Slovak University of Technology in Bratislava
Faculty of Mechanical Engineering
17th Conference on Applied Mathematics **APLIMAT 2018**

Proceedings

POSSIBLE COMPARISON BETWEEN TWO TIME SERIES

ŤOUPAL Tomáš (CZ), ŠEDIVÁ Blanka (CZ)

Abstract. This paper is focused on the possible methodologies for comparing two time series by estimating the probability that both time series will increase or decrease (of course in probability) in the same time.
It can be considered as the specific measure of dependence (more precisely concordance) between two random variables with non-parametric approach. Then the problem may arise with a statistical inference. The main idea of this approach is based on the transformation of observed data set into increasing and decreasing movement and then the Markov model (Markov chain) of transitions (increasing-increasing, increasing-decreasing, decreasing-increasing, decreasing-decreasing) is used with the removal of the assumption of independence.

Keywords: time series, Markov model, exchange rate, Czech crown, coefficient of concordance, dependence

Mathematics Subject Classification: Primary 62M05, 91B84; Secondary 62M10

1 Introduction

A frequent financial problem in the world is the assessment of how the financial assets denominated in some currency will behave in the currencies of that currency to the accounting currency (exchange rate loss or yield). This can be assessed on the base of knowledge of the set of past values and the assumption that there will be no significant change in the used probability model of the behaviour of such time series.
In general, it is a measure of the "power" of keeping or not keeping tendencies (monotony, growth, decrease) in two time series[1].

2 Basic terms

The presented text is motivated by the concept of positive and quadrant dependence (e.g. [1] and others). The bond strength between two time series can be measured in many ways. One of them can be preserving (or not preserving) the probability of monotone relationship. This is the

[1] Such theory can be done for more time series. There will be work only with pairs of time series.

probability that the values of one time series are increasing and the values of the second time series are also increasing and similarly for consistent decreases (analytically of course for discordance behaviour).
Next, here will be also work with time series whose values are continuous random variables (such a probability model, in many cases, does not distort reality). It means the following cases, which may occur:

Series A	increases	increases	decreases	decreases
Series B	increases	decreases	increases	decreases
Status code	1,1	1,0	0,1	0,0
Numeric code	3	2	1	0

Tab. 1. Possible time series states.

Of course, increase or decrease must be related to some, non-empty, time interval or to compare the value at one time point with the value at another time point. Here, time series are limited to time series with a discrete time axis and for comparison at two time points. Therefore, for some given τ the considered states are written in the following table.

Series A	$x_A(t) > x_A(t-\tau)$	$x_A(t) > x_A(t-\tau)$	$x_A(t) < x_A(t-\tau)$	$x_A(t) < x_A(t-\tau)$
Series B	$x_B(t) > x_B(t-\tau)$	$x_B(t) < x_B(t-\tau)$	$x_B(t) > x_B(t-\tau)$	$x_B(t) < x_B(t-\tau)$
Status code	1,1	1,0	0,1	0,0
Numeric code	3	2	1	0

Tab. 2. Possible comparison between two time series.

Thus, each realization of a pair of time series (for given τ and at each time t) is assigned its state. For the degree of consistency, the concordance rate, of these time series will then be considered the probability that there will be state $(1,1)$ or $(0,0)$ anytime. From this, it is obvious that there will be work with pairs of time series, for which this probability does not depend (functionally) on time t (but it may and it will be depend on time shift τ).

It is formally possible to write the relationship under the following assumptions:

$$\begin{aligned} concr(x_A, x_B; \tau) &= P([1,1]|_\tau \cup [0,0]|_\tau) = P([1,1]|_\tau) + P([0,0]|_\tau) \\ &\equiv P([3]|_\tau) + P([0]|_\tau). \end{aligned} \tag{2.1}$$

Obviously:

$$\begin{aligned} &concr(x_A, x_B; \tau) = concr(x_B, x_A; \tau), \\ &0 \leq concr(x_A, x_B; \tau) \leq 1, \\ &x_B(t) = h\big(x_A(t)\big) \Rightarrow concr(x_A, x_B; \tau) = 0,\ \ h(x)\ \textit{is strictly decreasing}, \\ &concr(x, x_*; \tau) = 1,\ \textit{where}\ x_*\ \textit{is a copy of}\ x. \end{aligned} \tag{2.2}$$

In analogy with Kendal tau [2] a coefficient of concordance can be introduced as a difference of probability matching and probability discrepancy.

$$\begin{aligned} concc(x_A, x_B; \tau) &= P([1,1]|_\tau \cup [0,0]|_\tau) - P([1,0]|_\tau \cup [0,1]|_\tau) \\ &= 2concr(x_A, x_B; \tau) - 1. \end{aligned} \tag{2.3}$$

Obviously:

$$
\begin{aligned}
&-1 \leq concc(x_A, x_B; \tau) \leq +1,\\
&concc(x_A, x_B; \tau) = 0, \textit{ if the decreases or increases in both series are independent,}\\
&concc(x_A, x_B; \tau) = -1, \textit{ if increases in one series implies decreases in second series}\\
&concc(x_A, x_B; \tau) = +1, \textit{ if increases in one series implies increases in second series.}
\end{aligned}
\tag{2.4}
$$

From theory's point of view, this introduction does not bring anything new. However, it is clear. If the coefficient is negative, the probability of discrepancy behaviour dominates the probability of the same. If the coefficient is positive, the probability of the same behaviour dominates over the probability of discrepancy. If it is not possible to make decision significantly from the behaviour of one series to the behaviour of the second series, it is zero or near zero.

3 Statistical inference

Under these assumptions, the state event are realized by a pair of time series in any state, at a given time (and at a given time shift), and equally distributed. However, the assumption of independence of observations will not satisfied for most real (pair) time series (this is needed if we want to use the approaches of random sample, i.i.d.). The pair of time series will contain a greater or lesser amount of memory (= probability dependency over time).
Therefore, we need to use some model of time dependence. This may be a Markov chain model (specifically for ordinary pairs of time series without other functional relationships, homogeneous and of course regular [2]). This can replace the assumption of the independence of observing states by assuming the independence of observed transitions from one state to another.

For this Markov chain model hold

$$p^T(t+1) = p^T(t)P, \tag{3.1}$$

where $p(t)$ is the probability vector of the occurrence of Markov chain in a given state, and P is the probability matrix of transition from state to state, obviously $dim\, p(t) = k$ and $dim\, P = k x k$, in our case $k = 4$.

The elements of the P - matrix can be (pointed) estimated by classical frequencies. So $P(i,j) = Probability\ (transition\ to\ state\ j\ from\ state\ i)$, i.e.:

$$\hat{P}(i,j) = \frac{n(i,j)}{n(i)}, \tag{3.2}$$

where $n(i,j)$ is the number of observed transitions from state i to state j and $n(i) = \sum_{l=1}^{k} n(i,l)$.

Against common tasks, here $\hat{P}(i,j)$ is not a proportion of random variable of the number of observations and a fixed range of random sample. This is the proportion of two random variables. This situation complicates the situation around this estimate. Nevertheless, such an estimate is a consistent[2] estimate of $P(i,j)$, which is sufficient for our purposes.

Another problem, in estimating, may be some specific properties of the matrix P. Here, from the declared use, is $P(i,j) > 0$. It means, that the probability of being in any state "in one step" can get into any other step or remain is nonzero. This would not be guaranteed by the above mentioned frequency approach (especially with a small number of observations = zero frequency problem). Because of this situation, it is used here a Bayesian approach assuming a uniform a priori distribution (e.g. [3]), where in this case is $k = 4$:

$$\hat{P}(i,j) = \frac{n(i,j)+1}{n(i)+k} = \frac{n(i,j)+1}{n(i)+4}; i,j = 0,1,2,3. \tag{3.3}$$

This is a consistent estimate again.

It is true (for regular strings) for any vector q of distribution of probability occurrences in the state of chain $\lim_{n\to\infty} q^T P^n = p^T$, where p is the single probability distribution vector in the state of chain with $p^T = p^T P$. It is only one stationary distribution vector $p_{stac} = p$.

Calculation in this way could be more complicated, and this is an important fact that stationary probabilities are the only ones for the regular chain. Therefore, another way of determining stationary probabilities (least squares method) is here possible:

$$p^T = p^T P \Rightarrow p^T(P - \boldsymbol{I}) = \boldsymbol{0}^{\boldsymbol{T}}, \tag{3.4}$$

where $\boldsymbol{I}$ is an identical matrix of the same dimensions as P and $\mathbf{0}$ is a zero vector of the same size as p.

[2] Proof can be done with using the characteristics of multinomial distribution.

4 Obtained results

4.1 Table results

Exchanges rates of different currencies in one source currency

Data set is daily foreign exchange rates announced by the CNB for the period 3. 2. 2011 - 6. 4. 2017, www.cnb.cz[3] . $\tau = 5$ time shift is 5 quotation days.

CZK per EUR and CZK per USD

Estimation of transition matrix, from state (row), to state (column)

From state (row)		0,0	0,1	1,0	1,1
		0	1	2	3
0,0	0	0,727	0,093	0,105	0,075
0,1	1	0,191	0,550	0,028	0,231
1,0	2	0,251	0,045	0,503	0,201
1,1	3	0,068	0,130	0,094	0,708

It is obvious that dominate the probability of staying in the state.

Stationary probablities		
p(0,0)=	*p(0)=*	***0,343***
p(0,1)=	*p(1)=*	***0,181***
p(1,0)=	*p(2)=*	***0,145***
p(1,1)=	*p(3)=*	***0,331***
Suma =		1,000
Concordance rate		**67,38%**
Concordance coeficient		**0,348**

With stationary probabilities, we can observe the dominance of probabilities of occurrence in concordant states. The probability of consistent behaviour is about 2/3.

Mean recurrence (stay) time		
t(0,0)=	*t(0)=*	3,67
t(0,1)=	*t(1)=*	2,22
t(1,0)=	*t(2)=*	2,01
t(1,1)=	*t(3)=*	3,43

Estimates of the mean times of stay in the state again prefer concordant states.

Matrix of mean transition times, from state (row), to state (column)

From state (row)		0,0	0,1	1,0	1,1
		0	1	2	3
0,0	0	2,92	10,58	11,47	8,42
0,1	1	6,96	5,52	13,26	6,22
1,0	2	6,06	11,18	6,89	6,83
1,1	3	8,48	9,48	12,01	3,02

These values indicate that both time series tend more to conservative behaviour (in terms of mutual relationship).

If we make some qualitative statements about the long-term behaviour of the CZK against EUR and USD, there is no need to distinguish between these two target currencies.

Tab. 3. CZK per EUR and CZK per USD.

[3] The revaluation (or devaluation) here means appreciating the CZK against the currency in which the time series is.

CZK per CNY and CZK per JPY

Estimation of transition matrix, From state (row), to state (column)

From state (row)		0,0	0,1	1,0	1,1
		0	1	2	3
0,0	0	0,762	0,088	0,063	0,087
0,1	1	0,228	0,534	0,016	0,222
1,0	2	0,184	0,029	0,505	0,283
1,1	3	0,082	0,089	0,094	0,735

It is obvious that dominate the probability of staying in the state too.

Stationary probablities		
p(0,0)=	*p(0)=*	*0,361*
p(0,1)=	*p(1)=*	*0,147*
p(1,0)=	*p(2)=*	*0,122*
p(1,1)=	*p(3)=*	*0,370*
Suma =		1,000
Concordance rate		73,18%
Concordance coeficient		0,464

With stationary probabilities, we can observe the dominance of probabilities of occurrence in concordant states. The probability of consistent behaviour is about 73 %.

Mean recurrence (stay) time		
t(0,0)=	*t(0)=*	4,20
t(0,1)=	*t(1)=*	2,15
t(1,0)=	*t(2)=*	2,02
t(1,1)=	*t(3)=*	3,77

Estimates of the mean times of stay in the state again prefer concordant states, especially $t(0,0)$.

Matrix of mean transition times, From state (row), to state (column)

From state (row)		0,0	0,1	1,0	1,1
		0	1	2	3
0,0	0	2,77	12,41	15,31	7,89
0,1	1	6,46	6,82	16,29	6,19
1,0	2	7,25	13,76	8,23	5,31
1,1	3	8,52	12,49	13,96	2,70

Both time series tend to concordant states.

There are time series of foreign exchange currencies from partners from a geographically close environment, so both behave (in terms of changes) significantly similarly.

Tab. 4. CZK per CNY and CZK per JPY.

CZK per EUR and CZK per DKK

Estimation of transition matrix, to state (column) (From state (row))

		0,0	0,1	1,0	1,1
		0	1	2	3
0,0	0	0,783	0,020	0,025	0,172
0,1	1	0,316	0,316	0,013	0,354
1,0	2	0,424	0,015	0,182	0,379
1,1	3	0,192	0,024	0,018	0,766

Significant dominance here have the probabilities of staying in concordant states.

Stationary probablities		
p(0,0)=	*p(0)=*	***0,493***
p(0,1)=	*p(1)=*	***0,031***
p(1,0)=	*p(2)=*	***0,026***
p(1,1)=	*p(3)=*	***0,450***
Suma =		1,000
Concordance rate		**94,38%**
Concordance coeficient		**0,888**

With stationary probabilities, we can observe the dominance of probabilities of occurrence in concordant states. The probability of consistent behaviour is very high. There is possible to say, that it is about 95 %.

Mean recurrence (stay) time		
t(0,0)=	*t(0)=*	4,61
t(0,1)=	*t(1)=*	1,46
t(1,0)=	*t(2)=*	1,22
t(1,1)=	*t(3)=*	4,27

Estimates of the mean times of stay in the state again prefer concordant states. Here is again significant dominance of concordant states.

Matrix of mean transition times, to state (column) (From state (row))

		0,0	0,1	1,0	1,1
		0	1	2	3
0,0	0	2,03	46,54	46,12	5,46
0,1	1	4,11	32,63	47,15	4,06
1,0	2	3,60	46,66	39,06	4,13
1,1	3	4,97	46,03	46,94	2,22

Both time series tend to concordant states. By the biasing from the state, there is a much faster movement to a concordant rather than unconcordant state.

It is clear, that the CZK per DKK exchange rate is set in strong relationship to the CZK per EUR exchange rate (mediated through the market or directly by the method of determination).

Tab. 5. CZK per EUR and CZK per DKK.

CZK per EUR and CZK per PLN

Estimation of transition matrix, to state (column); From state (row)

From state (row)		0,0	0,1	1,0	1,1
		0	1	2	3
0,0	0	0,676	0,128	0,140	0,056
0,1	1	0,165	0,611	0,034	0,190
1,0	2	0,208	0,044	0,580	0,168
1,1	3	0,072	0,137	0,143	0,648

Here, the probabilities of staying in position are greatly dominated.

Stationary probablities		
p(0,0)=	*p(0)=*	***0,305***
p(0,1)=	*p(1)=*	***0,219***
p(1,0)=	*p(2)=*	***0,210***
p(1,1)=	*p(3)=*	***0,266***
Suma =		1,000
Concordance rate		**57,16%**
Concordance coeficient		**0,143**

Probabilities of occurrence in individual states are practically uniformly distributed.

Mean recurrence (stay) time		
t(0,0)=	*t(0)=*	3,09
t(0,1)=	*t(1)=*	2,57
t(1,0)=	*t(2)=*	2,38
t(1,1)=	*t(3)=*	2,84

In terms of mean times of stay in the state, the individual states are similar.

Matrix of mean transition times, to state (column); From state (row)

From state (row)		0,0	0,1	1,0	1,1
		0	1	2	3
0,0	0	3,27	9,17	8,79	9,27
0,1	1	7,15	4,58	10,58	7,18
1,0	2	6,42	10,51	4,76	7,73
1,1	3	8,24	8,98	8,75	3,76

Both time series tend to concordant states. Deviating from the state is a much faster movement to a concordant rather than unconcerned state.

Here are listed (in some ways) the opposite exchange rates. From the value of one exchange rate it is not possible to make “correct” conclusions about the second exchange rate (in terms of non-quantified changes, only increases and decreases).

Tab. 6. CZK per EUR and CZK per PLN.

4.2 Graphical results

The graphical results of the previous prepositions are shown in the following figures. The figures on the left side show clearly the relative frequencies, i.e. point estimates of stationary probabilities of individual states (status codes). Conversely, the figures on the right side the correlation between the considered exchange rates.

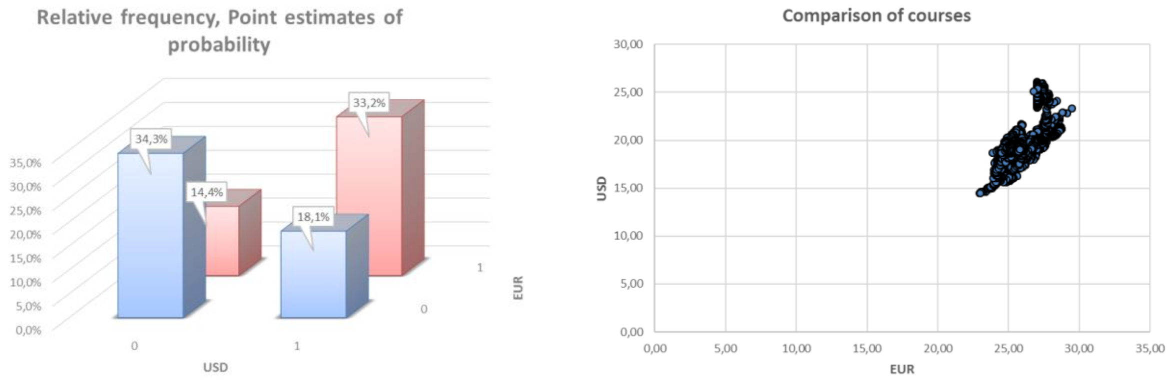

Fig. 1. Relative frequency and correlation among CZK, EUR and USD.

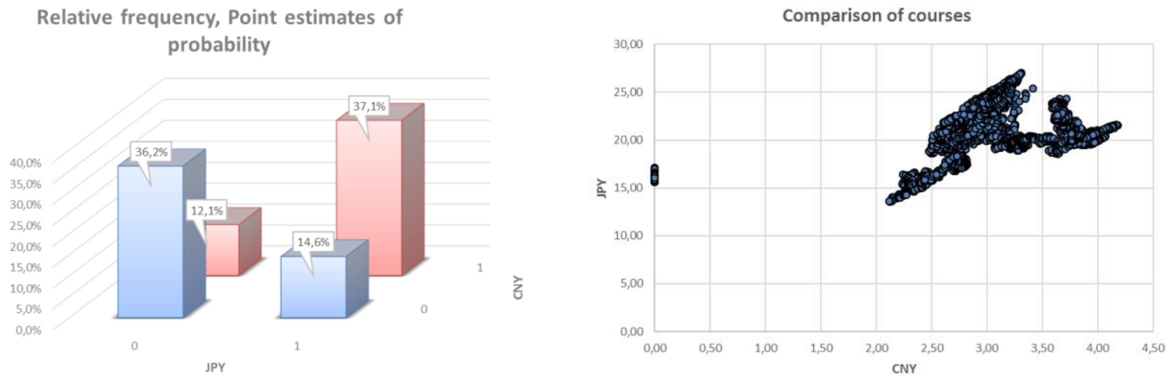

Fig. 2. Relative frequency and correlation among CZK, CNY and JPY.

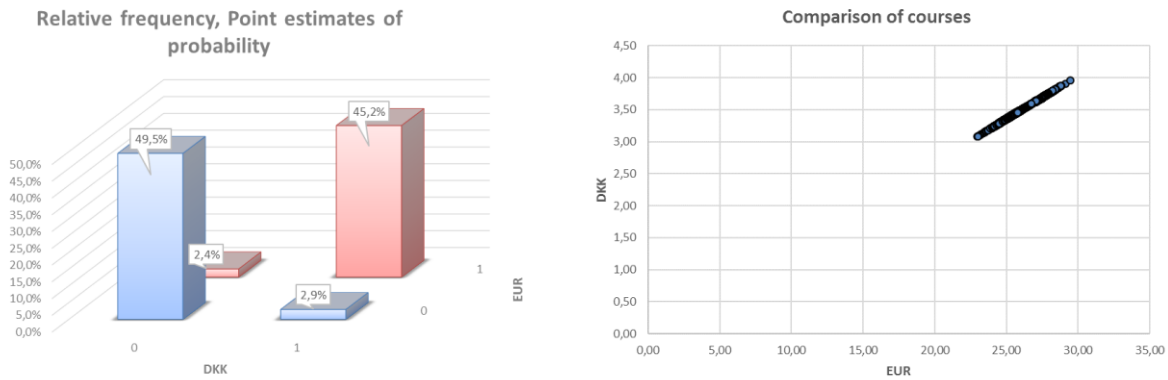

Fig. 3. Relative frequency and correlation among CZK, EUR and DKK.

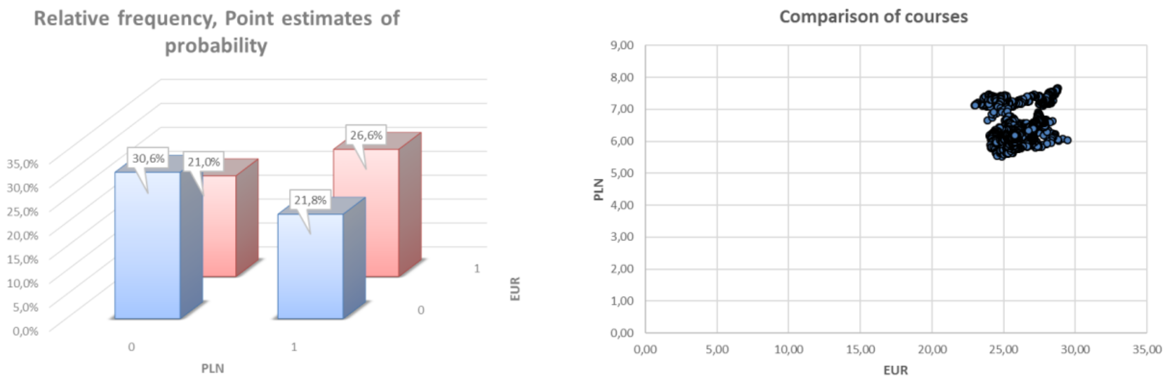

Fig. 4. Relative frequency and correlation among CZK, EUR and PLN.

Another problem in this approach may be the shift time τ. This is how the effect of the results of the shift time is affected. I.e. there is considered the question, how much time shift affect obtained results. The results are different for each observed pair of exchange rates. For better imagination, here is the first case i.e. CZK per EUR and CZK per USD.

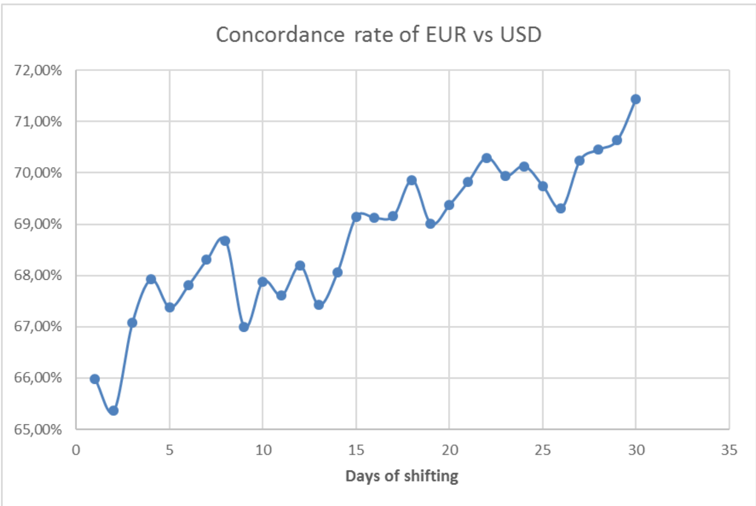

Fig. 5. Effect of time shift on concordance rates.

There is used a 5-day time shift for clarity on the base of observed simulations and for practical use. For the meaningfulness of these approaches, the maximum number of shifting days in simulations was limited to 30.

5 Conclusion

Computational modelling was done for daily published courses:

Source	Number of Exchange series	Source currency	From the period	to
ČNB	28	CZK	2. 1. 2007	6. 4. 2017

Tab. 7. Source data set.

Concordance can be described as "probability matching (qualitative)" practically (in the interval of 1 - 30 quoted days) independent of the "distance" of the date being compared. It appears that the concept of concordance between two time series is an appropriate tool for qualitative classification of the relationship between two time series.

Here are used 4 estimates of concordance rate and other statistics of all possibilities in this paper based on source data set from CNB. The first two are based on significant trading currencies (EUR, USD, CNY, JPY) and the other two are based on interesting results among European states (DKK, PLN). That means a pair of currencies with a high coefficient of concordance and a low coefficient of concordance.

Acknowledgement

This publication was supported by the project LO1506 of the Czech Ministry of Education, Youth and Sports.

References

[1] LEHMANN, E.L., *Some Concepts of Dependence*, 1966. [online] Ann. Math. Statist. Volume 37, Number p..1137-1153. Available at: <https://projecteuclid.org/euclid.aoms/1177699260> [Accessed 8 July 2017].

[2] GRINSTEAD, CH.M. and SNELL, J.L., *Introduction to Probability*, 2003 [online] Published by the American Mathematical Society. Available at: <https://math.dartmouth.edu/~prob/prob/prob.pdf> [Accessed 8 July 2017].

[3] WITTEN I. H. and BELL, T. C., *The Zero-Frequency Problem: Estimating the Probabilities of Novel Events in Adaptive Text Compression*, 1991. [online] IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 37, NO. 4. Available at: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=87000> [Accessed 8 July 2017].

Current address

Ťoupal Tomáš, Ing., Ph.D.
European Centre of Excellence NTIS – New Technologies for Information Society
Faculty of Applied Sciences
University of West Bohemia
Univerzitní 8, 301 00 Plzeň, Czech Republic
Tel. number: +420 377 63 2674, e-mail: ttoupal@kma.zcu.cz

Šedivá Blanka, RNDr., Ph.D.
European Centre of Excellence NTIS – New Technologies for Information Society
Faculty of Applied Sciences
University of West Bohemia
Univerzitní 8, 301 00 Plzeň, Czech Republic
Tel. number: +420 377 63 2618, e-mail: sediva@kma.zcu.cz