THE CALCULATION OF ENTROPY OF FINANCIAL TIME SERIES

Abstract: *A number of variations of calculation of entropy have been developed to measure the regularity of time series. The article deals with the calculation of approximate entropy used in physic. Some measures of time series can be made in financial branch. The article presents the used method of calculation programed in MATLAB. The case study is dedicated to evaluation of indexes of eastern states, such as Poland, Hungary, Estonia, Lithuania, Slovenia and Czechia.*

1. Introduction

The calculation of approximate entropy is possible to use in case of evaluation of time series including economy and financial ones. The value of entropy is „high"(approximate to 1) in case when the time is randomness. The value of entropy is „low"(approximate to 0) in case of deterministic time series (for example sinusoidal wave). It means that the entropy inform us about the level of determinism or randomness of time series and thus the "stability" of phenomena "producing" such time series, for example stock market, state economy etc. The calculation of entropy is described in literature [3,4]. The article presents the used method of calculation programed in MATLAB. The case study is dedicated to evaluation of indexes of eastern states, such as Poland, Hungary, Estonia, Lithuania, Slovenia and Czechia. The other possibilities of measure of time series are by Hurst exponent. See lit. [1].

2. Used Methods

The calculation is as follows. The parameters r and m are the filter and the size of the template vector one. Time series with N observations X_m, $n = 1, \ldots, N$ create embedding vector v_i, each made up of m consecutive values of X, $v(i) = [X_i, \ldots, X_{i+k}]$ with $k = 1, \ldots, m - 1$. For $i, j = 1, \ldots, N - m + 1$, each template vector $v(i)$ is compared to all conditioning vectors $v(j)$ (including itself) by computing L_1 norm. According to index k, the distance between the vector elements is measured sequentially according the formula

$$d[v(i), v(j)] = \max_{k=0,\ldots,m} \left| X_{i+k} - X_{j+k} \right|.$$

Quantification of the regularity of a particular pattern, count the relative frequency of distances between the template vector $v(i)$ to all vectors $v(j)$, that lie within the neighborhoods r. The formula is as follows

$$C_i^m(r) = \frac{1}{N - m + 1} \sum_{j=0}^{N-m-1} \Theta(r - d[v(i), v(j)]),$$

where $\Theta(\cdot)$ represents the binary Heaviside function, where $\Theta(\cdot) = 1$ if the condition $r > L_1$ is satisfied, else $\Theta(\cdot) = 0$. The average over i of the log amount of vectors $v(j)$ for which $r > L_1$ is calculated according formula

$$\phi_i^m(r) = \frac{1}{N - m + 1} \sum_{i=0}^{N-m-1} \ln C_i^m(r).$$

Finally, the conditional probabilities and ensure positivity is calculated according formulas

$$Ep(r, m, n) = \phi^m(r) - \phi^{m+1}(r) \quad \text{and} \quad Ep(r, 0, n) = -\phi^1(r).$$

3. The calculation programs

The fixed parameter is *r*. The program *Entropy.m* programmed in MATLAB do the calculation of value of entropy. See prog.1.

```
clear all
EN=(xlsread('DostalEn','Data'));
[ro,co]=size(EN);
for i=1:6
pp=[EN(1:ro,i)];
pot=[EN(1:ro,i)];
subplot(2,3,1)
if i==1 plot(pp,'Color',('red'));end;
subplot(2,3,2)
if i==2 plot(pp,'Color',('green'));end;
subplot(2,3,3)
if i==3 plot(pp,'Color',('black'));end;
subplot(2,3,4)
if i==4 plot(pp,'Color',('cyan'));end;
subplot(2,3,5)
if i==5 plot(pp,'Color',('blue'));end;
subplot(2,3,6)
if i==6 plot(pp,'Color',('yellow'));end;
r=.5;
 [n,p] = size(pp);
a = 0;
b = 0;
c = zeros(n,p);
for k=1:n
  for j=1:p
   e(:,j) = pp(k,j);
  end
  d = (abs( e - pp ) <= r );
  if p == 1
    cpp = d;
  else
    cpp = all(d');
  end
  pc = sum(cpp);
  a = a + log(pc);
  inds = find(cpp);
  pt = sum( abs( pot(cpp) - pot(k) ) < r );
  b = b + log(pt);
end
ep(i) = (a - b)/n;
end
ep
```

Prog.1 Program for calculation of entropy

4. Case study

The case study presents the calculation of entropy of time series of six indices of Hungary-Budapest (BUX price index), Lithuania-Riga (HSBC price index), Estonia-Tallinn (OMXT price index), Czechia-Prague SE PX price index), Slovakia-Bratislava (SAX 16 price index)

and Poland-Warsaw (General index price index). See fig.1 from left top to right down. The data were taken from 25.2.2000 to 25.2.2010.
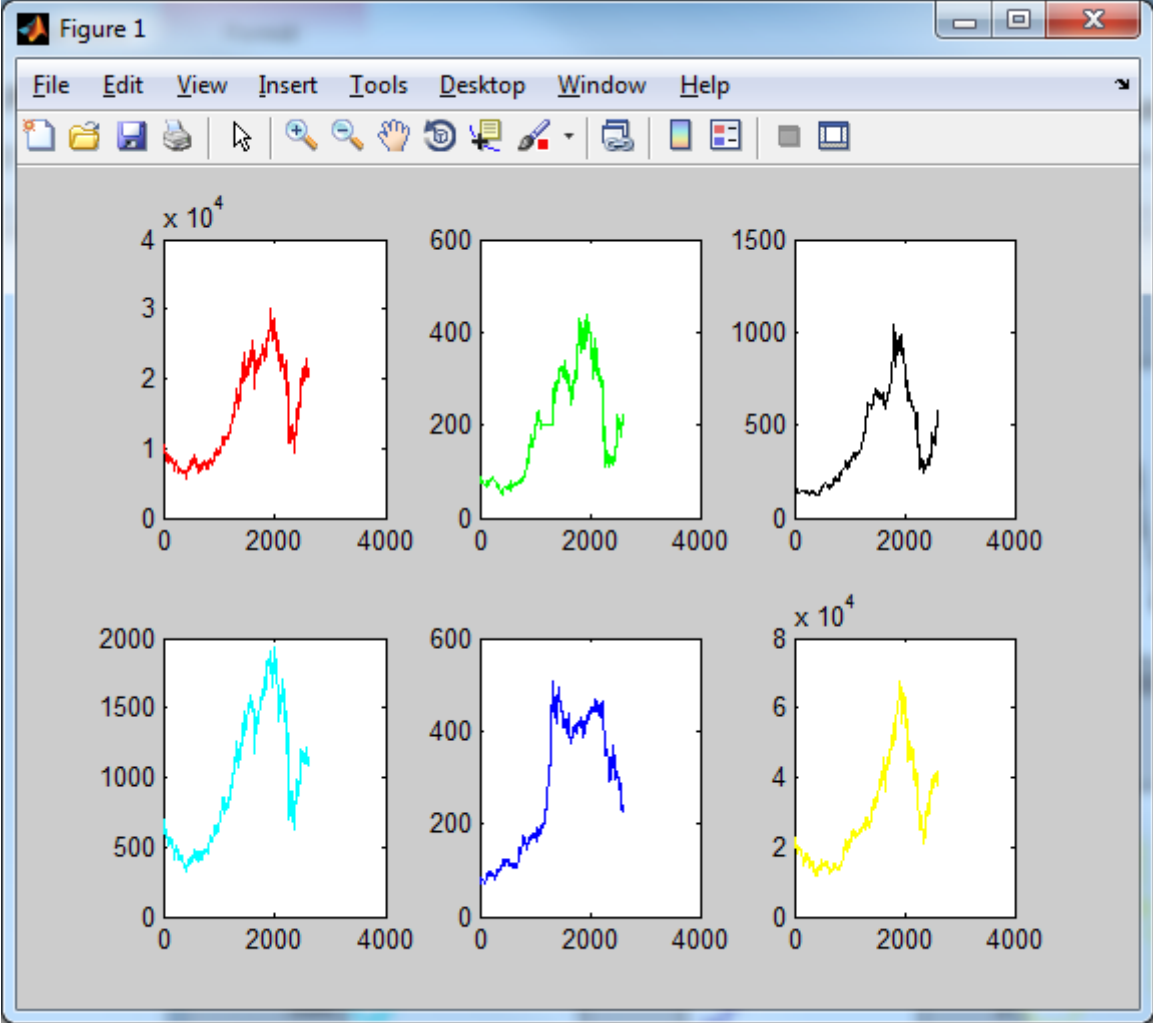

Fig.1 Searched time series

The results of entropy after calculation are as follows. See res.1.

ep =	0.0015	0.0045	0.0043	0.0463	0.0036	0.0033
	Hungary	Lithuania	Estonia	Czechia	Slovakia	Poland

Res.1 Values of calculated entropy

The case study represents the indices from the more stable to less one Hungary, Poland, Slovakia, Estonia, Lithuania, Slovakia and Czechia. The results can help the decision makers to solve the problem in which country to invest. Further search will be done to compare results with Hurst exponent.

5. Conclusion

The conclusion is as follows: Approximate entropy measures the relative frequency which blocks of length *m* that is close and remain close together in the next incremental comparison *m+1*. Small values of calculated entropy imply strong regularity or determinism, in a time

series; large value of entropy implies substantial irregularity or randomness. Approximate entropy has also been found to be a robust measure of regularity when applied to time series that contain noise and outliners. A number of variations of calculation of entropy have been developed but the general principle of measuring of regularities of time series is similar. The approximate entropy metric do not assume that only obvious pattern such as seasonality and trend are important, bur revel any form of regularity. Nor are they distorted when the series means is close to zero.

The use of calculation of approximate entropy will be further searched for the support of decision making in which country to invest.