MATHEMATICAL ECONOMICS

No. 13(20) 2017

ESTIMATION OF THE PROBABILITY OF INVERSION IN ASSESSING THE CREDIBILITY OF EXPERTS. A CASE STUDY[1]

Agnieszka Bukietyńska, Mariusz Czekała, Marek Gurak, Jacek Jagodziński, Jarosław Kłosowski

Abstract. The aim of the paper is to obtain an objective opinion from the experts. Experts are evaluated on the basis of the accordance of their preferences with the total preferences. The inversion distribution was applied. This distribution is associated with inversions occurring in the tau-Kendall test. The probabilities obtained as a result of estimation are needed to construct the estimator of demand.

Keywords: inversions, fashion industry, probability of mistake, estimation.

JEL Classification: C12, C46, D81, G17, N21.

DOI: 10.15611/me.2017.13.01.

1. Introduction

The paper presents the concept of estimating the probability to solve the problem of forecasting demand for a selected group of products from the fashion industry. The presented example of estimation is part of a problem being experienced by a company dealing in the distribution of linen industry goods.

Because fashion-related phenomena are often subject to sudden and unexpected changes, the use of subjective expert knowledge is a way to overcome certain specific difficulties.

Agnieszka Bukietyńska
Wroclaw School of Banking
e-mail: agnieszka.bukietynska
ORCID: 0000-0002-8318-7530

Mariusz Czekała
Wroclaw School of Banking
email: mariuszczekala@windowslive.com
ORCID: 0000-0003-1007-0806

Marek Gurak
MG4-LTD Co.
e-mail: mgurak@mg4.eu
ORCID: 0000-0002-0669-3384

Jacek Jagodziński
Wrocław University of Science and Technology
e-mail: jacek.jagodzinski@pwr.edu.pl
ORCID: 0000-0002-4693-7531

Jarosław Kłosowski
Wroclaw School of Banking
jarek.klosowski@gmail.com

[1] This is a part of the Program Operacyjny Inteligentny Rozwój 2014-2020 of Narodowe Centrum Badań i Rozwoju, number of application: POIR.01.01.01-00-0886/17.

To use expert knowledge, a reliable method of their assessment is needed. In this work the concept related to the inversion distribution will be used. In addition to the distribution associated with the inversion analysis, the method of estimating the probability of an event interpreted as an expert error will be used. The proposed estimator is unbiased.

As mentioned earlier, the goal of the entire project is to obtain an effective method of forecasting. The forecasting method is based on expert opinions. Their predictions are to be taken into account after taking into account the probabilities of making mistakes, therefore this is a very important step in building a predictor.

The first part of the work is theoretical. A distribution which is a natural generalization of the distribution occurring in the tau-Kendall test will be shown. Then, the estimation method will be considered. The next part will present data obtained from the MG4 company. In the final part the results of the estimation of inversion probabilities will be shown (the contractual name).

2. Inversions

Definition. Let $(X_1, \dots, X_n)$ be the sequence of observations. The inversion for i and j occurs, if and only if $i > j$, but $X_i < X_j$.

The observation sequence $(X_1, \dots, X_n)$ can be equated with the permutation of the first n natural numbers. In this work, it will be relevant to determine the preferences of experts. These preferences may be described using a sequence of natural numbers, the elements of which describe the order preferred by the expert.

Having a sample pattern (actually realized demand) for the learning sample, the preferences of the analyzed expert can be compared with this pattern. Then the inversion number will be interpreted as the number of errors committed by the expert.

The number of inversions for selected values of n is presented in Figure 1. In the considered problem, the case of $n = 9$, to a lesser extent cases $n = 4$ and $n = 3$, is the one of particular importance. A recursive method of calculating the inversion number is presented in the works of [Czekała, Bukietyńska 2017].

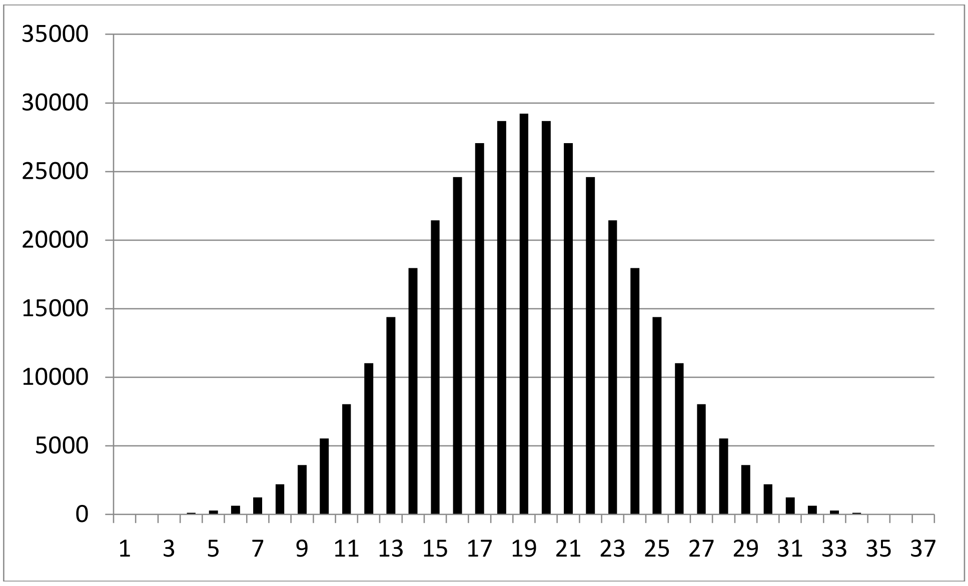

Fig. 1. Number of inversions, case of $n = 9$

Source: own calculations.

The number of permutations with a given inversion number varies from 1 (there is only one permutation with a zero inversion number – permutation being the identity) to 29,228 since there are so many permutations with 18 inversions. To obtain the probability, the number of permutations with a given number of inversion should be divided by 362,880 (in the general case by n!).

It should be remembered, however, that the distribution for parameter $p = 0.5$ is obtained.

Additional number of inversions is presented in Table 1.

Table 1. Distribution of Y_k for $p = 0.5$

Y_k	0	1	2	3	…	…	k-1
Probability	1/k	1/k	1/k	1/k	…	…	1/k

Source: [Feller 1961].

In Table 2 an associated case is considered, which is an important extension of the previous case.

Table 2. Distribution of Y_k (general case)

Y_k	0	1	2	3	…	k-1
Probability	q^{k-1}/w_k	$q^{k-2}p/w_k$	$q^{k-3}p^2/w_k$	$q^{k-4}p^3/w_k$	…	p^{k-1}/w_k

where $w_k = (q^k - p^k)/(q-p)$ for $q \neq p$

Source: [Czekała, Bukietyńska 2017].

The full distribution of the number of inversions in a general case (with p generally different from 0.5) is shown in Table 3.

Table 3. Distribution of I_n

I_n	0	1	…	k	…	N_n
Probability	$\left\{ {N_n \atop 0} \right\} \frac{q^{N_n}}{W_n}$	$\left\{ {N_n \atop 1} \right\} \frac{q^{N_n-1}p}{W_n}$	…	$\left\{ {N_n \atop k} \right\} \frac{q^{N_n-k}p^k}{W_n}$	…	$\left\{ {N_n \atop N_n} \right\} \frac{p^{N_n}}{W_n}$

Source: [Czekała, Bukietyńska, Schlichtinger 2018].

In Table 3, $N_n = \frac{n(n-1)}{2}$ is the maximal number of inversions, $\left\{ {N_n \atop k} \right\}$ is the number of permutation having exactly k inversions, p is the probability of inversions and $\mathrm{q} = 1 - p$. Moreover

$$W_n = \sum_{k=0}^{N_n} \left\{ {N_n \atop k} \right\} q^{N_n-k} p^k$$

are normalizing constants.

Numbers $\left\{ {N_n \atop k} \right\}$ may be computed using the method presented in the paper [Czekała, Bukietyńska 2017], where are the terms of the well-known sequence A008302.

3. Data presentation

Usually, in practical matters (this is also the case when evaluating experts) the p parameter is unknown. It should therefore be estimated. The estimation method will be presented and applied below.

In the first column of Table 4 the symbols from $e1$ to $e13$ denote the experts. In the discussed example there are 13 experts. In the first line of Table 4 there appear the groups of the products. Table 5 contains an explanation of the acronyms.

Table 4. The order of the experts

Product group	**MP**	**MX**	**LP**	**LH**	**TB**	**MN**	**TP**	**LG**	**LN**	**Sum**
Sum	92452	76717	67159	26914	24133	15933	7004	5642	5210	323317
***e*1**	***20160***	***19380***	***30690***	***2115***	***26430***	***1320***	***8760***	***60***	***2640***	111555
***e*2**	***0***	***0***	***0***	***2040***	***0***	***0***	***0***	***310***	***1980***	4330
***e*3**	***0***	***0***	***1080***	***0***	***1270***	***0***	***0***	***0***	***0***	2350
***e*4**	***16300***	***5500***	***5410***	***1340***	***3700***	***890***	***655***	***40***	***1145***	34980
***e*5**	***2490***	***850***	***990***	***365***	***615***	***640***	***0***	***20***	***275***	6245
***e*6**	***3740***	***2895***	***3670***	***445***	***1690***	***325***	***120***	***25***	***265***	13175
***e*7**	***8710***	***4705***	***12230***	***1110***	***3910***	***860***	***800***	***70***	***865***	33260
***e*8**	***14160***	***11140***	***8840***	***1825***	***16680***	***1015***	***1260***	***250***	***1340***	56510
***e*9**	***6630***	***2055***	***4975***	***1135***	***1715***	***720***	***250***	***60***	***640***	18180
***e*10**	***3580***	***890***	***3605***	***755***	***785***	***520***	***180***	***120***	***415***	10850
***e*11**	***5420***	***3640***	***4920***	***380***	***1595***	***1025***	***40***	***40***	***690***	17750
***e*12**	***4680***	***2050***	***3660***	***275***	***405***	***275***	***20***	***20***	***240***	11625
***e*13**	***10540***	***3960***	***18040***	***855***	***2680***	***250***	***30***	***45***	***735***	37135

Source: data obtained from the MG4 company.

Table 5. Groups of products and ranks for the total

Product group	Rank	Name
MP	1	Men’s Basic Pants
MX	2	Men’s Boxer Shorts
LP	3	Ladies’ Basic Pants
LH	4	Ladies’ Homewear Set
TB	5	Teens’ Basic Colour Pants &Bra
MN	6	Men’s Homewear Night&Day
TP	7	Teens’ Bikini
LG	8	Ladies’ Homewear Dressing Gown
LN	9	Ladies’ Homewear Night&Day

Source: data obtained from the MG4 company.

The basis of an experts' evaluation constitutes a comparison of their preferences with the total ones. This means endowing each of the experts with a rank. In Table 6, data about the expert $e2$ and $e3$ are omitted due to the fact that they made too small number of choices. This was also the case with ranking of the experts $e11$ and $e12$ because of ties. According to the theorem about the distribution of inversions, ties are not allowed. These two cases will be discussed separately.

Table 6. Ranks of the experts

Product group	MP	MX	LP	LH	TB	MN	TP	LG	LN	INVERSIONS
Rank of sum	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	**0**
$e1$	3	4	1	7	2	8	5	9	6	**10**
$e4$	1	2	3	5	4	7	8	9	6	**4**
e5	1	3	2	6	5	4	9	8	7	**7**
e6	1	3	2	5	4	6	8	9	7	**4**
$e7$	2	3	1	5	4	7	8	9	6	**6**
$e8$	2	3	4	5	1	8	7	9	6	**8**
$e9$	1	3	2	5	4	6	8	9	7	**4**
$e10$	2	3	1	5	4	6	8	9	7	**5**
$e13$	2	3	1	5	4	7	9	8	6	**7**

Source: own calculations.

4. Estimation

It may be proved that [Czekała, Bukietyńska, Schlichtinger 2018]:

$$EY_n = \begin{cases} \frac{p}{q-p} \cdot \frac{q^n - nqp^{n-1} + (n-1)p^n}{q^n - p^n} & \text{for} \quad p \neq q \\ \frac{n-1}{2} & \text{for} \quad p = q \end{cases},$$

where $q = 1 - p$.

It may be proved using elementary methods that function $EY_n(p)$ is strictly increasing. The expected values for the specified values of parameter p (for $n = 9$) are included in Table 7.

Table 7. Expected values ($n = 9$)

P	$E(Y_n)$
0.1	0.125
0.2	0.333299
0.3	0.745608
0.4	1.759637
0.5	4
0.6	6.240363
0.7	7.254392
0.8	7.666701
0.9	7.875

Source: own calculations.

Because $I_n = \sum_{k=1}^{n} Y_k$, hence

$$EI_n = \sum_{k=1}^{n} EY_k = F_n(p).$$

With regard to the proposition included for example in the paper [Czekała, Bukietyńska, Gurak 2017], the observed number of inversions should fulfil the equation

$$F_n(p) = \hat{I}_n.$$

Since function F_n is the sum of the strictly increasing functions, so there exists the inverse function F_n^{-1}. That is why, the unbiased estimator of parameter p is the following one:

$$\hat{p} = F_n^{-1}(I_n).$$

The predefined estimator is an unbiased estimator.
To obtain values of the estimator of observed numbers of inversions, one ought to solve equations:

$$\hat{p} = F_n^{-1}(I_n),$$

for $I_n \in \{4,5,6,7,8,10\}$. Values from this set have been observed according to data from Table 6 (the last column). To solve the abovementioned equation, the module *Solver* was applied. Additionally, the value of the estimator was computed for $I_n \in \{1,2,3,9\}$. The results of the computations are presented in the Table 8.

For $I_n = 18$ the estimated value of parameter p is equal to 0.5. In Figure 2 there are distributions of the number of inversions for selected numbers of inversions and the probabilities that correspond with them. Figure 2 illustrates the significant difference between the random case corresponding with the number of inversions equal to 18 and the probability endowed with a value 0.5.

Table 8. Results of estimation

Expert	I_n	$\hat{p}$
$e1$	**10**	0.402801
$e4$	**4**	0.267075
$e5$	**7**	0.349509
$e6$	**4**	0.267075
$e7$	**6**	0.326629
$e8$	**8**	0.369373
$e9$	**4**	0.267075
$e10$	**5**	0.299673
$e13$	**7**	0.349509
-	**9**	0.386959
-	**3**	0.226422
-	**2**	0.173811
-	**1**	0.102521

Source: own calculations.

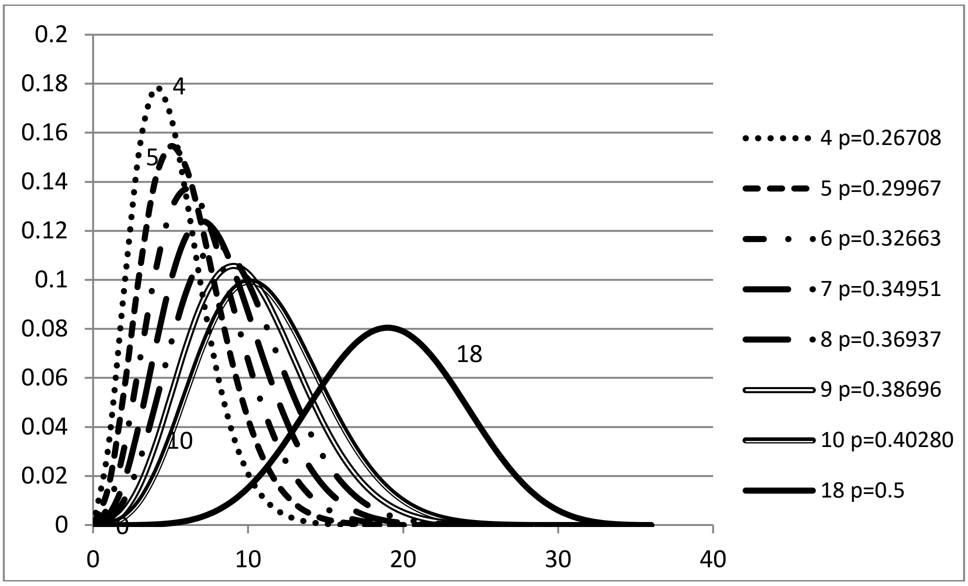

Fig. 2. Distribution of inversions number for selected values of p

Source: own work.

Similarly as shown in the papers [Bukietyńska 2017; Bukietyńska et al. 2017; Czekała, Bukietyńska, Gurak 2017], the distribution of the inversions number may be used to test the hypothesis $H\colon p = 0.5$. In this paper the validity of such a hypothesis means the total randomness of the experts' opinion. In this case, p-values are necessary to make a decision. These values are presented in Figures 3 and 4.

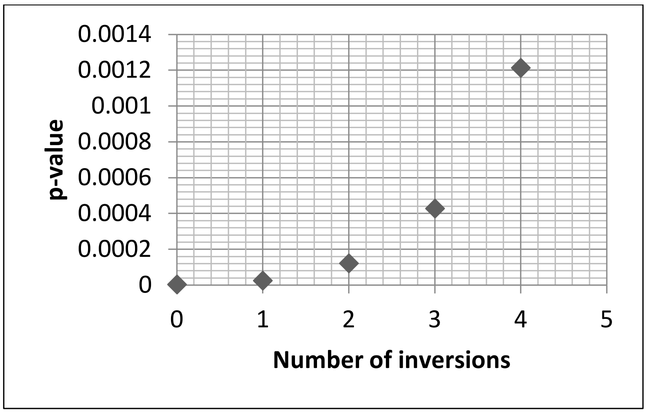

Fig. 3. p-value, case $n \leq 4$

Source: own work.

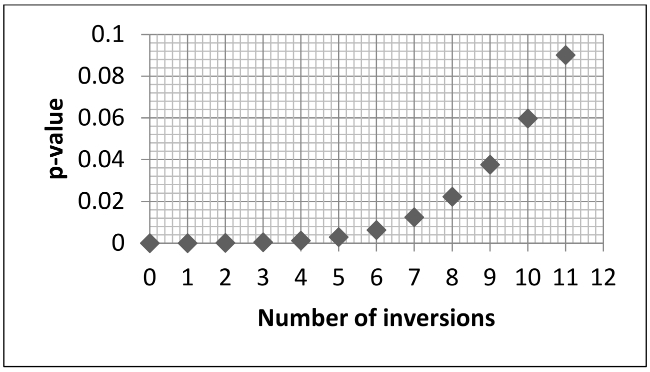

Fig. 4. p-value, case $n \leq 11$

Source: own work.

Knowing the values from Figures 3 and 4 make testing the hypothesis about usefulness of an expert possible.

5. Summary

The presented example gives the method assigning the probability of a mistake to each expert. This probability may be identified with probability p occurring in the distribution of the number of inversions. These probabilities shall be used to construct the predictor, where weights will be characterized by inverse dependence in relation to the probability of mistakes. In this paper the case of ties was omitted. The cause being the fact that the distribution of inversions does not allow such a possibility. In the case of the experts 11th and 12th, ties are occurring twice. The heuristic solution of such a problem may be based on the consideration, in each of these cases, of two permutations using a replacement of ranks. Therefore one may obtain two possibly observed values of inversion, which are included in the range presented in Table 8. The situation seems to be more complex when one has too small a number of observations. The presented method cannot be used in a situation that is similar to the one related to expert $e3$. The case when one has to do with three choices is on the borderline of applying the method, however this was considered in the paper [Czekała, Bukietyńska, Gurak 2018].

Bibliography

Bukietyńska A. (2017). *The test of inversion in the analysis of investment funds*. Central and Eastern European Journal of Management and Economics.

Bukietyńska A., Czekała M., Wilimowska Z., Wilimowski M. (2017). *The Inversion Test of the Investment Funds Efficiency Measures*. Advances in Intelligent Systems and Computing.

Czekała M., Bukietyńska A. (2017). *Distribution of Inversions and the Power of the τ-Kendall's Test*. Advances in Intelligent Systems and Computing.

Czekała M., Bukietyńska A., Gurak M. (2018). *Rozkład inwersji w analizie poprawności ocen eksperckich na przykładzie branży modowej* (accepted for publication).

Czekała M., Bukietyńska A., Schlichtinger A.M. (2018). *Estimation of the probability of inversion in the test of variable dependencies*, Advances in Intelligent Systems and Computing.

Feller W. (1961). *An Introduction to Probability Theory and its Application*. Wiley. New York.

Janjic M. (2014). *A generating function for numbers of insets*. J. Integer Seq. 17, #14.9.7.

Kendall M.G., Buckland W.R. (1960). *A Dictionary of Statistical Terms*. Oliver and Boyd. Edinburgh.

Netto E. (1927). *Lehrbuch der Kombinatorik*. 2nd ed. Leipzig. Teubner, pp. 96.

The On-Line Encyclopaedia of Integer Sequences, sequence A008302. Data related to sales obtained from MG4 Company.