Computer Science. Tom 1. 1999 517.977.5:519.816

*Wiesław Wajs**

OPTIMISTIC AND PESSIMISTIC RESULT OF PLANNING AND SCHEDULING DYNAMIC PROCESSES

1. Introduction

In many real-life circumstances decision problems arise. Optimisation problems can be formulated as decision problems as well. An optimisation problem can be expressed in terms of a model and a performance index. While the model describes the problem, the performance index assigns a value to each feasible realisation of the problem [1].

An algorithm is a method to solve a class of problems with computer. The computational complexity of an algorithm, which can be measured, is the cost. It is measured in runtime during which the algorithm is used to solve one of the problems. If the runtime is limited by a polynomial function of the amount of input data at most, the problem is said to be an easy one otherwise it is a hard problem. If a problem is easy it is enough to describe a method meeting such a constraint, when used to solve the problem. What does it mean that a problem is hard? The problem is hard when it is necessary to prove that it is impossible to find a fast method to perform the calculations which identify an optimal solution. There are a number of easy problems. Matrix inversion is easy: $n \times n$ matrix can be inverted with the Gaussian elimination method in time of cn^3 at most. Sorting problem is easy as well. The fact that a computational problem is hard does not imply that its every instance has to be hard. The problem is hard when no algorithm can be pointed at, which could ensure a high performance for all instances of the problem. Notice that the amount of input data to the computer in this example is small [7].

In recent years there has been a growth in research which deals with the development and complexity analysis of combinatorial algorithms. Complexity measures are of two kinds: *static*, independent of the size and characteristics of the input data, and *dynamic*, dependent on the input data [3].

* Katedra Automatyki, Akademia Górniczo-Hutnicza, Kraków

2. Optimisation for dynamic problems

Optimisation is aimed at finding the optimum sequence for the given form of the performance index. This section deals with the problems of optimisation for dynamic problems, including combinatorics optimisation. Three forms of the performance index for optimisation tasks are established. Now we are considering the scheduling problem.

The loss of the profit $p_i(C_{i-1})$ depends on the resources volume C_{i-1}. Special attention is given to finding polynomial algorithms used for combinatorics optimisation tasks. These optimisation tasks may be solved for a large number of dynamic processes.

The optimisation algorithms are built upon the sorting procedure. It is assumed that the performance index is additive. On the completion of i dynamic processes, the performance index value determines the initial condition for the subsequent dynamic problem under realisation which is discretionally chosen out of $n-i$ processes yet to be carried out.

In real physical or economic processes it is necessary to know a definite time interval for carrying out the elementary technical, technological or economic operation.

3. Applied Statistics

The plotting of experimental results to see if there is any orderly relation between variables is usually referred to as correlating the data [26]. The pairs of values of the variables associated with each data point are designated x_i and y_i with y assigned to the variable which is imprecisely known, or to the dependent variable. A straight line through the data is expressed as: $\hat{y} = ax + b = ax + (\overline{y} - a\overline{x})$, where $\hat{y}$ is estimated value of y for observed value of x. a is the slope of line, identical with regression coefficient, and b is intercept which gives the estimated value of y at $x = 0$. The values of a and b corresponding to the line are calculated with the minimum – squared deviation of y from $\hat{y}$. It is not obvious from a plot of the data whether or not it is reasonable to draw a straight line through the points. It is possible to test the estimated variance removed by the linear correlation against the estimated variance remaining after correlation. The total sum of squares of deviation variable from its mean is $\Sigma(y - \overline{y})^2$, and the total degrees of freedom are $n - 1$. r^2 is the fraction of the sum of squares of deviation removed by the correlation line, $(1- r^2)$ is the sum of squares of deviation from the least – squares line, equal to $\Sigma(y - \hat{y})^2$, with $n - 2$ degrees of freedom. We designate the estimate of variance removed by the correlation as $s^2(C)$. The ratio of $s^2(C)/s^2(\hat{y})$ may be tested by the F ratio test (Fisher test *R. A. Fisher, Frank Yates, Statistical Tables for Biological, Agricultural and Medical Research, Oliver and Boyd Ltd., Edinburg and London 1953*) [25], [24] for 1 and $n - 2$ degrees of freedom to see whether the variance removed by the correlation line is significant when compared to the residual variance of estimate. This test is equivalent to testing the significance of r, the correlation coefficient, since $s^2(C)/s^2(\hat{y}) = r^2(n - 1)/(1 - r^2)$. Since F at 1 degree of freedom is equal to t^2 (t test), the F ratio test for correlation $r^2(n - 2) / (1 - r^2)$, can be expressed in terms of t, that is $t = r^2\sqrt{n-2}/\sqrt{1-r^2}$. This t test, the F test, and the tabulated significant values for the correlation coefficient will all give identical results.

Example

An example has been written by Matlab.

(The optimistic and pessimistic result of a slope *a*)

There is given a vector *x* and a vector *y*.

```
x=[10 40 50 210 220 470 850];
y=[10 20 30 45 50 80 120];
n=7;
sr_x=mean(x)
sr_y=mean(y)
sy=sum(y)
Sx=sum(x)

sy2=sum(y*y')
sx2=sum(x*x')
sxy=sum(x*y')
Spx2=sum(x*x')-1/n*(sum(x))^2

spy2=sum(y*y')-1/n*(sum(y))^2
spxy=sum(x*y')-1/n*(sum(x)*sum(y))
a=spxy/spx2
B=sr_y-a*sr_x

r=spxy/sqrt(spx2*spy2)
r2=r*r
s2dash_y=(1-r2)*spy2/(n-2)
S2a=s2dash_y/spx2

sa=sqrt(s2a)
s2bar_y=s2dash_y/n
Sbar_y=sqrt(s2bar_y)

t005n_2=2.571
```

The value t005n_2=2.571 comes from *R. A. Fisher, Frank Yates, Statistical Tables for Biological, Agricultural and Medical Research, Oliver and Boyd Ltd., Edinburg and London 1953)*. It is given at the 0.05 level, and $n-2=5$.

```
bary_u=sr_y+t005n_2*sbar_y
bary_l=sr_y-t005n_2*sbar_y
a_u=a+t005n_2*sa
```

a_u is a pessimistic result for a slope a.

```
a_l=a-t005n_2*sa
```

a_l is an optimistic result for a slope a.

for $i=1:n$,

```
Y_UR(i) a*x(i)+b+t005n_2* ...
```

```
sqrt(s2dash_y*(1/n+(x(i)-sr_x)*(x(i)-sr_x)/spx2));
end
```

for $i=1:n$,

```
y_LR(i)=a*x(i)+b-t005n_2* ...
sqrt(s2dash_y*(1/n+(x(i)-sr_x)*(x(i)-sr_x)/spx2));
end
y_UR
y_LR
plot(x,y,'*')
hold on;
plot(x,a*x+b)
hold on;
plot(x,(a+t005n_2*sa)*x+sr_y-sr_x*(a+t005n_2*sa),'r')
hold on;
plot(x,(a-t005n_2*sa)*x+sr_y-sr_x*(a-t005n_2*sa),'b')
plot(sr_x,bary_l,'+');
hold on;
plot(sr_x,bary_u,'o');
hold on;
plot(x,y_UR);
hold on;
plot(x,y_LR);
title('Confidence intervals');
xlabel(' x axis');
ylabel(' y axis').
```

The results of the example are shown in Fig. 1.

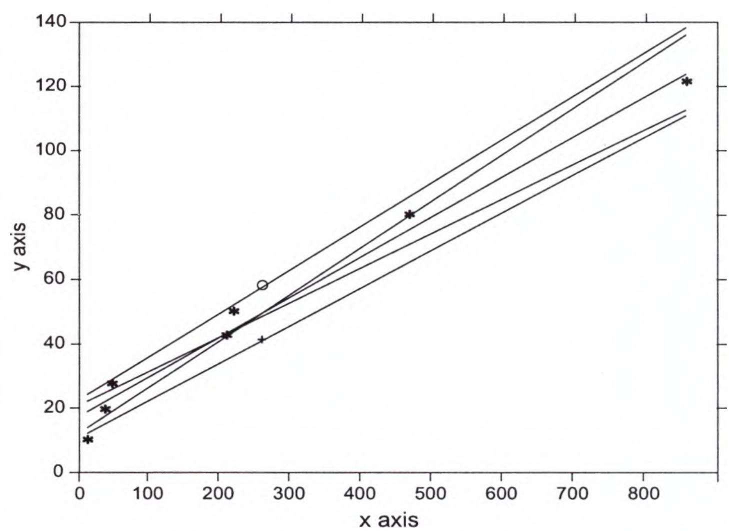

Fig. 1. Applied statistic results

4. Scheduling of linear dynamic processes

The model of the optimisation process is given as

$$n \,|\, p_i(C_{i-1}) = a_i C_{i-1} + b_i;\ a_i > 0;\ b_i > 0;\ C_0 \geq 0 \,|\, min \sum C_i .$$

In this section we examine a case of scheduling problem where the coefficients *a* and *b* are different for each model [21]. The relations $a_1 \leq a_2 \leq a_3 \leq ... \leq a_{n-1} \leq a_n$ are assumed to be fulfilled.

In Table 1 $p_i(C_{i-1}), C_i, \sum C_i$ are shown (for permutation of models in the form (1, 2, 3,..., $n-1$, n).

Table 1

i	$p_i(C_{i-1})$	C_i	ΣC_i
0	0	C_0	C_0
1	$a_1C_0+b_1$	$(a_1+1)C_0+b_1$	C_0+C_1
2	$a_2(a_1C_0+b_1)+b_2=a_2C_1+b_2$	$(a_2+1)C_1+b_2$	$C_0+C_1+C_2$
3	$a_3C_3+b_3$	$(a_3+1)C_2+b_3$	$C_0+C_1+C_2+C_3$
...	....	....	
n	$a_nC_{n-1}+b_n$	$(a_n+1)C_{n-1}+b_n$	$C_0+C_1+...+C_n$

PROBLEM

$$n \,|\, p_i(C_{i-1}) = a_i C_{i-1} + b_i;\ a_i > 0;\ b_i > 0;\ C_0 \geq 0 \,|\, min \sum C_i$$

In this section we examine a case of scheduling problem where the coefficients *a* and *b* are different for each model. The relations above are assumed to be fulfilled.

It is assumed that relations are fulfilled.

Theorem 1

If for the problem

$$n \,|\, p_i(C_{i-1}) = a_i C_{i-1} + b_i;\ a_i > 0;\ b_i > 0;\ C_0 \geq 0 \,|\, min \sum C_i$$

for $\forall_{j=1,2...n}$ and $i \neq j$

the relation

$$C_0(a_i+1) + b_i + ((a_j+1)\, b_i + b_j)\left(1 + \sum_{k=1}^{k=n-2} \prod_{l=1, l\neq i, l\neq j}^{l=k} (a_l+1)\right)$$

$$\Leftarrow C_0\,(a_j+1)+b_j+((a_i+1)\,b_j+b_i)\left(1+\sum_{k=1}^{k=n-2}\;\prod_{l=1,l\neq i,l\neq j}^{l=k}(a_l+1)\right)$$

is fulfilled

then the *i*th model occupies the first position of optimal permutation.

PROOF

For the sequence of models $(1, 2,..., n-1, n)$ the performance index is

$$\sum C_i' = C_0 + C_1 + C_2 +,...,+ C_{n-1} + C_n =$$

$$C_0 +$$

$$(a_2 + 1)C_0 + b_1 +$$

$$(a_2 + 1)\,(a_1 + 1)C_0 + (a_2 + 1)b_1 + b_2 +$$

$$...+$$

$$(a_2+1)(a_1+1)C_0\prod_{l=3}^{l=n-1}(a_l+1)+$$

$$b_1(a_1+1)\prod_{l=3}^{l=n}(a_l+1)+$$

$$\sum_{k=3}^{k=n-1} b_{k-1}\prod_{l=k}^{l=n-1}(a_l+1)+$$

$$b_{n-1}+$$

$$(a_2+1)(a_1+1)C_0\prod_{l=3}^{l=n}(a_l+1)+$$

$$b_1(a_2+1)\prod_{l=3}^{l=n}(a_l+1)+$$

$$\sum_{k=3}^{k=n} b_{k-1}\prod_{l=n}^{l=n}(a_l+1)+$$

$$b_n.$$

For the sequence of models (2, 1,..., $n-1$, n) the performance index is

$$\sum C_i'' = C_0 + C_2 + C_1 +,...,+ C_{n-1} + C_n =$$

$$C_0$$

$$(a_2 + 1)C_0 + b_2 +$$

$$(a_1 + 1)(a_2 + 1)C_0 + (a_1 + 1)\, b2 + b_1 +$$

$$... +$$

$$(a_1+1)(a_2+1)C_0 \prod_{l=3}^{l=n-1} (a_l+1)+$$

$$b_2 (a_1+1) \prod_{l=3}^{l=n-1} (a_l+1)+$$

$$\sum_{k=3}^{k=n-1} b_{k-1} \prod_{l=3}^{l=n-1} (a_l+1)+$$

$$b_{n-1} +$$

$$(a_1+1)(a_2+1)C_0 \prod_{l=3}^{l=n} (a_l+1)$$

$$b_2 (a_1+1) \prod_{l=3}^{l=n} (a_l+1) +$$

$$\sum_{k=3}^{k=2} b_{k-1} \prod_{l=k}^{l=2} (a_l+1) + bn.$$

Given the relation $\sum C_i' \Leftarrow \sum C_i''$,
we have

$$C_0 (a_i+1) + b_i + ((a_j+1)b_i + b_j) \left(1 + \sum_{k=1}^{k=n-2} \prod_{l=1, l \neq i, l \neq j}^{l=k} (a_l+1) \right)$$

$$\Leftarrow C_0 (a_j+1) + b_j + ((a_i+1)b_j + b_i) \left(1 + \sum_{k=1}^{k=n-2} \prod_{l=1, l \neq i, l \neq j}^{l=k} (a_l+1) \right)$$

Now, there are $n - 1$ models. The model on the first position of optimal permutation is fixed. We can calculate a new initial condition

$$C_0 := C_0 + a_i C_0 + b_i.$$

The models are assumed to be scheduled by the relation

$$a_1 \le a_2 \le a_3 \le \ldots \le a_i \le a_j \le \ldots \le a_{n-2} \le a_{n-1}.$$

Theorem 2

If for the problem

$$dn \left| p_i(C_{i-1}) = a_i C_{i-1} + b_i;\ a_i > 0;\ b_i > 0;\ C_0 \ge 0 \right| \ min \sum C_i$$

where $i \in \{1, 2, \ldots, m\}$ and $2 \le m \le n-2$

for$\forall_{j=1,2\ldots n}$ and $i \ne j$ the relation

$$C_0 (a_i + 1) + b_i + ((a_j + 1) b_i + b_j) \left(1 + \sum_{k=1}^{k=m} \prod_{l=1, l\ne 1, l \ne j}^{l=k} (a_l + 1) \right)$$

$$\Leftarrow C_0 (a_j + 1) + b_j + ((a_i + 1) b_j + b_i) \left(1 + \sum_{k=1}^{k=m} \prod_{l=1, l\ne i, l \ne j}^{l=k} (a_l + 1) \right)$$

is fulfilled then the model M_i occupies the ith position of optimal permutation.

PROOF

The proof is analogous to the proof of theorem 1.

Example

The algorithm resulting from the theorems1 and 2 is tested. The data for the problem

$$n \left| p_i(C_{i-1}) = a_i C_{i-1} + b_i;\ a_i > 0;\ b_i > 0; C_0 \ge 0 \right| min \sum C_i$$

are presented in Table 2.

Table 2

i	a_i	b_i	C_0
0			1
1	1.00000	3.13799	
2	2.86105	20.25810	
3	3.27292	67.16544	
4	4.31869	16.17955	
5	5.37224	42.56738	

We examine the relation:

$$l_1 = C_0\,(a_i+1) + b_i + ((a_j+1)b_i + b_j)\left(1 + \sum_{k=1}^{k=5} \prod_{l=1, l\neq i, l\neq j}^{l=k} (a_l+1)\right) \Leftarrow$$

$$r_1 = C_0\,(a_j+1) + b_j + ((a_i+1)b_j + b_i)\left(1 + \sum_{k=1}^{k=5} \prod_{l=1, l\neq i, l\neq j}^{l=k} (a_l+1)\right)$$

for $j = 1, 2, \ldots, 5$ and $i \neq j$.

The results of relations are shown in Table 3.

Table 3

value l_1	value r_1	i	j	result of relation
5599.92	7568.304	1	2	true
12595.22	21551.679	1	3	true
4162.741	4511.450	1	4	true
6831.216	9680.073	1	5	true
7568.304	5599.992	2	1	false
12540.749	22835.790	2	3	true
8203.488	5481.717	2	4	false
9808.306	10571.599	2	5	true
21551.679	12595.228	3	1	false
22835.790	12549.749	3	2	false
22449.685	8189.802	3	4	false
24443.491	12948.251	3	5	false
4511.450	4162.741	4	1	false
5481.717	8203.488	4	2	true
8189.802	22449.685	4	3	true
63897.827	10654.154	4	5	true
9680.073	6831.216	5	1	false
10571.599	9808.306	5	2	false
12948.251	24443.491	5	3	true
10654.154	6389.782	5	4	false

The model M_1 occupies the first position of optimal permutation. There are four models that are described in Table 4.

Table 4

i	a_i	b_i	C_0
0			5.13799
1	2.86105	20.25810	
2	3.27292	67.16544	
3	4.31869	16.17955	
4	5.37224	42.56738	

We can write the relation

$$l_2 = C_0\,(a_i+1)+b_i+((a_j+1)b_i+b_j)\left(1+\sum_{k=1}^{k=4}\prod_{l=1,l\neq i,l\neq j}^{l=k}(a_l+1)\right) \Leftarrow$$

$$r_2 = C_0\,(a_j+1)+b_j+((a_i+1)b_j+b_i)\left(1+\sum_{k=1}^{k=4}\prod_{l=1,l\neq i,l\neq j}^{l=k}(a_l+1)\right)$$

for $j = 1, 2, 3, 4$ and $i \neq j$,
where

$$C_0 := C_0 + C_0(a_1+1) + b_1.$$

The results of relations for $i, j = 1, 2, 3, 4$ are shown in Table 5.

Table 5

value l_2	*value r_2*	*i*	*j*	*result of relation*
6221.548	11331.502	1	2	true
4067.817	2732.252	1	3	false
4846.361	5244.331	1	4	true
11331.502	6221.547	2	1	false
11091.537	4059.092	2	3	false
12039.865	6400.437	2	4	false
2732.252	4067.817	3	1	true
4059.509	11091.537	3	2	true
3154.815	5256.623	3	4	true
5244.331	4846.361	4	1	false
6400.437	12039.865	4	2	true
5256.623	3154.815	4	3	false

The model M_4 occupies the second position of optimal permutation. Now, we have only 3 models, and $C_0 = 43.0695$, see Table 6.

Table 6

i	a_1	b_1	C_0
0			43.0695
1	2.86105	20.25810	
2	3.27292	67.16544	
3	5.37224	42.56738	

We examine the relation:

$$l_3 = C_0\,(a_i+1)+b_i+((a_j+1)b_i+b_j)\left(1+\sum_{k=1}^{k=3}\prod_{l=1,l\neq i,l\neq j}^{l=k}(a_l+1)\right) \Leftarrow$$

$$r_3 = C_0\,(a_j+1)+b_j+((a_i+1)b_j+b_i)\left(1+\sum_{k=1}^{k=3}\prod_{l=1,l\neq i,l\neq j}^{l=k}(a_l+1)\right)$$

for $j = 1, 2, 3, 4$ and $i \neq j$,

where $$C_0 := C_0 + C_0(a_3 + 1) + b_3.$$

The results of relations for $i, j = 1, 2, 3$, are shown in Table 7.

Table 7

value l_3	•value r_3	i	j	result of relation
1361.226	2358.158	1	2	true
1133.049	1358.733	1	3	true
2358.158	1361.226	2	1	false
2584.398	1595.941	2	3	false
1358.733	1133.049	3	1	false
1595.941	2584.398	3	2	true

The model M_2 occupies the third position of optimal permutation. At the end there are only two models, and $C_0 = 188.24$, see Table 8.

Table 8

i	a_i	b_i	C_0
0			188.24
1	3.27292	67.16544	
2	5.37224	42.56738	

We examine the relation $l_4 = C_0(a_1 + 1) + (a_2 + 1)b_1 \quad r_4 = C_0(a_2 + 1) + (a_1 + 1)b_2)$.

The model M_3 occupies the fourth position of optimal permutation.

The optimal permutation has the form M_1, M_4, M_2, M_3, M_5.

The computational complexity of scheduling problem

$$n \mid p_i(C_{i-1}) = a_i C_{i-1} + b_i;\ a_i > 0;\ b_i > 0;\ C_0 \geq 0 \mid \min \sum C_i$$

is

$$O\left(\sum_{i=2}^{i=n} (i^2 - i) n \log n\right).$$

5. Remarks

All sequencing problems are concerned with the assignment of tasks to appropriate servers over time, so as to meet certain structural and time constraints and some performance optimisation criteria. Over the last four decades considerable study has been given by scientists and engineers to sequencing and scheduling problems due to their theoretical interest (combinatorial optimisation, complexity analysis, worst-case analysis) and practical value (e.g., in FMS, CIM and multiprocessor systems) [20].

Bibliography

[1] Aho A.V., Hopcroft J.E., Ullman J.D.: *The Design and Analysis of Computer Algorithms*, Addison – Wesley, Reading, MA. 1974

[2] Berge C.: *The Theory of Graphs and its Applications* New York, John Wiley 1962

[3] Cook J.: *The Complexity of theorem proving procedures* Proc. Third Annual ACM Symp. on Theory of Computing Machinery, New York, 1971, 151–158

[4] Edmonds J.: *Paths, trees, and flowers* Canad. J. Math., 17, 1965, 449–467

[5] Edmonds J., Karp R.M.: *Theoretical improvements in algorithmic efficiency for network flow problems* J. Assoc. Comput. Mach., 19, 1972, 248–264

[6] Garey M.R., Johnson D.S.: *Approximation algorithms for combinatorial problems: an annotated bibliography* Algorithms and Complexity: New Directions and Recent Results, J.F. Traub, ed., New York, Academic Press 1976, 41–52

[7] Garey M.R., Johnson D.S., Stockmeyer L.: *Some simplified NP – complete graph problems* Theoret. Comput. Sci., 1, 1976, 237–267

[8] Karp R.M.: *Reducibility among combinatorial problems Complexity and Computer Computations*, R.E. Miller, and J.W. Thatcher, eds., New York, Plenum Press 1972, 85–104

[9] Knuth D.E.: *The art of computer programming* Volume 1: Fundamental Algorithms, Addison – Wesley, Reading, MA. 1968

[10] Knuth D.E.: *The art of computer programming* Volume 2: Seminumerical Algorithms. Addison – Wesley, Reading, MA. 1973

[11] Knuth D.E.: *Optimum binary search trees*, Acta Informat. 1, 1971, 14–25

[12] Turing A.M.: *On computable numbers, with an application to the Entcheidungs problem* Proc. London Math. Society, ser. 2, 42. 1937, 230–265; corrections, 43, 544–546

[13] Wirth N.: *The programming language Pascal* Acta Informat., 1, 1971, 35–63

[14] Baker K.R.: *Introduction to Sequencing and Scheduling*. New York, Wiley 1974

[15] Conway, R.W., Maxwell W.L., Miller L.W.: *Theory of scheduling*. Addison – Wesley, Reading Mass. 1967

[16] French S.: *Sequencing and Scheduling*. New York, Wiley 1986

[17] Filar T.: *Application of the Green Function in a Numeric Method for Heating Control of Massive Charges*. AGH Kraków 1982 Scientific Bulletins No. 799, 1982

[18] Graham R.L., Lawler E.I., Lenstra J.K., Rinnooy Kan A.H.G.: *Optimisation and approximation in deterministic sequencing and scheduling*. Annals of Discrete Mathematics. 5, 1981, 287–326

[19] Knuth D.E.: *The Art of Computer Programming (Sorting and Searching)*. Reading, Mass. Addison – Wesley 1973

[20] Rinnooy Kan A. H. G.: *A Machine scheduling problems*. Stenfet Krose B.V., Leiden, Netherlands, 1976

[21] Wajs W.: *Optimal Control for Linear and Non-linear Dynamic Processes*. IFAC Symposia Series, XI World Congress IFAC, Tallinn, vol. III, pp. 517–523, 1991

[22] Cobham A.: *The intrinsic computational difficult of functions*. Proc. 1964 Congress for Logic Methodology and Philosophy of Science North–Holland 24–30, 1964
[23] Borodin A.: *Computational complexity: theory and practice, in A. V. Aho (ed.) Currents in the Theory of Computing*. Prentice–Hall 35–89. 1973 Proc. 1964 Congress for Logic Methodology and Philosophy of Science North–Holland 24–30, 1964
[24] Fisher R.A.: *Statistical Methods for Research Workers.,* Oliver and Boyd Ltd., Edinburgh and London, 1948
[25] Fisher R.A., Yates F.: *Statistical Tables for Biological, Agricultural and Medical Research.* Edinburg and London, Oliver and Boyd Ltd. 1953
[26] Volk W.: *Applied Statistics for Engineers.* New York, McGraw–Hill Inc. 1969

Recenzent
prof. dr hab. inż. Konrad Wala