RESEARCHES IN IDENTIFICATION OF LOGICAL AND PROBABILISTIC RISK MODELS WITH GROUPS OF INCOMPATIBLE EVENTS

Solojentsev E.D., Rybakov A.V.

Institute of Mechanical Engineering Problems of RAS,
sol@sapr.ipme.ru

Abstract: In this paper the results of the researches in identification of the logical and probabilistic (LP) risk models with groups of incompatible events are presented. The dependence of the criterion function on several parameters has been investigated. The parameters include: the total number of optimisations, the amplitude of parameters increments, the initial value of the criterion function (CF), the choice of identical or different amplitudes of increments for different parameters, objects risks distribution. An effective technology of defining the global extreme in the identification of LP-risk model for the calculation time, appreciable to practice has been suggested.

Key words: risk, logic, probability, model, identification, incompatible events

The logical and probabilistic risk models are almost twice as accurate and have seven times better robustness than other known classification methods [1,2]. However the task of multi-parameter and multi-criteria optimisation for training LP-models is characterised by exclusive difficulty [1-3]. In the process of identification of LP-risk models in business according to statistical data there arise a number of additional features and difficulties [1,2]:

- The criterion function $F_{\text{max}}$ (CF) is a number of correctly recognised good and bad objects, i.e. it accepts the integer values and it is stepped;
- CF has some local extrema, and depends on the high number of real positive arguments;
- The derivatives of the criterion function with respect to probabilities $P_{1jr}$ cannot be computed.

For each event-grade in GIE we consider three probabilities: $W_{jr}$ is the relative frequency of the grade in the objects of the “object-signs” table, $P_{1jr}$ is the probability of the event-grade in GIE, $P_{jr}$ is the probability of the event-grade to be substituted into the probability formula. The sums of the probabilities both $W_{jr}$ and $P_{1jr}$ in GIE equal 1. Connection of these probabilities are considered in [1].
The criterion function $F_{\text{max}}$, presented in Fig.1, depends only on two arguments and changes with steps equal to 2. The platforms have different sizes. The arguments $P_{11}$ and $P_{12}$ belong to the interval [0,1], but their sizes can differ substantially. While approaching the extreme the platforms decrease in size.

The optimisation can get «stick» at any «platform», not reaching the maximum or crossing the maximum. The character of changing the criterion function in the multivariate space remains the same. Let us remind that the optimisation arguments space dimension for the credit risk LP-model equals 94 [1].

1. IDENTIFICATION OF LP-RISK MODELS

The risk object is described by a large number of signs, every sign has several grades. These signs and grades correspond to random events, which lead to a failure [1,2]. The events-signs ($j=1,n$) have logical connections and events-grades for each event-sign ($r=1,N_j$) form groups of incompatible events (GIE).

The identification of the P-risk model consists in the determination of optimal probabilities $P_{jr}$, $r=1,N_j; j=1,n$, corresponding to events-grades. Let us formulate the identification (training) problem for a B-risk model [1,2].

Available data: the ‘object-signs’ table with $N_g$ good and $N_b$ bad objects and the risk B-model;

Expected results: to determine the probabilities of $P_{jr}$, $r=1,N_j; j=1,n$ for events-grades and the acceptable risk $P_{ad}$, dividing the objects into good and bad according the amount of risk.

We need: to maximise the criterion function, which is the number of correctly classified objects:

$$F = N_{bs} + N_{gs} \Rightarrow \text{MAX},$$

where $N_{gs}$ and $N_{bs}$ are the numbers of objects classified as good and bad using both by statistics and the P-risk model (both estimates should coincide). From (1) it follows, that the errors or accuracy indicators of the P-risk model in the classification of good $E_g$ and bad $E_b$ objects and in the classification of the whole set $E_m$ are equal:

$$E_g = (N_g - N_{gs}) / N_g; E_b = (N_b - N_{bs}) / N_b; E_m = (N - F) / N.$$

Assumed restrictions:

1) probabilities $P_{jr}$ and $P_{1jr}$ must satisfy the stipulation:

$$0 < P_{jr} < 1, j=1,n; r=1,N_j.$$

2) the average risks of objects $P_m$ based on the P-risk model and on the table $P_{av}$ must be equal; while training the P-risk model we must correct the $P_{jr}$ probabilities on every step of iterative training under the formula

$$P_{jr} = P_{jr} \cdot (P_{av} / P_m); j=1,n; r=1,N_j.$$

3) the acceptable risk $P_{ad}$ must be determined with the given ratio of incorrectly classified good and bad objects, because of non-equivalence losses at their wrong classification:

$$E_{gb} = (N_g - N_{gs}) / (N_b - N_{bs}).$$

2. OPTIMISATION IN THE IDENTIFICATION TASK

Identification of the LP-risk model by the random search method is based on the ideas used in the training of neural networks [4]. With reference to the identification task of the LP-risk model, the following formula for the calculation of the changes of events-grades probabilities may be put down:

$$dP_{1jr} = K_1 \cdot (1 / N_j) \cdot tg(K_3); j=1,n; r=1,N_j.$$
where: \( K_1 \) is a coefficient; \( N_t \) is the current number of optimisation; \( K_3 \) is a random number from \(-\pi/2, +\pi/2\), \( n \) is a number of events-signs, \( N_j \) is a number of events-grades in each GIE, i.a. for every event–sign.

In the formula (6) the CF is a current error in training. The number of optimisations \( N_t \) before the end of the training process, can be very big. The «tangent» operation is the consequence of the training error distribution recording to Cauchy. Theoretically, this error is distributed according to the normal law, but not spend a lot of time on tabulated values calculation, we use the distribution of the training error under the Cauchy’s law. It allows to reduce in 100 times the calculation time, which otherwise, for real problems, would continue for days and weeks.

For failure risk LP-model training the following modification of the formula (6) is suggested [1]:

\[
(7) \\
\frac{dP_1}{P} = K_3 \cdot (N_{opt} - N_i) \cdot \tan(K_3) , \\
j = 1, n_f = 1, N_j,
\]

where: \( N_{opt} \) is the given number of optimisations. The new values of \( P_{1,f} \) and \( P_{f} \), obtained at \( F > F_{max} \) on every step \( N_t \) of optimisation are considered optimal and saved.

In the LP-risk model identification task, the criterion function cannot exceed the total number of objects in the statistical data. The formula (7) is quite applicable, but the time of calculation is too big (about 10 hours for a session of optimisation).

To reduce the time of calculation, in the formula (7) the "tangent" operation is eliminated. As a result the following expression is obtained [3]:

\[
(8) \\
dP_1 = K_1 \cdot (N_{opt} - N_i) \cdot K_3 , \\
j = 1, n_f = 1, N_j.
\]

Using (7,8) the optimization happens so: if \( F > F_{max} \), then we remember the new \( P_{1,f} \) and \( P_{f} \). If the criterion function does not strictly increase after the chosen number of trials \( N_{mc} \) in Monte-Karlo, then \( F_{max} \) is reduced by 2-4 units and optimisation continues.

In spite of the investigation in optimisation, carried out before, where the formulas (7) and (8) were used [1,2], the problem of optimisation in the identification task of LP-risk models is far from the final solution. The following fact proves it. In one of the research with the huge number of optimisations \( N_{opt}=245000 \) and with the constant, almost optimal, value of the increment \( dP_{1,f} \), we obtained \( F_{max} = 824 \) instead of \( F_{max} = 810 \) at the usual number of optimisations \( N_{opt} \approx 245 \). We had to carry out special investigations, the results of which are adduced below.

3. INVESTIGATIONS IN IDENTIFICATION / OPTIMISATION

If we generate a random number \( K_3 \) in the interval \([-1, +1]\), then the absolute values of increments of probabilities \( dP_{1,f} \), multiplied by 100, are transformed in percents (%). It is convenient, for practically it solves the problem of the evaluation of probabilities \( P_{1,f} \) accuracy. For example, if the increment is \( dP_{1,f}=0.0005 \), it equals 0.05 %. We can say that the probability \( P_{1,f} \) with the accuracy 0.05 % is evaluated.

Using the formula (8), in the beginning of optimisation we have the following maximum amplitude of probabilities increments:

\[
(9) \\
\frac{AP_{1}}{max} = K_1 \cdot N_{opt}.
\]

In the end of optimisation the maximum amplitude of probabilities increments equals 0. Let us designate the current amplitude of probabilities increments as \( AP_1 \). There is an optimal interval \( OPT \) of the amplitudes increments \( AP_1 \), which position and width are unknown (Fig. 2). For the big values of \( AP_1 \) there is a small probability of increasing \( F_{max} \), and for small values of \( AP_1 \) there is a high probability to stop at the local extreme of the reached value \( F_{max} \) (see Fig.1).
The optimisation process (of training the LP-risk model) should be long enough in the optimal OPT interval. The value of \(dN_{opt}\) duration in the optimal OPT interval is equal

\[
dN_{opt} = \frac{(OPT \times N_{opt})}{AP_{1,\max}}.
\]

It also depends on the number of optimisations \(N_{opt}\) and the maximum amplitude of the increment \(AP_{1,\max}\). The more \(N_{opt}\) is and the less \(AP_{1,\max}\) is, the longer is the duration of \(dN_{opt}\). The purpose of this work is the investigation of the dependence of the criterion function (accuracy of LP-risk model) on the following parameters in the training formula (8):

1. The number of optimisations \(N_{opt}\);
2. The increment minimum amplitude \(AP_{1,\min}\), at which the optimisation is still possible;
3. The initial value of the criterion function \(F_{beg}\);
4. The choice of identical or different amplitudes \(AP_{1}\) for different grades;
5. The increment maximum amplitude \(AP_{1,\max}\);
6. Objects risk distribution in the statistical data.

Let us illustrate it. A question arises, whether to choose the identical or different values of increments amplitudes \(AP_{1}\) for all events-grades? In other words, whether the amplitudes \(AP_{1,fr}\) should depend on the values of probabilities \(P_{1,fr}\)? In the training formulas of the LP-risk model (7) and (8) the increments amplitudes \(AP_{1,fr}\) are identical for all events-grades and do not depend on the values of their probabilities \(P_{1,fr}\). The increments \(dP_{1,fr}\) differ only because of the random simulation of the \(K_{3}\) coefficient.

The model investigations for the LP-model of the credit risk were made on the PC. The credit risk structural LP-model has 20 events-signs (correspondingly GIE) and 94 events-grades. The credit risk L-function is [1,2] :

\[
Y = X_{1} \cup X_{2} \cup \ldots \cup X_{20}
\]

Verbally it can be formulated as follows: a failure occurs, if any one, or any two, … or all initiating events happen. After the orthogonalization of the L-function (11) the following P-risk model for the evaluation of the credit risk has been obtained:

\[
P = P_{1} + P_{2}Q_{1} + P_{3}Q_{2} + \ldots
\]

The investigations were carried out in a set of 1000 credits, 700 of which were good and 300 - bad [5]. For calculation investigations we used the Software, designed in the object-oriented languages Visual C++ and Java.
3.1 The choice of parameters $N_{opt}$, $API_{min}$, $F_{beg}$

In comparison with the optimal variant $F_{max} = 824$, the initial variant had the probabilities $P_{1jr}$ without the last four signs. So the optimisation starts at $F_{beg} = 690-760$. Such solution allowed to reduce calculation time.

The calculations were made for two values of increments maximum amplitudes: 1) $API_{max} = 0.05$ (5%), 2) $API_{max} = 0.1$ (10%). We used the following numbers of optimisations $N_{opt}$: 150, 300, 500, 750, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000.

The results of investigations presented in Table 1 (Var.2-21) and Fig.3, allow to make the following conclusions:

1) The criterion function $F_{max}$ (column 6 in Table 1 and Fig.3) asymptotically increases with the growth of the number of $N_{opt}$ optimisation;

2) The minimum amplitude $API_{min}$ (column 9) equals approximately 0.0025 (0.25%); at the smaller values of $API_{min}$ the optimisation does not happen and the number of the last optimisation $N_{end}$ (column 10) is less, than the given number of $N_{opt}$ optimisations. It is necessary to modify the law of the change of $API$ during the training process, adding the constant line $API_{min}$ (Fig.4). It increases the chance to get the greater value of $F_{max}$;

3) The big value of $N_{opt}$ can lead to the disappearance of the B-C line (Fig. 4), which undoubtedly will deteriorate the process of optimisation.

4) The initial value of $F_{beg}$ (column 5) should not be lowered, as it often leads to low final values of $F_{max}$ (Fig. 5) because of the unsuccessful trajectory of optimisation process; in the considered case it is possible to accept $F_{beg} = 750-760$.

Taking into consideration the just made conclusions, instead of the formula (8) the following formula for training the LP-risk model is suggested:

\[ (13) \quad \text{If} \quad API < API_{min}, \quad \text{then} \quad dP_{1jr} = API_{min} \times K_{3}, \]

\[ \text{If} \quad API > API_{min}, \quad \text{then} \quad dP_{1jr} = K_{1} \times (N_{opt} - N_{b}) \times K_{3}. \]

The optimisation results using the formula (13) under $API_{min} = 0.0025$ (0.25%), different $API_{max} = 0.098, 0.09, 0.03$ (9.8%, 9%, 3%), a rather large number of optimisations $N_{opt} = 5000-12000$ and the high $F_{beg} = 745$ in Table 1 (var. 22-24) are shown. In all variants high values of $F_{max} = 812-822$ have been obtained.

### Table 1. The investigations results in the choice of optimisation parameters

<table>
<thead>
<tr>
<th>$N$</th>
<th>$N_{opt}$</th>
<th>$K_{1}$</th>
<th>$API_{max}$</th>
<th>$F_{beg}$</th>
<th>$F_{max}$</th>
<th>$dP_{c}$</th>
<th>$API_{min}$</th>
<th>$N_{end}$</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2000</td>
<td>0.0001</td>
<td>0.2</td>
<td>776</td>
<td>786</td>
<td>0.204</td>
<td>0.1987</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>0.000165</td>
<td>0.05</td>
<td>756</td>
<td>794</td>
<td>0.1969</td>
<td>0.00198</td>
<td>289</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>300</td>
<td>0.00033</td>
<td>0.1</td>
<td>712</td>
<td>790</td>
<td>0.221</td>
<td>0.00429</td>
<td>288</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>750</td>
<td>0.0000665</td>
<td>0.05</td>
<td>756</td>
<td>802</td>
<td>0.1641</td>
<td>0.00545</td>
<td>669</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>750</td>
<td>0.000133</td>
<td>0.1</td>
<td>692</td>
<td>790</td>
<td>0.2052</td>
<td>0.01316</td>
<td>652</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1000</td>
<td>0.00005</td>
<td>0.05</td>
<td>750</td>
<td>802</td>
<td>0.1867</td>
<td>0.00350</td>
<td>931</td>
<td>(3)</td>
</tr>
<tr>
<td>7</td>
<td>1000</td>
<td>0.0001</td>
<td>0.1</td>
<td>708</td>
<td>792</td>
<td>0.2174</td>
<td>0.01580</td>
<td>843</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2000</td>
<td>0.000025</td>
<td>0.05</td>
<td>776</td>
<td>808</td>
<td>0.1595</td>
<td>0.00747</td>
<td>1702</td>
<td>(3)</td>
</tr>
<tr>
<td>9</td>
<td>2000</td>
<td>0.0005</td>
<td>0.1</td>
<td>724</td>
<td>798</td>
<td>0.1802</td>
<td>0.01405</td>
<td>1720</td>
<td>(3)</td>
</tr>
<tr>
<td>10</td>
<td>3000</td>
<td>0.0000166</td>
<td>0.05</td>
<td>748</td>
<td>806</td>
<td>0.1867</td>
<td>0.00699</td>
<td>2581</td>
<td>(3)</td>
</tr>
<tr>
<td>11</td>
<td>3000</td>
<td>0.00033</td>
<td>0.1</td>
<td>708</td>
<td>806</td>
<td>0.1867</td>
<td>0.00501</td>
<td>2849</td>
<td>(3)</td>
</tr>
<tr>
<td>12</td>
<td>4000</td>
<td>0.000025</td>
<td>0.05</td>
<td>744</td>
<td>812</td>
<td>0.1945</td>
<td>0.00791</td>
<td>3368</td>
<td>(3)</td>
</tr>
<tr>
<td>13</td>
<td>4000</td>
<td>0.00025</td>
<td>0.1</td>
<td>740</td>
<td>802</td>
<td>0.2121</td>
<td>0.00862</td>
<td>3563</td>
<td>(3)</td>
</tr>
<tr>
<td>14</td>
<td>5000</td>
<td>0.00001</td>
<td>0.05</td>
<td>754</td>
<td>806</td>
<td>0.1663</td>
<td>0.00556</td>
<td>4445</td>
<td>(3)</td>
</tr>
<tr>
<td>15</td>
<td>5000</td>
<td>0.0002</td>
<td>0.1</td>
<td>738</td>
<td>803</td>
<td>0.1586</td>
<td>0.00400</td>
<td>4801</td>
<td>(3)</td>
</tr>
<tr>
<td>16</td>
<td>6000</td>
<td>0.00016</td>
<td>0.1</td>
<td>710</td>
<td>810</td>
<td>0.1598</td>
<td>0.00625</td>
<td>5610</td>
<td>(3)</td>
</tr>
<tr>
<td>17</td>
<td>6000</td>
<td>0.000133</td>
<td>0.109</td>
<td>736</td>
<td>810</td>
<td>0.1618</td>
<td>0.00495</td>
<td>5730</td>
<td>(3)</td>
</tr>
<tr>
<td>18</td>
<td>7000</td>
<td>0.000071</td>
<td>0.05</td>
<td>764</td>
<td>810</td>
<td>0.2096</td>
<td>0.00407</td>
<td>6430</td>
<td>(3)</td>
</tr>
<tr>
<td>19</td>
<td>7000</td>
<td>0.000142</td>
<td>0.1</td>
<td>734</td>
<td>810</td>
<td>0.1692</td>
<td>0.00745</td>
<td>6479</td>
<td>(3)</td>
</tr>
<tr>
<td>20</td>
<td>8000</td>
<td>0.000062</td>
<td>0.05</td>
<td>764</td>
<td>810</td>
<td>0.1735</td>
<td>0.00985</td>
<td>6425</td>
<td>(3)</td>
</tr>
<tr>
<td>N</td>
<td>N_{opt}</td>
<td>K_1</td>
<td>AP_1^{max}</td>
<td>F_{beg}</td>
<td>F_{max}</td>
<td>dP_c</td>
<td>AP_1^{min}</td>
<td>N_{end}</td>
<td>Notes</td>
</tr>
<tr>
<td>----</td>
<td>---------</td>
<td>-------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
<td>------------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>21</td>
<td>8000</td>
<td>0.0000125</td>
<td>0.1</td>
<td>718</td>
<td>814</td>
<td>0.1802</td>
<td>0.00286</td>
<td>7772</td>
<td>(3)</td>
</tr>
<tr>
<td>22</td>
<td>12000</td>
<td>0.0000075</td>
<td>0.09</td>
<td>772</td>
<td>812</td>
<td>0.1737</td>
<td>0.0025</td>
<td>11754</td>
<td>(10)</td>
</tr>
<tr>
<td>23</td>
<td>8000</td>
<td>0.00000375</td>
<td>0.03</td>
<td>780</td>
<td>820</td>
<td>0.1526</td>
<td>0.0025</td>
<td>7662</td>
<td>(10)</td>
</tr>
<tr>
<td>24</td>
<td>8000</td>
<td>0.00000875</td>
<td>0.07</td>
<td>744</td>
<td>814</td>
<td>0.173</td>
<td>0.0025</td>
<td>7801</td>
<td>(10)</td>
</tr>
<tr>
<td>25</td>
<td>5000</td>
<td>0.0000043</td>
<td>0.0215</td>
<td>812</td>
<td>820</td>
<td>0.1462</td>
<td>0.0025</td>
<td>23</td>
<td>(13)</td>
</tr>
<tr>
<td>26</td>
<td>5000</td>
<td>0.00000043</td>
<td>0.0025</td>
<td>810</td>
<td>824</td>
<td>0.1511</td>
<td>0.0025</td>
<td>34</td>
<td>(13)</td>
</tr>
<tr>
<td>27</td>
<td>8000</td>
<td>0.00000002</td>
<td>0.0025</td>
<td>810</td>
<td>826</td>
<td>0.1538</td>
<td>0.0025</td>
<td>678</td>
<td>(13)</td>
</tr>
<tr>
<td>28</td>
<td>8000</td>
<td>0.0000025</td>
<td>0.00458</td>
<td>806</td>
<td>822</td>
<td>0.1604</td>
<td>0.00609</td>
<td>507</td>
<td>(13)</td>
</tr>
<tr>
<td>29</td>
<td>8000</td>
<td>0.00000312</td>
<td>0.00572</td>
<td>806</td>
<td>822</td>
<td>0.1677</td>
<td>0.00452</td>
<td>1757</td>
<td>(13)</td>
</tr>
</tbody>
</table>

**Fig. 3.** The dependence of the criterion function $F_{\text{max}}$ on the number of optimizations $N_{\text{opt}}$.  

**Fig. 4.** The graph of the current amplitude of increment $AP_1$ modification.  

**Fig. 5.** Dependence of the criterion function $F_{\text{max}}$ on its initial value.
3.2 Different amplitudes $API_{jr}$ of increments for different grades

It should be noted, that the probabilities $P1_{jr}$ depend on: a number of grades in GIE, the frequencies of $W_{jr}$ grades in objects and the grades contributions in the classification errors of objects. In the formula of training the LP-risk model (8) the increments amplitudes $API_{jr}$ are identical for all events-grades and do not depend on the magnitude of their probabilities $P1_{jr}$.

Let us change the formula of the training LP-risk model so that it takes into account the value of probability for each grade

$$dP1_{jr} = K_1*(N_{opt} - N_1)*K_3 * P1_{jr}.$$  \(14\)

Here the amplitudes for every event grade are equal

$$API_{jr} = K_1*(N_{opt} - N_1)*P1_{jr}$$

and the formula (14) can be the following:

$$dP1_{jr} = API_{jr} * K_3.$$  \(16\)

Let us also put down the formula (14) with the following modification:

$$dP1_{jr} = K_1*(N_{opt} - N_1)*((1-a) + a)*P1_{jr})* K_3,$$  \(17\)

where $a$ is a coefficient from the interval $0 < a < 1$. It determines the formula (8) at $a=0$, the formula (14) at $a=1$ and all the modifications at other values of $a$.

In the formula (13) let us take into account the limitations, introduced earlier in the formula (8), and we shall get the following expression for training the LP-risk model:

$$\begin{align*}
\text{If } API_{jr} < API_{min} , \text{ then } dP1_{jr} = API_{min} , \\
\text{If } API_{jr} > API_{min} , \text{ then } dP1_{jr} = K_1*(N_{opt} - N_1)*((1-a) + a)*P1_{jr})* K_3,
\end{align*}$$

The investigations results in optimisation using the formula (18) at $a=1 (API_{max} = 2.15 \%, 0.25 \%, 0.45 \%, 0.57\%)$ are represented in Table 1 (Var.25-29). They show that the high values of the $F_{max} = 822-826$ can be obtained at the limited number of optimisation attempts $N_{end}$ (column 10). Actually the first optimisation already gives the high value of CF ($F_{beg}=806-810$). The optimisation process ends at $N_{end}=23-1750$ instead of the given numbers of optimisations $N_{opt}=5000-8000$ (column 6). It seems, that the number of optimisations $N_{opt}$ can be essentially reduced. To verify this hypothesis some extra investigations have been carried out.

The investigations were carried out at small numbers of optimisations $N_{opt} = 600, 450, 300, 150, 100, 50$ and $K_1=0.00033, 0.00025, 0.00015, 0.0001$. The increments maximum amplitude $API_{max}$ varied in an interval $0.5\% - 20\%$ from $P1_{jr}$. In Table 2 the CF values and the difference between maximum and minimum risks of objects in the statistics $F_{max} / APc$ are shown. The results of the investigations should be considered as good ($F_{max} = 810-822$) and completely confirming the effectiveness of the formulas (14), (17) and (18).

Also the investigations of the influence of $a$ parameter on the optimisation results have been carried out. It was done at the small numbers of optimisations $N_{opt}=150$ and $K_1=0.00015$. The maximum amplitude of an increment $API_{max}$ equals $0.0225*P1_{jr}$.

<table>
<thead>
<tr>
<th>Number of optimisations, $N_{opt}$</th>
<th>$K_1=0.00033$</th>
<th>$K_1=0.00025$</th>
<th>$K_1=0.00015$</th>
<th>$K_1=0.0001$</th>
</tr>
</thead>
<tbody>
<tr>
<td>600</td>
<td>798 / 0.248</td>
<td>796 / 0.225</td>
<td>810 / 0.180</td>
<td>810 / 0.149</td>
</tr>
<tr>
<td>450</td>
<td>802 / 0.217</td>
<td>804 / 0.187</td>
<td>814 / 0.162</td>
<td>819 / 0.161</td>
</tr>
<tr>
<td>300</td>
<td>810 / 0.146</td>
<td>810 / 0.174</td>
<td>816 / 0.147</td>
<td>820 / 0.162</td>
</tr>
<tr>
<td>225</td>
<td>810 / 0.154</td>
<td>811 / 0.152</td>
<td>818 / 0.148</td>
<td>821 / 0.146</td>
</tr>
<tr>
<td>150</td>
<td>816 / 0.145</td>
<td>820 / 0.156</td>
<td>822 / 0.148</td>
<td>822 / 0.147</td>
</tr>
<tr>
<td>100</td>
<td>818 / 0.146</td>
<td>820 / 0.149</td>
<td>820 / 0.151</td>
<td>820 / 0.153</td>
</tr>
<tr>
<td>50</td>
<td>822 / 0.151</td>
<td>820 / 0.146</td>
<td>820 / 0.152</td>
<td>820 / 0.148</td>
</tr>
</tbody>
</table>

The investigations results, represented in Table 3, also confirm the effectiveness of the formulas (14),(17) and (18) at $a=1$. Really, at $a=1 F_{max}$ equals 820, and at $a=0 F_{max}$ equals 802.
### Table 3. Values $F_{\max}$ at different values of $a$

<table>
<thead>
<tr>
<th>Value</th>
<th>$a$</th>
<th>0.0</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
<th>0.6</th>
<th>0.7</th>
<th>0.8</th>
<th>1.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>$F_{\max}$</td>
<td>802</td>
<td>800</td>
<td>798</td>
<td>804</td>
<td>808</td>
<td>810</td>
<td>808</td>
<td>810</td>
<td>818</td>
<td>820</td>
</tr>
</tbody>
</table>

#### 3.3 Determination of the amplitude $AP_{1\max}$ and the global extreme $F_{\max}$

Let us consider again the choice of the increment maximum amplitude of probabilities $AP_{1\max}$. The results of the change of $F_{\max}$ at the change of $AP_{1\max}=K_1*N_{opt}$ in the interval 0.5-20% of $P_{1jr}$ are represented in Table 2. They demonstrate that the higher is $AP_{1\max}$, the less is $F_{\max}$. In Fig.6 the dynamics and the results of optimisation for five variants, having $N_{opt}=2000$, are shown:

- **Variant 1:** $AP_{1\max}=0.05$, $F_{\max}=808$ (Var.8 in Table 1), training under the formula (3);
- **Variant 2:** $AP_{1\max}=0.1$, $F_{\max}=804$ (Var.9 in Table 1), training under the formula (3);
- **Variant 3:** $AP_{1\max}=0.05$, $F_{\max}=820$, training under the formula (14) with $a=1$;
- **Variant 4:** $AP_{1\max}=0.1$, $F_{\max}=786$, training under the formula (14) with $a=1$;
- **Variant 5:** $AP_{1\max}=0.2$, $F_{\max}=776$ (Var.1 in Table 1), training under the formula (14).
Variants 4 and 5 with high $API_{max}$, despite using the effective formula (18) and $a=1$, have bad training dynamics and results. In these variants CF are correspondingly 786 and 804. The optimisation process finishes early, ($N_{end}$=1608 and $N_{end}$=20). Additional optimisation attempts $N_{opt} - N_{end}$ have not increased CF. This example confirms that the increment amplitude $API_{max}$ should not be more than 0.02 - 0.05 (2-5 %).

We check the calculation of the global extreme of the CF by the graph (Fig.7). The function $F_{max}$ has an extreme at some value of the difference $APc$ between the maximum risk and the minimum risk of objects in statistics [2]. This difference, constructed for variants of computational investigations, presented in Table 1 and 2, demonstrates the robustness (stability) of solutions at a small dispersion of $APc$ in the area of the global extreme of CF.

4. CONCLUSION

In the investigations the following main results have been obtained:
1. The effective technology of the criterion function global extreme search in the tasks of identification of LP-risk models under statistical data has been offered. It permits to solve the task of multi-parameter multi-criteria optimisation with integer CF for the time, applicable to practice (less than before).
2. We suggest to generate in the training formula a random number $K_{3}$ in the interval [-1, +1]. It permits to consider the absolute values of increments $dP_{1jr}$, multiplied by 100, in percents (%) and to estimate the accuracy of probabilities $P_{1jr}$.
3. In the technology of the CF global extreme search, the following regularities of changing the CF should be used:
   • The CF asymptotically increases with the growth of $N_{opt}$ optimisation number;
   • The minimum amplitude $API_{min}$ of probabilities $P_{1jr}$ increments is established by 2-3 test calculations; at smaller values of $API_{min}$ the optimisation does not happen (less than 0.25 %);
   • The initial CF $F_{beg}$ should not be lowered, as low values more often result in low final values of $F_{max}$ because of the unsuccessful trajectory of the optimisation process;
   • Maximum amplitude of increments of $API_{max}$ must not exceed 0.02 - 0.05 (2-5%), as the training speed lows down and the value of the CF $F_{max}$ becomes less.
4. For the criterion function global extreme search new, more effective formulas of training (14), (17), (18) have been suggested; they use different amplitudes of increments for probabilities of different events-grades.
5. It has been confirmed that we can test the determination of the global extreme of CF $F_{max}$ by the graph of change of $F_{max}$ in the function of difference $APc$ between maximum and minimum risks of objects in statistics. The function $F_{max}$ has an extreme at a certain value of $APc$.

REFERENCES