Вісник Тернопільського національного технічного університету
https://doi.org/10.33108/visnyk_tntu
Scientific Journal of the Ternopil National Technical University
2023, № 3 (111) https://doi.org/10.33108/visnyk_tntu2023.03
*ISSN 2522-4433. Web: **visnyk.tntu.edu.ua***

UDC 51-7

CORRESPONDENCE ANALYSIS FOR DETECTING RISK FACTORS FOR CRIMINAL RECIDIVISM

Olha Kovalchuk

West Ukrainian National University, Ternopil, Ukraine

Summary. *Correspondence analysis was used in the work to identify associations between criminal recidivism and the following elements of criminal histories of criminals: sex, age at the time of the first conviction to the actual degree of punishment, age at the time of the first conviction to the suspended or actual sentence, educational level, type of employment at the time of conviction, availability of early releases, availability of suspended sentences, availability of motivation for the release. The conducted empirical analysis made it possible to draw conclusions about the existence of a direct relationship between the risk of criminal recidivism with the age at the time of the first conviction to the suspended and/or actual sentence, the level of education obtained, the type of employment, the presence of early releases, previous conditional convictions and the lack of correlation between the fact of committing repeated criminal offenses and the gender and motivation of the convicts for release.*

Key words: *correspondence analysis, associations, internal security, χ^2-test, criminal recidivism.*

https://doi.org/10.33108/visnyk_tntu2023.03.035 *Received 10.07.2023*

Statement of the problem. Scientific analysis of criminal offenses and international judicial practice indicate that the risk of criminal recidivism by persons who have already had previous convictions is significantly higher than the risk of criminal offenses by persons who have not committed criminal offenses. Prevention of repeated criminal offenses is one of the basic priorities of ensuring the personal safety of citizens and public safety in general. An integral element of an effective internal security strategy of the country is the development and improvement of effective tools for assessing the risks of criminal recidivism by previously convicted persons, the basis for which is useful knowledge extracted from the criminal histories of criminals. This knowledge can be obtained using mathematical modeling tools and modern information technologies based on the individual socio-demographic characteristics of convicts and the conditions of their microsocial environment. Such information should be stored in consolidated data warehouses, be reliably protected, in particular, against various types of cyber attacks [1–5].

Analysis of recent research. The fact of repeated criminal offenses committed by persons who have already served a sentence for a committed offense is one of the most important assessments of the effectiveness of criminal court sentences [6]. In the scientific literature, multifaceted studies on the identification of individual static and dynamic characteristics of convicts, which increase the risk of criminal recidivism [7, 8, 9], are widespread. D. Yukhymenko and others conducted a systematic review of risk factors for recidivism in previously convicted persons [10]. L. Jacobs and others studied the risk of recidivism in individuals suffering serious mental disorders with or without co-occurring disorders due to the use of psychoective substance [11]. C. Cuevas and others investigated the relationship of dynamic risk factors of the «risk-need-response» model with time to relapse and differences for young men and women [12]. K. Garritsen and others analyzed the risk factors of committing violent relapses in male forensic psychiatric patients [13]. J. Navarro-Perez and others determined the possibility of predicting the risk of recidivism

Corresponding author: Olha Kovalchuk; e-mail: olhakov@gmail.com ……

in juvenile offenders based on protective factors and/or risk factors. He also determined the relative weight of both types of factors in such predictions [14]. R. Heffernan and T. Ward investigated the risks and protective factors of sexual offenses based on the analysis of risk practices (purposeful actions), in which values (priorities, motivators, norms) are embedded, and which are based on human abilities [15]. Today, the analysis of various socio-economic factors associated with criminal events, the detection of deviations, the classification of patterns, and the development of effective computer models for predicting crimes applying the methods of intelligent data analysis and machine learning [16] are extremely relevant.

The objective of the paper is the identification of non-obvious relationships between the propensity of convicts to re-commit criminal offenses and their following socio-demographic characteristics and conditions of the micro-social environment: gender; age at the time of the first conviction before the actual punishment; age at the time of the first conviction to a real or conditional punishment; obtained level of education; type of employment; availability of early dismissals; the presence of suspended sentences; availability of motivation to release.

Formulation of the problem. Correspondence analysis was used to establish implicit relationships between the propensity of convicts to commit criminal recidivism and their socio-demographic characteristics and previous criminal histories [17]. The empirical analysis was carried out on the basis of official information from the criminal records of 13,000 people serving sentences in penitentiary institutions in Ukraine [18]. The following variables were used in the study:

- *Recidivism*: 0 – no,1 – yes;
- *Sex*: 0 – woman, 1 – man;
- *AFA* – age at the time of the first conviction to a real punishment: 1 – to 18 years, 2 – from 18 to 30 years; 3 – from 30 to 45 years; 4 – more than 45 years;
- *AFC* – age at the time of the first conviction to a real or conditional punishment: 1 – to 18 years, 2 – from 18 to 30 years; 3 – from 30 to 45 years; 4 – more than 45 years;
- *Education*: 0 – incomplete secondary education, 1 – secondary education, 2 – special secondary education, 3 – incomplete higher education, 4 – higher education;
- *Employment*: 0 – unemployed, 1 – part-time employment, 2 – full-time employment;
- *Early Dismissals*: 0 – no, 1 – yes;
- *Suspended Convictions*: 0 – no, 1 – yes;
- *Motivation* – motivation to release: 0 – no, 1 – yes.

Correspondence analysis is an exploratory analysis method designed to visually and numerically study the structure of high-dimensional connectivity tables. Computationally, the goal of correspondence analysis is to represent the distances between points in a lower-dimensional space. The task is to find a plane that can reproduce the original image as accurately as possible.

Correspondence analysis uses a specific interpretation of the Pearson χ^2 statistic. To evaluate the quality of the presentation of a multidimensional table in a space of smaller dimensions, the value of statistic χ^2 is calculated. In essence, correspondence analysis is a factor analysis of nominal data [17] or a dimensionality reduction method. The rows or columns of the initial table are represented by points in space; the distance χ^2 is calculated between them. In the next step, the *n*-dimensional space with a small value of *n* (usually $n = 2$) is sought, in which the calculated distances are minimally distorted, and in this sense, the structure of the initial table is reproduced as accurately as possible while preserving the relationships between features. The source information for the application of the correspondence analysis method is the Frequency Table. The following terminology is used in the compliance analysis:

Mass. The observations in the table are normalized: the relative frequencies for the table are calculated. The sum of all elements of the table becomes equal to 1 (each element is divided by the

total number of observations). The resulting standardized table shows how the mass is distributed across the cells of the table or across points of space. The sums of the rows and columns in the matrix of relative frequencies are called the mass of the row and column, respectively.

Quality. In the context of correspondence analysis, it is the quality of representation of the corresponding line point in the coordinate system defined by the corresponding number of dimensions. The quality of a point is defined as the ratio of the square of the distance from the point to the origin of the coordinates in the selected number of dimensions. The maximum number of dimensions is determined by the square of the distance from the origin of coordinates in space. Low quality means that the selected number of dimensions does not represent the corresponding row (or column).

Relative inertia. Inertia is defined as Pearson's χ^2 for a 2×2 table divided by the total number of observations. Relative inertia represents the dimensions of the user. A partial solution may represent a point reasonably well (high quality), but the same point may contribute very little to the overall inertia (a point-row whose elements are relative frequencies is similar to some row whose elements are averages over all rows).

Row & column profile. If the rows and columns of the table are completely independent, then the elements of the table can be represented using row and column sums or, in correspondence analysis terminology, using row and column profiles.

Relative Dim. *n*. This column displays the relative contribution of the corresponding point-line to the magnitude of inertia due to the corresponding dimension. This value is given for each point (row or column) and for each dimension.

Cosine ? – quality, or quadratic correlations with each dimension. This column contains the quality for each point due to the corresponding dimension. Cosine squared can be interpreted as the «correlation» between the corresponding point and dimension. This value is the square of the cosine of the angle formed by this point and the corresponding axis.

Metric coordinate system. In many cases, the term distance is used to denote the differences between the rows and columns of the matrix of relative frequencies, which, in turn, are represented in a space of smaller dimensions because of the use of correspondence analysis methods. Coordinates in the space of the smallest dimension represent distances. Unlike Euclidean distances, calculated from the relative frequencies of columns and rows, these distances are weighted.

Graphical analysis of the results is the most important part of the analysis. Usually, the horizontal axis corresponds to the maximum inertia. The graph shows the percentage of total inertia explained by a specific eigenvalue. The smaller the distance between points of the same type (rows or columns), the closer the connection. In order to establish a relationship between points of different types (between rows and columns), you need to consider the angles between them with the vertex at the center of gravity.

General rule of visual assessment of the degree of dependence:

- two arbitrary points of different types are connected by line segments with the center of gravity – a point with coordinates (0,0);
- if the angle formed is acute, the row and column are positively correlated;
- if the angle is obtuse, the correlation between the variables is inverse.
- if the angle is straight, there is no correlation.

Analysis of numerical results. We applied correspondence analysis to identify non-obvious relationships between the propensity of convicts to commit repeated criminal offenses and their following socio-demographic characteristics and conditions of the micro-social environment: gender, age at the time of the first conviction to the actual penalty, age at the time of the first conviction to the actual or suspended sentences, education, type of employment at the time of conviction, early release, suspended sentences, motivation for release.

Correspondence analysis is essentially a component decomposition of the χ^2 statistics. The main purpose is to find the space of the smallest dimension, which makes it possible to present deviations from the expected values. In Tables 1–9, the calculated eigenvalues are presented, which provide information about the number of measurements sufficient to represent qualitatively the information of the initial data tables for each of the pairs of analyzed variables.

Table 1

Eigenvalues and Inertia for all Dimensions for Variables *Recidivism* and *Sex*

Number of Dims.	Total Inertia=1.0000				
	Singular Values	Eigen-Values	Perc. of Inertia	Cumulatv Percent	Chi Squares
1	0.748679	0.560521	56.05206	56.05210	14787.78
2	0.662932	0.439479	43.94794	100.0000	11592.09

For a qualitative presentation of the Table of connections of the values of the variables *Recidivism* and *Sex*, two dimensions are sufficient. In this case, the first measurement «extracts» 56% of the total inertia, the second – 44%. Pearson's χ^2 criterion is an objective assessment of the closeness of empirical distributions to theoretical ones. An obtained level of 0.001 was reached – the results are statistically significant. Number of degrees of freedom *df* = 9, χ^2_{emp} = 26139.8. χ^2_{theor} (0.001;9) = 28.88. $\chi^2_{emp} > \chi^2_{theor}$. Therefore, it can be stated that the predicted values are close enough to the observed ones.

Table 2

Eigenvalues and Inertia for all Dimensions for Variables *Recidivism* and *AFA*

Number of Dims.	Total Inertia=2.0000				
	Singular Values	Eigen-Values	Perc. of Inertia	Cumulatv Percent	Chi Squares
1	0.833501	0.694724	34.73620	34.7362	19435.60
2	0.707107	0.500000	25.00000	59.7362	13988.00
3	0.707107	0.500000	25.00000	84.7362	13988.00
4	0.552518	0.305276	15.26380	100.0000	8540.41

Four dimensions are required to represent the point-line representation of the relational Table of values of the *Recidivism* and *AFA* variables. The first dimension represents almost 35% of the total inertia, the inclusion of the second dimension increases the "extracted" inertia to 60%, the inclusion of the third – to 85%, the fourth – to 100%; *df* = 25, χ^2_{emp} = 55952. χ^2_{theor} (0.001;25) = 52.62. $\chi^2_{emp} > \chi^2_{theor}$.

Table 3

Eigenvalues and Inertia for all Dimensions for Variables *Recidivism* and *AFC*

Number of Dims.	Total Inertia=2.0000				
	Singular Values	Eigen-Values	Perc. of Inertia	Cumulatv Percent	Chi Squares
1	0.837290	0.701054	35.05271	35.0527	19634.33
2	0.707107	0.500000	25.00000	60.0527	14003.43
3	0.707107	0.500000	25.00000	85.0527	14003.43
4	0.546759	0.298946	14.94729	100.0000	8372.54

A 4-dimensional space is used to display most of the information about the differences between the rows representing the Table of relational values of the *Recidivism* and *AFC* variables. The first dimension represents more than 35% of the total inertia, the inclusion of the second dimension increases the «extracted» inertia to 60%, the inclusion of the third – to 85%, and the fourth – to 100%; *df* = 25, χ^2_{emp} = 56013.7. χ^2_{theor} (0.001;25) = 52.62. $\chi^2_{emp} > \chi^2_{theor}$.

Table 4

Eigenvalues and Inertia for all Dimensions for Variables *Recidivism* and *Education*

Number of Dims.	Total Inertia=2.5000				
	Singular Values	Eigen-Values	Perc. of Inertia	Cumulatv Percent	Chi Squares
1	0.751062	0.564094	22.56376	22.5638	14754.90
2	0.707107	0.500000	20.00000	42.5638	13078.40
3	0.707107	0.500000	20.00000	62.5638	13078.40
4	0.707107	0.500000	20.00000	82.5638	13078.40
5	0.660232	0.435906	17.43624	100.0000	11401.91

Qualitative representation of the deviations of the expected frequencies of the values of the variables *Recidivism* and *Education* is possible only in 5-dimensional space. The first dimension represents almost 23% of the total inertia, the inclusion of the second dimension increases the «extracted» inertia to 43%, the inclusion of the third – to 63%, the fourth – to 83%, the fifth – to 100%; *df* = 36, χ^2_{emp} = 65392. χ^2_{theor} (0.001;36) = 61.3. $\chi^2_{emp} > \chi^2_{theor}$.

Table 5

Eigenvalues and Inertia for all Dimensions for Variables *Recidivism* and *Employment*

Number of Dims.	Total Inertia=1.5000				
	Singular Values	Eigen-Values	Perc. of Inertia	Cumulatv Percent	Chi Squares
1	0.739183	0.546392	36.42612	36.4261	14288.82
2	0.707107	0.500000	33.33333	69.7594	13075.62
3	0.673504	0.453608	30.24055	100.0000	11862.41

A qualitative representation of the coordinates of the points of the pair of variables *Recidivism* and *Employment* requires three dimensions. The first dimension represents more than 36% of the total inertia, the inclusion of the second dimension increases the «extracted» inertia to 70%, the third to 100%; *df* = 9, χ^2_{emp} = 32233.8. χ^2_{theor} (0.001;9) = 27.88. $\chi^2_{emp} > \chi^2_{theor}$.

Table 6

Eigenvalues and Inertia for all Dimensions for Variables *Recidivism* and *Early Dismissals*

Number of Dims.	Total Inertia=1.0000				
	Singular Values	Eigen-Values	Perc. of Inertia	Cumulatv Percent	Chi Squares
1	0.864104	0.746675	74.66753	74.05210	24068.15
2	0.503314	0.253325	25.33247	100.0000	8165.61

Three dimensions are required to represent the row-dot table of the values of the variables *Recidivism* and *Early Dismissals*. The first dimension represents almost 75% of the total inertia, the inclusion of the second dimension increases the «extracted» inertia – up to 100%; *df* = 9, χ^2_{emp} = 32233.8. χ^2_{theor} (0.001;9) = 27.88. $\chi^2_{emp} > \chi^2_{theor}$.

Table 7

Eigenvalues and Inertia for all Dimensions for Variables *Recidivism* and *Suspended Convictions*

Number of Dims.	Total Inertia=1.0000				
	Singular Values	Eigen-Values	Perc. of Inertia	Cumulatv Percent	Chi Squares
1	0.764971	0.585181	58.51812	58.5181	15657.50
2	0.644064	0.414819	41.48188	100.0000	11099.17

Three dimensions are required for a qualitative representation of the initial values of the *Recidivism* and *Suspended Convictions* variables. The first dimension represents almost 25% of the total inertia, the inclusion of the second dimension increases the «extracted» inertia – up to 100%. *df* = 9, χ^2_{emp} = 26756.7. χ^2_{theor} (0.001;9) = 27.88. $\chi^2_{emp} > \chi^2_{theor}$.

Table 8

Eigenvalues and Inertia for all Dimensions for Variables *Recidivism* and *Motivation*

Number of Dims.	Total Inertia=1.0000				
	Singular Values	Eigen-Values	Perc. of Inertia	Cumulatv Percent	Chi Squares
1	0.733453	0.537953	53.79533	53.7953	14061.97
2	0.679740	0.462047	46.20467	100.0000	12077.78

Two dimensions are sufficient to represent the values of the *Recidivism* and *Motivation* variables. The first dimension represents almost 25% of the total inertia, the inclusion of the second dimension increases the «extracted» inertia – up to 100%. χ^2_{emp} = 26139.8. χ^2_{theor} (0.001;9) = 27.88. $\chi^2_{emp} > \chi^2_{theor}$.

Evaluation of the quality of the solution. Special statistics are used to assess the quality of the obtained solution. All or most of the points should be correctly represented – the distances between them should not be distorted as a result of applying the correspondence analysis procedure. In Tables 9–16, the results of statistics calculation are presented based on available line coordinates for all analyzed pairs of variables.

Table 9

Column Coordinates and Contributions to Inertia for Variables *Recidivism* and *Sex*

Row Name	Row Number	Coordin. Dim.1	Coordin. Dim.2	Mass	Quality	Relative Inertia
Sex:0	1	-2.67685	2.37027	0.03628	1.00000	0.46373
Sex:1	2	0.20940	-0.18541	0.46373	1.00000	0.03628
Recidivism:0	3	-0.73691	-0.65251	0.25397	1.00000	0.24604
Recidivism:1	4	0.76064	0.673524	0.24604	1.00000	0.25396

For the two-dimensional solution of the conducted correspondence analysis for the variables *Recidivism* and *Sex*, a high value was obtained for all groups of convicts: Quality is

equal to 1. The selected number of dimensions sufficiently represents all the rows and columns of the initial data table.

Table 10

Column Coordinates and Contributions to Inertia for Variables *Recidivism* and *AFA*

Row Name	Row Number	Coordin. Dim.1	Coordin. Dim.2	Coordin. Dim.3	Coordin. Dim.4	Mass	Quality	Relative Inertia
Recidivism:0	1	-0.82040	0.00000	0.00000	-0.54383	0.25396	1.00000	0.12302
Recidivism:1	2	0.84681	0.00000	0.00000	0.56134	0.24604	1.00000	0.12698
AFA:1	3	1.38065	0.27015	2.246372	-0.91522	0.05641	1.00000	0.22180
AFA:2	4	0.27792	0.01796	-0.693192	-0.18423	0.31407	1.00000	0.09297
AFA:3	5	-1.17138	-1.20037	0.735499	0.77650	0.10087	1.00000	0.19957
AFA:4	6	-1.64112	3.49744	0.586119	1.08788	0.02865	1.00000	0.23568

A high-quality representation of all groups of initial data was obtained for the variables *Recidivism* and *AFA*. The 4-dimensional space provides a qualitative representation of the initial data point cloud.

Table 11

Column Coordinates and Contributions to Inertia for Variables *Recidivism* and *AFC*

Row Name	Row Number	Coordin. Dim.1	Coordin. Dim.2	Coordin. Dim.3	Coordin. Dim.4	Mass	Quality	Relative Inertia
Recidivism:0	1	-0.82472	0.00000	0.00000	-0.53855	0.25379	1.00000	0.12311
Recidivism:1	2	0.85005	0.00000	0.00000	0.55510	0.24622	1.00000	0.12690
AFC:1	3	1.12389	-1.12646	1.12980	-0.73391	0.09351	1.00000	0.20325
AFC:2	4	0.17200	0.63217	-0.49888	-0.11232	0.29573	1.00000	0.10213
AFC:3	5	-1.30216	-1.48622	-0.51923	0.85033	0.08479	1.00000	0.20761
AFC:4	6	-1.75393	1.70922	3.30782	1.14533	0.02597	1.00000	0.23710

For the variables *Recidivism* and *AFS*, the use of 4 dimensions is the optimal solution for a qualitative representation of the coordinates of the points representing the input data.

Table 12

Column Coordinates and Contributions to Inertia for Variables *Recidivism* and *Education*

Row Name	Coordin. Dim.1	Coordin. Dim.2	Coordin. Dim.3	Coordin. Dim.4	Coordin. Dim.5	Mass	Quality	Relative Inertia
Recidivism:0	0.73942	0.00000	0.00000	0.00000	-0.64100	0.25391	1.00000	0.09844
Recidivism:1	-0.76289	0.00000	0.00000	0.00000	0.67063	0.24610	1.00000	0.10156
Education:0	-0.56435	-0.47618	1.59036	1.42582	-0.49607	0.07870	1.00000	0.16852
Education:1	-0.37918	-0.20106	-1.02309	0.10660	-0.33333	0.21246	1.00000	0.11502
Education:2	0.24151	0.97005	0.54331	-0.81393	0.21231	0.16655	1.00000	0.13338
Education:3	0.89072	-3.98202	0.92121	-2.43375	0.78300	0.01997	1.00000	0.19201
Education:4	2.99999	-0.08253	-0.74721	2.20951	2.63718	0.02232	1.00000	0.19107

A 5-dimensional space was used for the qualitative representation of the coordinates of the initial data points of the variables *Recidivism* and *Education*, which provides the necessary representation of the data in the initial table.

Table 13

Column Coordinates and Contributions to Inertia for Variables *Recidivism* and *Employment*

Row Name	Row Number	Coordin. Dim.1	Coordin. Dim.2	Coordin. Dim.3	Mass	Quality	Relative Inertia
Recidivism:0	1	-0.72762	0.00000	-0.66297	0.25394	1.00000	0.16404
Recidivism:1	2	0.75093	0.00000	0.68421	0.24606	1.00000	0.16930
Employment:0	3	0.55413	1.23211	-0.50489	0.16233	1.00000	0.22511
Employment:1	4	0.28343	-1.04491	-0.25824	0.22333	1.00000	0.18445
Employment:2	5	-1.34034	0.29164	1.22125	0.12434	1.00000	0.25711

For the three-dimensional solution for the variables Recidivism and Employment, the quality for all groups of convicts is high. 3 dimensions reasonably represent all the rows and columns of the data table.

Table 14

Column Coordinates and Contributions to Inertia for Variables *Recidivism* and *Early Dismissals*

Row Name	Row Number	Coordin. Dim.1	Coordin. Dim.2	Mass	Quality	Relative Inertia
Early Dismissals:0	1	-0.55909	0.32565	0.35245	1.00000	0.14755
Early Dismissals:1	2	1.33553	-0.77790	0.14755	1.00000	0.35245
Recidivism:0	3	-0.85127	-0.49584	0.25374	1.00000	0.24626
Recidivism:1	4	0.87714	0.51090	0.24626	1.00000	0.25374

For the two-dimensional solution for the *Recidivism* and *Early Dismissals* variables, the quality of the representation of the initial data points is high. 2 dimensions are the optimal solution for this case.

Table 15

Column Coordinates and Contributions to Inertia for Variables *Recidivism* and *Suspended Conviction*

Row Name	Row Number	Coordin. Dim.1	Coordin. Dim.2	Mass	Quality	Relative Inertia
Recidivism:0	1	-0.75300	-0.63399	0.25394	1.00000	0.24606
Recidivism:1	2	0.77713	0.65430	0.24606	1.00000	0.25394
Suspended Cinviction:0	3	-0.66146	0.55691	0.28609	1.00000	0.21391
Suspended Cinviction:1	4	0.88469	-0.74485	0.21391	1.00000	0.28609

Calculated statistics on the coordinates of the rows and columns of the two-dimensional solution for the variables *Recidivism* and *Suspended Conviction* testify to the high quality of the representation of points in 2-D space.

Table 16

Column Coordinates and Contributions to Inertia for Variables *Recidivism* and *Motivation*

Row Name	Row Number	Coordin. Dim.1	Coordin. Dim.2	Mass	Quality	Relative Inertia
Recidivism:0	1	0.72186	-0.66900	0.25399	1.00000	0.24602
Recidivism:1	2	-0.74523	0.69066	0.24602	1.00000	0.25398
Motivation:0	3	-1.94535	-1.80120	0.62332	1.00000	0.43767
Motivation:1	4	0.27679	0.74485	0.43767	1.00000	0.06233

A two-dimensional solution is optimal for representing data points for the *Recidivism* and *Motivation* variables. The quality of the representation of the coordinates of the points of the initial data table in 2-dimensional space is high.

A graphical analysis of the relationship between the propensity of convicts to commit criminal recidivism (prone/non-prone) and the levels of the variables used in this analysis was conducted. The analysis of 2-D graphs of row and column coordinates for the corresponding pairs of variables gives grounds for drawing the following conclusions:

- there is a direct relationship between the fact of committing repeated criminal offenses (*Recidivism* = 1) and conviction in adolescence or young age to a real/conditional punishment (*AFA* = 1, *AFA* = 2, *AFC* = 1, *AFC* = 2) (Fig. 1, 2);
- there is a direct relationship between the convictions of criminal recidivism and the level of education they have acquired: persons who commit repeated criminal offenses mainly have incomplete secondary or secondary education (*Education* = 0, *Education* = 1) (Fig. 3).
- unemployed convicts (*Employment* = 0) and persons with partial employment (*Employment* = 1) are more prone to criminal recidivism (Fig. 4);
- there is a direct relationship between the commission of repeated criminal offenses by convicts with the presence of previous early dismissals (*Early Dismissals* = 1) and suspended convictions (*Suspended Conviction* = 1) (Fig. 5, 6);
- there is no relationship between convicts committing criminal recidivism and their motivation for release, as well as with the person's gender (Figs. 7, 8).

These findings confirmed the results obtained in our previous studies [19, 20].

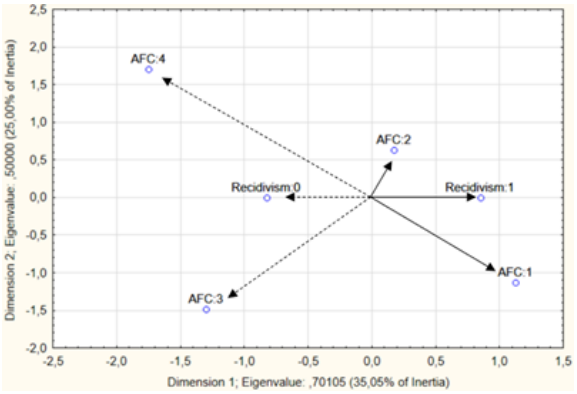

Figure 1. 3D Plot of Row and Column Coordinates for Variables Recidivism and AFA

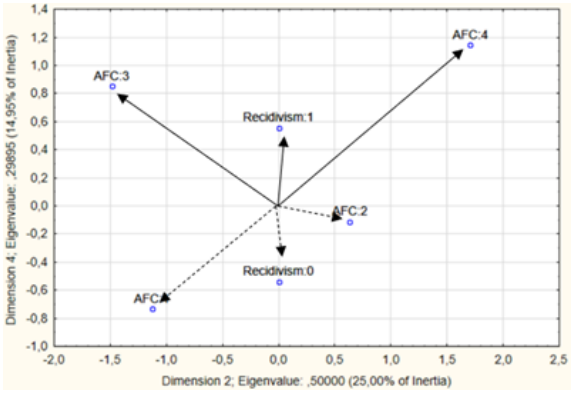

Figure 2. 3D Plot of Row and Column Coordinates for Variables Recidivism and AFC

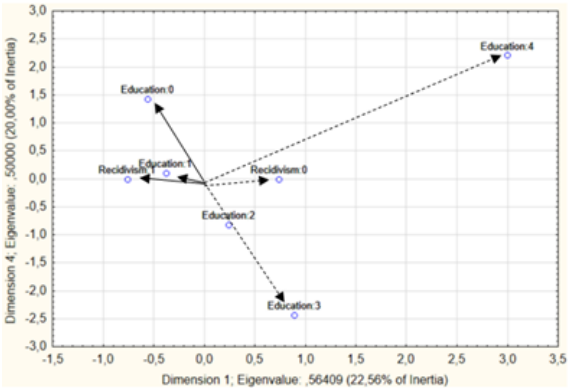

Figure 3. 3D Plot of Row and Column Coordinates for Variables Recidivism and Education

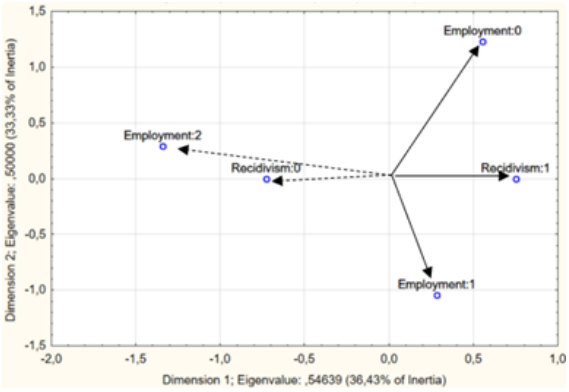

Figure 4. 3D Plot of Row and Column Coordinates for Variables Recidivism and Employment

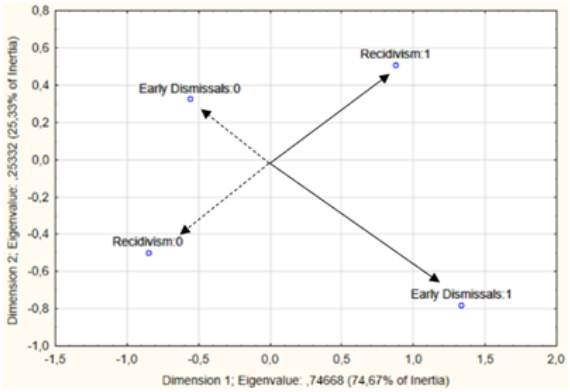

Figure 5. 2-D Plot of Row and Column Coordinates for Variables Recidivism and Early Dismissals

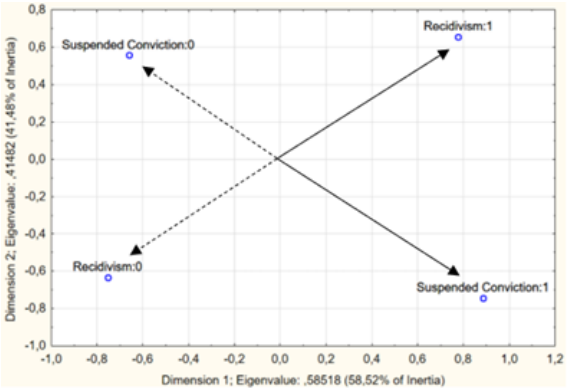

Figure 6. 2-D Plot of Row and Column Coordinates for Variables Recidivism and Suspended Conviction

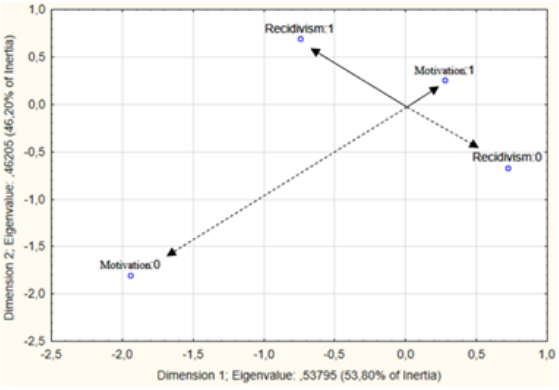

Figure 7. 2-D Plot of Row and Column Coordinates for Variables Recidivism and Motivation

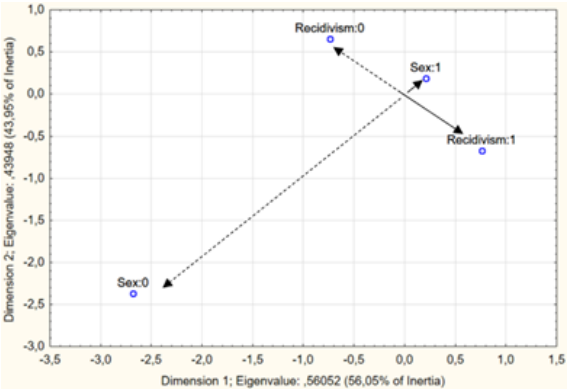

Figure 8. 2-D Plot of Row and Column Coordinates for Variables Recidivism and Sex

Conclusions. Correspondence analysis was applied to identify non-obvious relationships between the convictions of criminal recidivism and their following socio-demographic characteristics and elements of criminal history: gender, age at the time of the first conviction to the actual penalty, age at the time of the first conviction to the actual or conditional penalty, level of education, type of employment, early dismissals, suspended sentences, motivation for release. It is proved that persons sentenced to a real/suspended

sentence in adolescence or young adulthood, have incomplete secondary or secondary education, are unemployed or partially employed, have previous early releases and/or suspended sentences have a greater risk of committing repeated criminal offenses. It is cleared that there is no relationship between criminal recidivism and a person's gender and motivation to be released. The obtained results can provide important information to law enforcement agencies for the development of effective strategies to prevent and solve criminal offences. The continuation of our research will be the development of a discrimination model based on criminal histories of criminals to identify convicts with a high risk of recidivism.

References

1. Volkov S., Prokopenko A., Asabashvili S., Volkov K. Some aspects of autonomous cyber-physical systems diagnostics by their qualitative state. Scientific Journal of TNTU. 2022. Vol. 108. No. 4. P. 122–130. https://doi.org/10.33108/visnyk_tntu2022.04.122
2. Kulyna S. Evaluation of the reverse transformation methods complexity of the residual number system for secure data storage. Scientific Journal of TNTU. 2022. Vol. 107. No. 3. P. 21–28. https://doi.org/10.33108/visnyk_tntu2022.03.021
3. Hladiy G., Khoma N., Zakaliak R., Mohylska M. Website dependability evaluation model based on a multi-criteria approach. Scientific Journal of TNTU. 2022. Vol. 107. No. 3. P. 105–114. https://doi.org/10.33108/visnyk_tntu2022.03.105
4. Stadnyk M., Palamar A. Project management features in the cybersecurity area. Scientific Journal of TNTU. 2022. Vol. 106. No. 2. P. 54–62. https://doi.org/10.33108/visnyk_tntu2022.02.054
5. Prus R., Yatsyuk S., Hlynchuk L., Mulyar V. Economic aspects of information protection under present large-scale cyber-attacks conditions. Scientific Journal of TNTU. 2022. Vol. 106. No. 2. P. 63–74. https://doi.org/10.33108/visnyk_tntu2022.02.063
6. Kovalchuk O., Kasianchuk M., Karpinski M., Shevchuk R. Decision-Making Supporting Models Concerning the Internal Security of the State. INTL Journal of Electronics Telecommunications. 2023. Vol. 96. No. 2. P. 301–307. URL: https://doi.org/10.24425/ijet.2023.144365.
7. Yu R., Langstrom N., Forsman M., Sjolander A., Fazel S., Molero Y. Associations between prisons and recidivism: A nationwide longitudinal study. National Center for Biotechnology Information. PLoS ONE, 2022, 17, e0267941. https://doi.org/10.1371/journal.pone.0267941
8. Berezka K. M., Kovalchuk O. Ya., Banakh S. V., Zlyvko S. V., Hrechaniuk R. A Binary Logistic Regression Model for Support Decision Making in Criminal Justice. Folia Oeconomica Stetinensia. 2022. Vol. 22. No. 1. P. 1–17. https://doi.org/10.2478/foli-2022-0001
9. Yukhnenko D., Blackwood N., Fazel S. Risk factors for recidivism in individuals receiving community sentences: A systematic review and meta-analysis. CNS Spectrums. 2020. Vol. 25. No. 2. P. 252–263. https://doi.org/10.1017/S1092852919001056
10. Associative Rule Mining for the Assessment of the Risk of Recidivism, 4th International Workshop “Intelligent Information Technologies & Systems of Information Security”. Khmelnytskyi, Ukraine, 2023, 3373, pp. 376–387. https://ceur-ws.org/Vol-3373/paper24.pdf.
11. Jacobs L. A., Fixler A., Labrum T., Givens A., Newhill C. Risk Factors for Criminal Recidivism Among Persons With Serious Psychiatric Diagnoses: Disentangling What Matters for Whom. Front Psychiatry. 2021. Vol. 12. 778399. https://doi.org/10.3389/fpsyt.2021.778399
12. Cuevas C., Wolff K. T., Baglivio M. T. Dynamic risk factors and timing of recidivism for youth in residential placement. Journal of Criminal Justice. 2019. Vol. 60. P. 154–166. https://doi.org/10.1016/j.jcrimjus.2018.10.003
13. Garritsen K., Jankovic M., Masthoff E., Caluwé E.D., Bogaerts S. The Role of Dynamic Risk and Protective Factors in Predicting Violent Recidivism: Intellectual Ability as a Possible Moderator? International Journal of Offender Therapy and Comparative Criminology, 2022, 52271. https://doi.org/10.21428/cb6ab371.72aac051
14. Navarro-Pérez J.-J., Viera M., Calero J., Tomas J. M. Factors in Assessing Recidivism Risk in Young Offenders. Sustainability. 2020. Vol. 12. No. 3. 1111. https://doi.org/10.3390/su12031111
15. Heffernan R., Ward T. Dynamic Risk Factors, Protective Factors and Value-Laden Practices. Psychiatry, Psychology and Law. 2019. Vol. 26. No. 2. https://doi.org/10.1080/13218719.2018.1506721
16. Saravanan P., Selvaprabu J., Raj L. A., Khan A., Sathick K. Survey on crime analysis and prediction using data mining and machine learning techniques. Lect. Notes Electr. Eng. 2021. Vol. 688. P. 435–448. https://doi.org/10.1007/978-981-15-7241-8_31

17. Riani M., Atkinson A. C., Torti F., Corbellini A. Robust Correspondence Analysis. Journal of the Royal Statistical Society Series C: Applied Statistics. 2022. Vol. 71. No. 5. P. 1381–1401. https://doi.org/10.1111/rssc.12580
18. Unified register of pre-trial investigations. URL: https://erdr.gp.gov.ua. (accessed: 13.03.2023) [In Ukrainian].
19. Kovalchuk O., Karpinski M., Banakh S., Kasianchuk M., Shevchuk R., Zagorodna N. Prediction Machine Learning Models on Propensity Convicts to Criminal Recidivism. Information. 2023. Vol. 14. No. 3. P. 161. https://doi.org/10.3390/info14030161
20. Kovalchuk O. Modeling the risks of the confession process of the accused of criminal offenses based on survival concept. Scientific Journal of TNTU. 2022. Vol. 108. No. 4. P. 27–37. https://doi.org/10.33108/visnyk_tntu2022.04.027

Список використаної літератури

1. Volkov S., Prokopenko A., Asabashvili S., Volkov K. Some aspects of autonomous cyber-physical systems diagnostics by their qualitative state. Scientific Journal of TNTU. 2022. Vol. 108. No. 4. P. 122–130. https://doi.org/10.33108/visnyk_tntu2022.04.122
2. Kulyna S. Evaluation of the reverse transformation methods complexity of the residual number system for secure data storage. Scientific Journal of TNTU. 2022. Vol. 107. No. 3. P. 21–28. https://doi.org/10.33108/visnyk_tntu2022.03.021
3. Hladiy G., Khoma N., Zakaliak R., Mohylska M. Website dependability evaluation model based on a multi-criteria approach. Scientific Journal of TNTU. 2022. No. 107 (3). P. 105–114. https://doi.org/10.33108/visnyk_tntu2022.03.105
4. Stadnyk M., Palamar A. Project management features in the cybersecurity area. Scientific Journal of TNTU. 2022. No. 106 (2). P. 54–62. https://doi.org/10.33108/visnyk_tntu2022.02.054
5. Prus R., Yatsyuk S., Hlynchuk L., Mulyar V. Economic aspects of information protection under present large-scale cyber-attacks conditions. Scientific Journal of TNTU. 2022. No. 106 (2). P. 63–74. https://doi.org/10.33108/visnyk_tntu2022.02.063
6. Kovalchuk O., Kasianchuk M., Karpinski M., Shevchuk R. Decision-Making Supporting Models Concerning the Internal Security of the State. INTL Journal of Electronics Telecommunications. 2023. No. 96 (2). P. 301–307.
7. Yu R., Langstrom N., Forsman M., Sjolander A., Fazel S., Molero Y. Associations between prisons and recidivism: A nationwide longitudinal study. National Center for Biotechnology Information. PLoS ONE, 2022, 17, e0267941. https://doi.org/10.1371/journal.pone.0267941
8. Berezka K. M., Kovalchuk O. Ya., Banakh S. V., Zlyvko S. V., Hrechaniuk R. A Binary Logistic Regression Model for Support Decision Making in Criminal Justice. Folia Oeconomica Stetinensia. 2022. Vol. 22. No. 1. P. 1–17. https://doi.org/10.2478/foli-2022-0001
9. Yukhnenko D., Blackwood N., Fazel S. Risk factors for recidivism in individuals receiving community sentences: A systematic review and meta-analysis. CNS Spectrums. 2020. Vol. 25. No. 2. P. 252–263. https://doi.org/10.1017/S1092852919001056
10. Associative Rule Mining for the Assessment of the Risk of Recidivism, 4th International Workshop "Intelligent Information Technologies & Systems of Information Security". Khmelnytskyi, Ukraine, 2023, 3373, pp. 376–387. https://ceur-ws.org/Vol-3373/paper24.pdf.
11. Jacobs L. A., Fixler A., Labrum T., Givens A., Newhill C. Risk Factors for Criminal Recidivism Among Persons With Serious Psychiatric Diagnoses: Disentangling What Matters for Whom. Front Psychiatry. 2021. Vol. 12. 778399. https://doi.org/10.3389/fpsyt.2021.778399
12. Cuevas C., Wolff K. T., Baglivio M. T. Dynamic risk factors and timing of recidivism for youth in residential placement. Journal of Criminal Justice. 2019. Vol. 60. P. 154–166. https://doi.org/10.1016/j.jcrimjus.2018.10.003
13. Garritsen K., Jankovic M., Masthoff E., Caluwé E.D., Bogaerts S. The Role of Dynamic Risk and Protective Factors in Predicting Violent Recidivism: Intellectual Ability as a Possible Moderator? International Journal of Offender Therapy and Comparative Criminology, 2022, 52271. https://doi.org/10.21428/cb6ab371.72aac051
14. Navarro-Pérez J.-J., Viera M., Calero J., Tomas J. M. Factors in Assessing Recidivism Risk in Young Offenders. Sustainability. 2020. Vol. 12. No. 3. 1111. https://doi.org/10.3390/su12031111
15. Heffernan R., Ward T. Dynamic Risk Factors, Protective Factors and Value-Laden Practices. Psychiatry, Psychology and Law. 2019. Vol. 26. No. 2. https://doi.org/10.1080/13218719.2018.1506721
16. Saravanan P., Selvaprabu J., Raj L. A., Khan A., Sathick K. Survey on crime analysis and prediction using data mining and machine learning techniques. Lect. Notes Electr. Eng. 2021. Vol. 688. P. 435–448. https://doi.org/10.1007/978-981-15-7241-8_31

17. Riani M., Atkinson A. C., Torti F., Corbellini A. Robust Correspondence Analysis. Journal of the Royal Statistical Society Series C: Applied Statistics. 2022. Vol. 71. No. 5. P. 1381–1401. https://doi.org/10.1111/rssc.12580
18. Єдиний реєстр досудових розслідувань. [Електронний ресурс]. URL: https://erdr.gp.gov.ua.
19. Kovalchuk O., Karpinski M., Banakh S., Kasianchuk M., Shevchuk R., Zagorodna N. Prediction Machine Learning Models on Propensity Convicts to Criminal Recidivism. Information. 2023. Vol. 14. No. 3. P. 161. https://doi.org/10.3390/info14030161
20. Kovalchuk O. Modeling the risks of the confession process of the accused of criminal offenses based on survival concept. Scientific Journal of TNTU. 2022. Vol. 108. No. 4. P. 27–37. https://doi.org/10.33108/visnyk_tntu2022.04.027

УДК 57-1

АНАЛІЗ ВІДПОВІДНОСТІ ДЛЯ ВИЯВЛЕННЯ ФАКТОРІВ РИЗИКУ КРИМІНАЛЬНИХ РЕЦИДИВІВ

Ольга Ковальчук

Західноукраїнський національний університет, Тернопіль, Україна

Резюме. *Використано аналіз відповідності для виявлення взаємозв'язків між кримінальними рецидивами й наступними елементами попередніх кримінальних історій засуджених: стать, вік на момент першого засудження до реальної міри покарання, вік на момент першого засудження до реальної або умовної міри покарання, рівень освіти, тип зайнятості, наявність дострокових звільнень, наявність умовних засуджень, наявність мотивації до звільнення.*

Проведений емпіричний аналіз дає можливість зробити висновки про існування прямого зв'язку ризику скоєння кримінальних рецидивів з віком, в якому обвинуваченому було призначено перше покарання (як до умовної, так і до реальної міри). Особи, які вперше отримали вирок у підлітковому чи молодому віці, мають більший ризик скоєння повторних кримінальних злочинів. Засуджені, які були вперше засуджені у віці старше 30 років до реальної міри покарання чи до реальної або умовної міри покарання є менш схильними до скоєння кримінальних рецидивів. Встановлено, що особи з високим ризиком скоєння повторних кримінальних злочинів мають незакінчену середню або середню освіту. Особи з середньою спеціальною, незакінченою вищою та вищою освітою є менш схильними до скоєння повторних кримінальних злочинів. Безробітні засуджені та особи з частковою зайнятістю схильні до вчинення кримінальних рецидивів на відміну від осіб з повною зайнятістю. Доведено, що ризик скоєння кримінальних рецидивів є вищим у осіб, що мали попередні дострокові звільнення чи/та умовні засудження. Підтверджено відсутність взаємозв'язку між схильністю до скоєння засудженими кримінальних рецидивів у майбутньому з наявністю у них мотивації до звільнення та статтю особи. Отримані результати можуть надати вагому інформацію для розроблення ефективних стратегій профілактики, запобігання та розкриття кримінальних злочинів.

Ключові слова: *аналіз відповідності, взаємозв'язки, критерій* χ^2*, внутрішня безпека, кримінальний рецидивізм.*

https://doi.org/10.33108/visnyk_tntu2023.03.035 *Отримано 10.07.2023*