0.355	0.035471



Get k smallest elements						
Samples	Pvalues*Samples	Repeats	FinMath		Matlab	
			Total Time	Single Time	Total Time	Single Time
100	1	100000	1.34	1.34E-05	11.678	1.17E-04
100	2	100000	1.345	1.35E-05	11.613	1.16E-04
100	3	100000	1.342	1.34E-05	11.561	0.000116
100	4	100000	1.352	1.35E-05	11.618	1.16E-04
100	5	100000	1.34	1.34E-05	11.561	1.16E-04
10000	200	1000	1.281	0.001281	0.983	0.000983
10000	400	1000	1.279	0.001279	0.967	0.000967
10000	600	1000	1.274	0.001274	0.977	0.000977
10000	800	1000	1.286	0.001286	0.962	0.000962
10000	1000	1000	1.288	0.001288	0.965	0.000965
1000000	200	10	1.621	0.1621	0.613	0.061266
1000000	400	10	1.63	0.163	0.639	0.063869
1000000	600	10	1.635	0.1635	0.614	0.061365
1000000	800	10	1.615	0.1615	0.599	0.059901
1000000	1000	10	1.619	0.1619	0.613	0.06135







Correlation								
Factors	Objects	Repeats	FinMath			Matlab		
			Total Time	MFLOPS	Single Time	Total Time	MFLOPS	Single Time
100	10	500	0.008	6250	0.00002	1.412	36	0.003
100	20	500	0.009	11112	0.00002	4.56	22	0.009
100	30	500	0.017	8824	0.00003	9.905	16	0.02
100	50	500	0.037	6757	0.00007	26.965	10	0.054
100	100	500	0.104	4808	0.00021	106.003	5	0.212
1000	10	125	0.006	208334	0.00005	0.59	2120	0.005
1000	100	60	0.063	95239	0.00105	24.679	244	0.411
1000	200	20	0.058	68966	0.0029	32.645	123	1.632
1000	500	10	0.122	40984	0.0122	101.804	50	10.18
1000	1000	5	0.286	17483	0.0572	204.085	25	40.817
5000	10	15	0.003	1250000	0.0002	0.23	16296	0.015
5000	100	15	0.045	833334	0.003	21.634	1734	1.442
5000	500	3	0.153	245099	0.051	106.621	352	35.54
5000	1000	2	0.425	117648	0.2125	280.601	179	140.301
5000	1500	1	0.439	85422	0.439	353.3	107	353.3