



9/2 (2011), 261–274

tmcs@math.klte.hu
http://tmcs.math.klte.hu

Teaching
Mathematics and
Computer Science

Longest runs in coin tossing. Teaching recursive formulae, asymptotic theorems and computer simulations

ZSOLT KARÁCSONY* and JÓZSEFNÉ LIBOR

Abstract. The coin tossing experiment is studied, focusing on higher education. The length of the longest head run can be studied by asymptotic theorems ([3]), by recursive formulae ([10]) or by computer simulations. In this work we make a comparative analysis of recursive formulas, asymptotic results and Monte Carlo simulation for education. We compare the distribution of the longest head run and that of the longest run (i.e. the longest pure heads or pure tails) studying fair coin events. We present a method that helps to understand the concepts and techniques mentioned in the title, which can be a useful didactic tool for colleagues teaching in higher education.

Key words and phrases: run, recursive formula, asymptotic theorem, simulation.

ZDM Subject Classification: K50, K60, B40.

1. Introduction

Recursive formulae, asymptotic theorems and simulations play an important role in several fields of math-teaching in higher education. But the exact knowledge of these notions leaves a lot to be desired. The study of some illustrative

*This research was carried out as part of the TAMOP-4.2.1.B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund.

problems may help to understand them. While solving these problems students have an opportunity to compare these definitions. In this paper we shall use the well-known coin tossing experiment to compare the results of recursive formulae, asymptotic theorems and computer simulations. In connection with coin tossing we examine the length of the longest run. For a sequence of independent coin tosses with p ($0 \leq p \leq 1$), the longest run of consecutive heads in the first n tosses is a natural object of study. In our work the longest run means the longest sequence of the same objects, for example the longest head-run means the longest sequence among the sequences of homogeneous heads, which is not interrupted by tail. Our research includes the examination of the length of the longest head-run and the length of the longest whatever – head or tail – run, studying fare coin.

The following well-known example by T. Varga can be an interesting introduction to the problem for students. We can read this for example in Révész [8]. A class of school children is divided into two sections. In one of them each child is given a coin which he throws two hundred times, recording the resulting head and tail sequence on a piece of paper. In the other section the children do not receive coins but are told instead that they should try to write down a ‘random’ head and tail sequence of length in 200. Collecting and mixing all the slips of paper, the teacher then tries to subdivide them into their original groups. Most of the time he succeeds quite well. His secret is that he has observed the following. In a randomly produced sequence of length in two hundred, there are, say, head-runs of length seven (knowing Rényi’s $\log_2 200$ result). On the other hand, he has also observed that most of those children who had to write down an imaginary random sequence are usually afraid of writing down runs of longer than four. Hence, in order to find the slips coming from the coin tossing group, he simply selects the ones which contain runs longer than five. When we had the privilege of meeting professor Révész, he talked about the continuation of Varga’s experiment. He made his students acquainted with Varga’s experiment and the result of it. Then they made the trial again. Professor Révész was successful in subdividing the slips of paper with quite good accuracy again. His secret was very

easy. The students focused on the length of pure head run for example, but they did not pay attention to the length of pure tail run or the length of the head-tail pairs run. These experiments led us to ask the following questions. What is the length of the longest head run or the longest whatever run in coin tossing?

For the investigation of the mentioned definitions (recursive formula, asymptotic theorem, simulation) our chosen field – coin tossing – is perfect in many respects. We know that there is no exact and closed formula to give the length of the longest run in case of any n . The recursive formulae are exact, – they give correct results – but using them for large n , is very time- and storage-consuming even if we use very good computer. For large n we can use the asymptotic theorems, but these give only approaching results. And finally, the simulation gives an average value from a lot of repetitions of the trial. Considering these, we can examine the meaning of these definitions by studying the coin tossing trial. Our graphs will show the differences between the values in case of small and large n too and the limits of the scope of the notions.

Consider n independent tosses of a fair coin, and let R_n represent the length of the longest run of heads, and similarly let R'_n be the length of the longest whatever (head or tail) run. As we have got a regular (fair) coin, the probability of head-tossing equals $p = 0.5$, and of course the probability of tail-tossing is equal to $q = 1 - p = 0.5$. The problem facing us is the following. What is the length of the longest head or whatever run?

2. Recursive formulae

Sometimes it is difficult to define an object explicitly. However, it may be easy to define the object in terms of itself. The recursive functions, which form a class of computable functions, take their name from the process of 'recurrence' or 'recursion'. In its most general numerical form the recursion process consists in defining the value of a function by using other - the previous - values of the same function. A recursive sequence is a sequence of numbers $f(n)$ indexed by

an integer n and generated by solving a recurrence equation. First we have to give the initial condition that tells where the sequence starts, then the recursion formula that tells how any term of the sequence relates to the preceding term.

And now back to our coin tossing problem.

2.1. Longest head run

Let R_n represent the length of the longest run of heads and let $A_n(x)$ be the number of sequences of length n in which the longest head run does not exceed x . The (cumulative) distribution function of R_n is

$$F_n(x) = P(R_n \leq x) = \frac{A_n(x)}{2^n}. \tag{1}$$

Using the results of Schilling [10], we have the following recursive formula for $A_n(x)$

$$A_n(x) = \begin{cases} \sum_{j=0}^x A_{n-1-j}(x), & \text{if } n > x, \\ 2^n, & \text{if } 0 \leq n \leq x. \end{cases} \tag{2}$$

To see how this works, consider the case in which the longest head run consists of at most four heads. If $n \leq 4$, then clearly $A_n(4) = 2^n$ since any outcome is a favorable one. For $n > 4$, each favorable sequence begins with either T, HT, HHT, HHHT or HHHHT and is followed by a string having not more than four consecutive heads. Thus, $A_n(4) = A_{n-1}(4) + A_{n-2}(4) + A_{n-3}(4) + A_{n-4}(4) + A_{n-5}(4)$ for $n > 4$. From this we can calculate the values of $A_n(4)$:

n	0	1	2	3	4	5	6	7	8	...
$A_n(4)$	1	2	4	8	16	31	61	120	236	...

REMARK 1. For $n = 1, 2, 3, \dots$, the number $A_n(1)$ of sequences of length n where n does not continue two consecutive heads is the $(n + 2)$ nd Fibonacci number.

REMARK 2. The values of $A_n(k)$ can be given with the help of k^{th} degree Fibonacci numbers. Moreover, we can study the case of biased coins using the application of k^{th} degree Fibonacci polynomials. (See Philippou–Makri, [6].)

2.2. Longest whatever run

Consider n independent tosses of a fair coin, and let R'_n represent the length of the longest run of heads or tails (whatever). If $B_n(x)$ is the number of sequences of length n in which the longest run does not exceed x , then the (cumulative) distribution function is $F'_n(x) = P(R'_n \leq x) = \frac{B_n(x)}{2^n}$. Schilling [10] has proved that $B_n(x) = 2A_{n-1}(x - 1)$, for $x \geq 1$.

To prove it consider a sequence of length n consisting of signs H and T (representing ‘head’ and ‘tail’, respectively). Write below of this sequence another sequence of signs S and D (representing ‘same’ and ‘different’, respectively). We write S if two consecutive signs in the first sequence are the same, and write D if they are different. Now a run of length $x - 1$ consisting of signs S represents a run of length x in the first sequence. We also see that any sequence containing signs S and D belongs to two sequences of signs H and T. Finally, observe that $A_{n-1}(x - 1)$ can be considered as the number of sequences of length $n - 1$ and containing signs S and D but not having S-run of length x .

The example below shows one of the strings that contributes to $B_{13}(4)$:

H	H	H	T	H	T	H	T	T	T	T	H	H
S	S	D	D	D	D	D	S	S	S	D	S	

So we can reduce this case to the case of the longest head run

$$F'_n(x) = P(R'_n \leq x) = \frac{B_n(x)}{2^n} = \frac{2A_{n-1}(x - 1)}{2^n} = \frac{A_{n-1}(x - 1)}{2^{n-1}} = F_{n-1}(x - 1). \quad (3)$$

The implication of (3) is that for n tosses of a fair coin the longest run tends to be one longer than the longest run of heads alone.

3. Asymptotic theorems

First we recall two well-known asymptotic results of probability theory. The law of large numbers asserts that the arithmetical mean of independent observations converges to the expectation as the number of observations tends to infinity. In statistics the consistency of estimators is based on the law of large numbers. An important particular case is the Bernoulli law of large numbers, i.e. the relative frequency converges to the probability. The formal statement is the following (Kolmogorov’s strong law of large numbers (SLLN)). Let X_1, X_2, \dots , be independent identically distributed (i.i.d.) random variables with $E|X_i| < \infty$ and $E(X_i) = m$. Then $\frac{X_1+X_2+\dots+X_n}{n} \rightarrow m$ almost surely.

The second basic result is the central limit theorem (CLT). It means that the sequence of the standardized partial sums converges to the standard normal distribution. In statistics the asymptotic normality of estimators (e.g. the maximum likelihood estimator) and the asymptotic distribution of certain statistics (e.g. χ^2 statistic) are based on CLT’s. The formal statement is the following. Let X_1, X_2, \dots be i.i.d. random variables, let $0 < \sigma^2 = D^2(X_i) < \infty, E(X_i) = m$. The distribution of $\frac{X_1+X_2+\dots+X_n-nm}{\sqrt{n}\sigma}$ converges to the standard normal law.

For the longest head run we have both strong limit theorems and results on the limiting behavior of the distribution. A well-known theorem of Rényi says that the length of the longest run in n tosses of a fair coin is about $\log_2 n$. This was generalized to excessive blocks by the Erdős–Rényi [2] laws of large numbers.

For the limiting behavior of the distributions there are several known results. Let us see the case of the longest head run. The asymptotic behavior of R_n is described by the following theorem.

THEOREM 1 (Földes (1979) [3]). *For any integer k we have*

$$P(R_n - \lfloor \log_2 n \rfloor < k) = \exp\left(-2^{-(k+1-\{\log_2 n\})}\right) + o(1), \quad (4)$$

where $\lfloor a \rfloor$ denotes the integer part of a and $\{a\} = a - \lfloor a \rfloor$.

Similarly we can see the asymptotic behavior of R'_n . Using the above theorem of Földes and (3) we have the following asymptotic result for R'_n .

THEOREM 2. *For any integer k we have*

$$P(R'_n - \lfloor \log_2(n-1) \rfloor < k) = \exp\left(-2^{-(k-\{\log_2(n-1)\})}\right) + o(1). \quad (5)$$

where $\lfloor a \rfloor$ denotes the integer part of a and $\{a\} = a - \lfloor a \rfloor$.

We see that both of the above theorems are different from the CLT. We have no asymptotic normality. Moreover, we have no fixed limiting distribution. The only possibility is to construct a sequence of 'accompanying laws'. That is the distribution of $R_n - \lfloor \log_2 n \rfloor$ close to a certain sequence as $n \rightarrow \infty$, i.e. $P(R_n - \lfloor \log_2 n \rfloor < k) - \exp(-2^{-(k+1-\{\log_2 n\})}) \rightarrow 0$ as $n \rightarrow \infty$. The periodic behavior of $R_n - \lfloor \log_2 n \rfloor$ is presented in the following figure (the numerical calculation behind it is based on the recursion (1) and (2)).

4. Simulation

Stochastic simulation is a well-known method of numerical mathematics. This method is also called the Monte Carlo method after the city in the Monaco Principality, because of roulette, a simple random number generator. The name and the systematic development of this method dates from about 1944. Simulation methods can play an important role in teaching several topics not only mathematics.

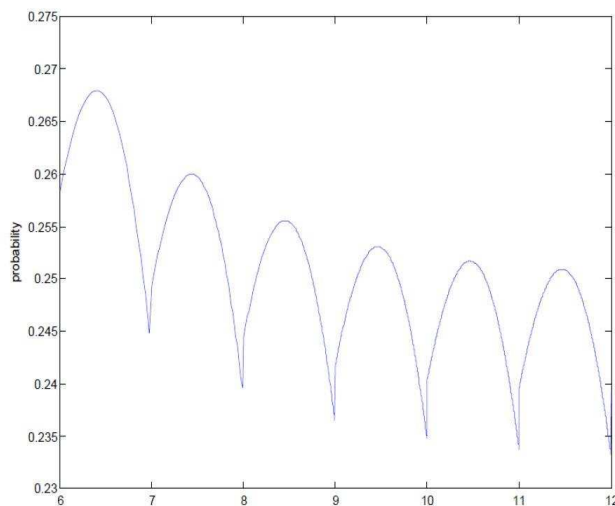
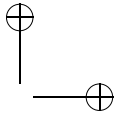
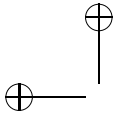


Figure 1. $P(R_n = \lceil \log_2 n \rceil - 1)$, $n = 2^6, \dots, 2^{12}$, $p = 0.5$

“The main idea of the law of large numbers - the larger the sample size, the greater the likelihood that a specific experimental result will be closer to the theoretical one - is misunderstood by pupils. Teaching through simulations can catalyze comprehension of existing concepts among pupils and develop their stochastic reasoning. Through simulations, it is possible to develop an understanding of particularly elementary concepts, and to learn that drawing conclusions should be based on large samples, whereas small samples usually lead to the wrong conclusions.” (See [1]) “The Monte Carlo method provides approximate solutions to a variety of mathematical problems by performing statistical sampling experiments on a computer. The method applies to problems with no probabilistic content as well as to those with inherent probabilistic structure.” (See [7]) Applying the popular Markov Chain Monte Carlo (MCMC) method we can calculate certain hard-to-compute quantities (in say physics).

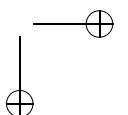
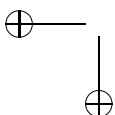
Here we deal with the simplest version of the stochastic simulation. We need the numerical value of a quantity p . Moreover, p can be considered as the probability of an event A : $P(A) = p$. Repeat the experiment N times and

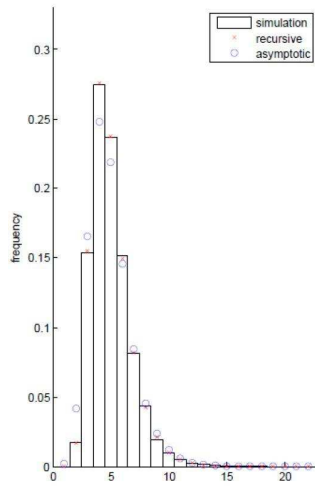


calculate the relative frequency of A . Denote it by $\frac{k_A}{N}$. Then, by the Bernoulli law of large numbers, $\frac{k_A}{N} \rightarrow p$ almost surely as $N \rightarrow \infty$. However, usually we make our experiment on a computer. Using our computer programme, we build the experiment, and we record if A occurs or not. Repeat the computer experiment N times. It means that we run our programme N times with N consecutive random numbers. However, as the coin tossing experiment is a random experiment itself, we develop our computer model as follows. Using the first n random numbers, we can perform the coin tossing experiment of length n . Then repeat it N times (N is a large integer, $N = 1000$, say). Therefore, for any fixed n , the distribution of R_n can be obtained as $N \rightarrow \infty$. (It is an important issue, because in the previous section we considered the asymptotic distribution of R_n , as $n \rightarrow \infty$.)

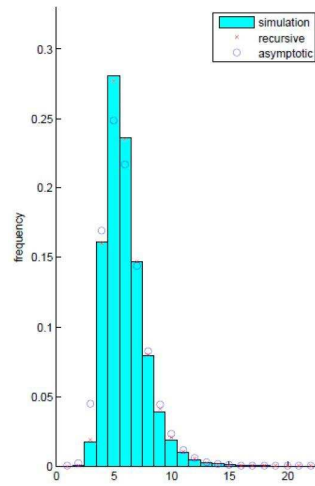
5. Comparison of numerical results

Now let us see our problem. For numerical calculations we used MATLAB software and a high-capacity PC (INTEL Core Quad Q9550 processor, 4Gb. DDR3 memory). We calculated the distribution of R_n and R'_n . We considered the precise values obtained by recursions, the asymptotic values offered by asymptotic theorems, and used simulation with 20,000 repetitions. On the figures below \times denotes the result of the recursion, \circ belongs to the asymptotic result, while the histogram shows the relative frequencies calculated by simulation. On the left side you can see the case of the longest head run and on the right side you can see the case of the longest whatever run.

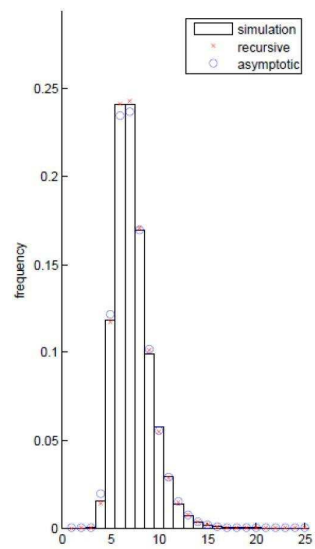




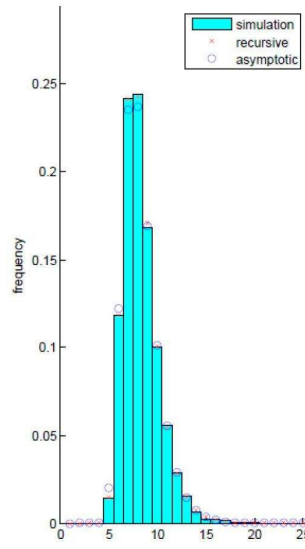
(a) Distribution of the longest head run $p = 0.5, n = 50$



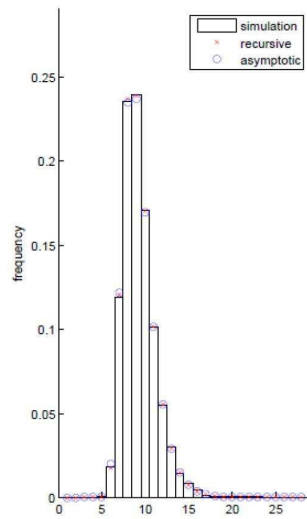
(b) Distribution of the longest head run $p = 0.5, n = 50$



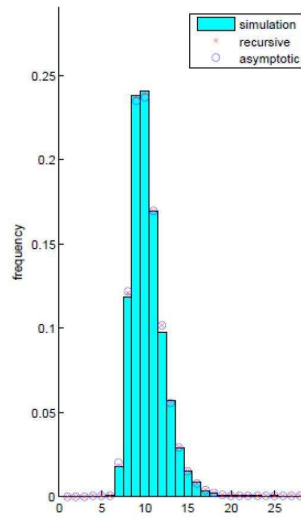
(c) Distribution of the longest head run $p = 0.5, n = 250$



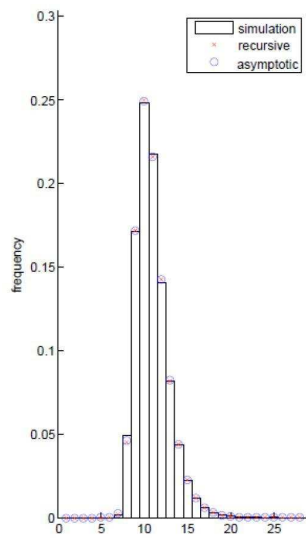
(d) Distribution of the longest head run $p = 0.5, n = 250$



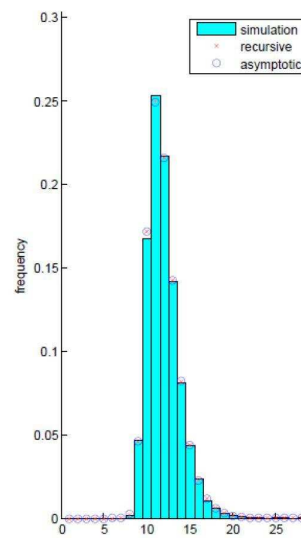
(e) Distribution of the longest head run $p = 0.5, n = 1000$



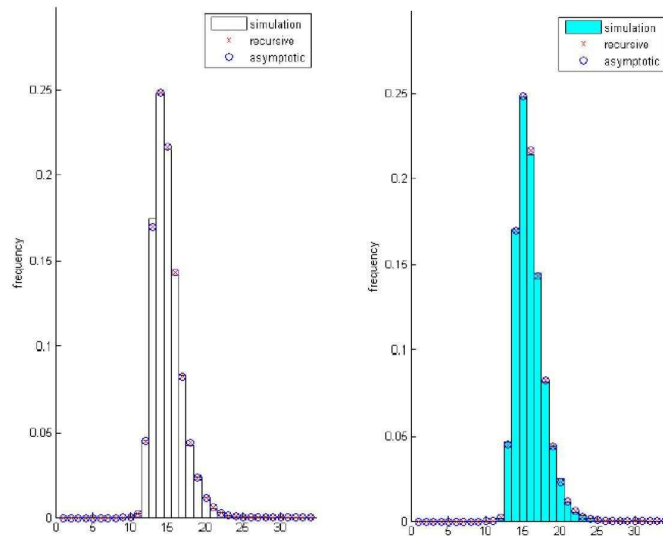
(f) Distribution of the longest head run $p = 0.5, n = 1000$



(g) Distribution of the longest head run $p = 0.5, n = 3100$



(h) Distribution of the longest head run $p = 0.5, n = 3100$



(i) Distribution of the longest head run $p = 0.5, n = 50000$

(j) Distribution of the longest head run $p = 0.5, n = 50000$

6. Conclusions

First we emphasize that the precise values of the distributions are given by the recursive formulae. The practical drawbacks of recursive formulae are the following. They are limited by the capacity of the computer and they can be influenced by numerical errors. On the other hand, asymptotic theorems always offer approximations but they are easy to compute. The approximations are quite bad for a small n but they are almost precise for a large n . Finally, simulation always gives random approximations but it is quite precise if the number of repetitions is large. The above figures show the properties of the procedures mentioned. For a fair coin we studied short trials ($n = 50$), medium size trials ($n = 250$), long trials, ($n = 1000$ and $n = 3100$) and very long trials ($n = 50,000$).

If n is small or medium sized, we can see that the recursive results are closer to the simulated values, than the results given by asymptotic theorems. Although we can say that if n is small, the recursive algorithm is fast, but it slows down if

n increases. The next table demonstrates this statement showing some running times.

n	repet.	running time
100,000	20,000	773.575832 s.
10,000	20,000	31.795056 s.
5,000	20,000	14.398654 s.
3,100	20,000	9.009479 s.
1,000	20,000	3.984981 s.
500	20,000	3.240019 s.
250	20,000	2.412824 s.
50	20,000	2.092010 s.
30	20,000	1.935557 s.

For a large n the asymptotic values are closer to the simulated results, so we can use them instead of the recursive values. The asymptotic value is a good approximation if $n \geq 1000$, and it is practically precise if $n \geq 10,000$.

The figures show that the distribution of R'_n can practically be obtained from the distribution of R_n by shifting it to the right by 1, as we have seen in (3).

References

- [1] M. Amit and I. Jan, *Autodidactic learning of probabilistic concepts*, Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education, Vol. 2, Prague: PME, 49–56.
- [2] P. Erdős and P. Révész, On a new law of large numbers, *J. Analyse Math.* **23** (1970), 103–111.
- [3] A. Földes, The limit distribution of the length of the longest head-run, *Period. Math. Hungar.* **10**, no. 4 (1979), 301–310.
- [4] N. Metropolis, *The Beginning of the Monte Carlo Method*, Los Alamos Science, Special Issue, 1987.
- [5] N. Metropolis and S. Ulam, The Monte Carlo Method, *Journal of the American Statistical Association* **44**, no. 247 (Sep. 1949), 335–341.

- [6] A. N. Philippou and F. S. Makri, Longest success runs and Fibonacci-type polynomials, *The Fibonacci Quarterly*, no. 23 (1985).
- [7] S. Pillana, History of Monte Carlo Method,
<http://stud2.tuwien.ac.at/e9527412/history.html>.
- [8] P. Révész, Regularities and irregularities in a random 0,1 sequence, *Statistical Papers* **31** (1990), 95–101.
- [9] J. A. Rochowicz, Learning Binomial Probability Concepts With Simulation, Random Numbers and a Spreadsheet, 2005-07-26,
<http://www.toodoc.com/recursive-algorithm-in-discrete-mathematics-ebook.html>.
- [10] M. F. Schilling, The longest run of heads, *College Math. J.* **21**, no. 3 (1990), 196–207.
- [11] I. Stahl, *Teaching the Classics of Simulation to Beginners*, Proceedings of the 2003 Winter Simulation Conference, 1941–1951.

ZSOLT KARÁCSONY
DEPARTMENT OF APPLIED MATHEMATICS
UNIVERSITY OF MISKOLC
H-3515 MISKOLC-EGYETEMVÁROS
HUNGARY

E-mail: matkzs@uni-miskolc.hu

JÓZSEFNÉ LIBOR
DEPARTMENT FOR METHODOLOGY OF ECONOMIC ANALYSIS
UNIVERSITY COLLEGE OF SZOLNOK
H-5000 SZOLNOK, TISZALIGETI SÉTÁNY
HUNGARY

E-mail: liborne@szolf.hu

(Received January, 2011)