

---

УДК 343.973+ 519.2

*П. П. Сердюк*

### **ЛОЖНЫЕ КОРРЕЛЯЦИИ В ИССЛЕДОВАНИИ НАСИЛЬСТВЕННОЙ ПРЕСТУПНОСТИ**

**Постановка проблемы.** Примеров использования корреляционного анализа в установлении коррелятов в связи с преступностью много. Однако самой раздражающей корреляцией, которая была якобы доказана в криминологическом исследовании, остаётся корреляция между динамикой уровня насильственной преступности и динамикой неравенства в распределении доходов населения. О проблеме неравномерного распределения ресурсов среди населения как криминогенном факторе говорили чуть ли не все криминологи в мире, поэтому совершенно нет никакой необходимости в том, чтобы приводить здесь их список, но эти утверждения не получили надлежащей эмпирической проверки, поскольку их результаты тяжеловерно противоречивы.

Посему существовала необходимость проверки изысканий по этому вопросу. Но как-то у криминологов до этого не доходили руки. Нужно было, всего-навсего, проверить математические модели и таблицы линейной регрессии с произвольно выбранным явлением. Например, каким-либо явлением, связь с которым выглядела бы вздорной. Вот и было решено в качестве такового избрать показатели производства молока за релевантный период для того, чтобы исключить возможность ложной корреляции при таком же подходе.

**Анализ последних исследований и публикаций.** Наиболее последовательными сторонниками использования регрессионного и корреляционного анализа по зависимой, коей является насильственная преступность, и независимой, то есть неравенство в распределении доходов населения (индекс Джини), переменной были и остаются представители школы С.Г. Олькова. Нужно отметить, что этот криминолог сделал очень много для криминологической науки, поскольку всемерно способствует тому, чтобы вытянуть её из болота умозрительности. Его работу, а также труды его учеников трудно переоценить, хотя они зачастую и не приводили к точным, непротиворечивым результатам, но их подход, который рвётся из уз умозрительности в криминологию, сделал очень многое.

Ещё до публикаций по этому вопросу С.Г. Олькова его ученица Э.Г. Юзиханова в 2002 г. защитила кандидатскую диссертацию, в которой произвела расчёты и установила корреляцию между динамикой преступности и индексом Джини [1, с. 129–131]. Не вызывает сомнений, что делалось это с подачи С.Г. Олькова, который в 2003 г. опубликовал тезисы доклада, посвящённого влиянию социального неравенства населения на преступность, а после, в 2004 г., и 2011 г. предложил нашему вниманию две обстоятельные статьи, в которых на основе корреляционного анализа пришёл к утверждению, что существует правило, в соответствии с которым в странах с более высокой степенью дифференциации народонаселения по уровню доходов существует более высокий уровень насилия, в том числе экстремального, выражающегося в умышленных убийствах [2, с. 73–78; 3, с. 240–252].

Идеи С.Г. Олькова были развиты его учеником И.С. Скифским в диссертации, защищённой в 2006 г., в которой были взяты временные рамки в 25-ть лет, с 1980 по 2004 гг. Однако исследователь сам порядок подсчёта не показал, ни в тексте, ни в приложениях к диссертации, поэтому можно было взять только данные, которыми он пользовался, и увидеть общий результат. Это самое масштабное исследование вопроса о влиянии социально-экономического неравенства на насильственную преступность, посему оно представляет самый острый интерес для проверки тех результатов, к которым пришёл указанный исследователь [4]. Впрочем, рамки этой статьи не позволяют мне проверить все результаты, и было принято решение остановиться на проверке общей гипотезы о связи между динамикой уровня насильственной преступности и динамикой неравенства в распределении доходов населения.

Результаты изысканий С.Г. Олькова и Э.Г. Юзихановой необходимо проверять отдельно, поскольку они основаны на более узкой эмпирической базе. Работа же И.С. Скифского особо заслуживает внимания из-за своего масштаба и непосредственной близости к интересующему нас вопросу. Нужно сказать, что публикации его учителя – С.Г. Олькова, хотя и имеют формат научных статей, всё же сносно описаны, а вот диссертация И.С. Скифского лишь описывает применяемые формулы и результат. Сразу даже непонятно, какие показатели состояния насильственной преступности автор использовал, абсолютные или относительные? И лишь ознакомление с работами С.Г. Олькова и конкретные перерасчёты говорят о том, что И.С. Скифский использовал относительные показатели насильственной преступности, исходя из индекса на 100 000 населения.

Исследователь использовал уравнение парной регрессии:

$$y = a + bx,$$

где  $\tilde{y}$  – среднее значение результативного признака  $y$  при определённом значении факторного признака  $x$ , другими словами – среднее значение насильственной преступности при определённом значении индекса Джини;  $a$  – свободный член уравнения;  $b$  – коэффициент регрессии, измеряющий среднее отношение отклонения результативного признака от

его средней величины к отклонению факторного признака от его средней величины на одну единицу его измерения,  $\sigma_y$  вариация  $y$ , приходящаяся на единицу вариации  $x$  [4, с. 90].

В своём выступлении на конференции, состоявшейся 15 мая 2013 г. в Балтийском институте экологии, политики и права (РФ), Я.И. Гишинский особо отметил роль исследований указанных авторов по этому вопросу. Он остановился на мысли о том, что первоначально результаты И.С. Скифского вызвали сомнения в показателе уровня корреляции (0,92), близкому к функциональной связи, то есть (1,0), но, по его словам, исследователь развеял сомнения на отдельной встрече, в которой он пояснил свои результаты. Примечательно то, что поставили под сомнение результат, а не сам подход, которым воспользовался исследователь [5].

Дело в том, что существует риск так называемых «ложных корреляций», наподобие тех, которые были установлены в США. Например, исследователи нашли корреляционную связь между количеством смертей в бассейнах и количеством ролей Николаса Кейджа в кино, возрастом победительниц конкурса «Мисс Америка» и количеством смертей из-за горячего пара, уровнем производства пчелиного мёда в США и уровнем разводов в Южной Каролине. Не нужно забывать, что математика – это не волшебство. Когда сравниваются числовые параметры, то может стать так, что корреляция существует только между числами.

Очевидно, необходимо помнить, что метод корреляционного анализа довольно часто даёт противоречивые, и даже смешные результаты. Это должно заставить искать другие способы определения влияния факторов на преступность. Либо же необходимы усовершенствования в означенном методе.

Стало быть, нужно быть осторожным с математикой, и важно помнить, что определённые методы математики могут подтверждать ложные связи, которых в действительности нет, и они вряд ли могли быть в такой последовательности.

**Цель статьи** (задачи) состоит в том, чтобы проверить гипотезу о существовании корреляции между динамикой уровня насильственной преступности и динамикой неравенства в распределении доходов населения.

**Изложение основного материала исследования.** По выводам И.С. Скифского, который в построенной линейной регрессии, несомненно, использовал относительные показатели насильственной преступности, исходя из показателя на 100 000 населения, изменение коэффициента Джини на 0,1 или на 1% должно изменять показатель насильственной преступности на 48 300 преступлений в России, то есть на абсолютный показатель, а не относительный. При этом исследователь отдельно не переводил на базу сравнения в 100 000 населения полученный результат  $y = 482,57x - 52,855$ . Как раз вот эта цифра 482,57 $x$  и есть эти 48 300 насильственных преступлений, что означает абсолютный показатель, а не относительный [4, с. 97]. Но так не может быть, ведь относительные данные на входе должны давать относительные показатели на выходе. Подобной ошибки не

допускал его учитель С.Г. Ольков, поскольку он использовал относительные показатели насильственной преступности на 100 000 населения, и его выводы говорили об изменении индекса или коэффициента умышленных убийств на 100 000 населения. Построенное им регрессионное уравнение по данным за 2000–2008 гг., которое описывало связь между числом умышленных убийств и индексом Джини, имело вид:  $\sigma = -19,71 + 0,6865x$ , где  $\sigma$  – оценка числа умышленных убийств, приходящегося на 100 000 народонаселения,  $x$  – индекс Джини (в процентах). Отсюда он сделал вывод, что коэффициент умышленных убийств, приведённый на 100 000 народонаселения, изменяется на 0,6865 преступлений при изменении индекса Джини на единицу измерения (1%). Отрицательное значение свободного члена (сдвига) в уравнении исследователь объяснял тем, что при низких значениях индекса Джини, близких к нулю, совершение умышленных убийств практически невероятно (Ольков, 2011). Это предположение, конечно же, притянута, и его можно объяснить нежеланием исследователя усомниться в корректности использования этого исследовательского метода.

Проверим якобы доказанную гипотезу, представленную в диссертации И.С. Скифского, о том, что изменение индекса Джини или неравенства в распределении доходов населения коррелирует с изменениями уровня насильственной преступности. При этом использованы те же исходные данные, что и в его публикации. Используются те же эконометрические методы (табл. 1).

Если индекс Джини в таблице представить в процентах, например, как средний показатель не 0,323, а 32,3%, то тогда создаётся отрицательное значение  $b$ , что говорило бы об обратной корреляции. А в случае проделывания того же во второй таблице  $b = 3,6$ , что означает больше, чем функциональная связь, что невозможно. Это должна быть цифра, не больше 1.

Таблица 1

год	x	y	xу	x <sup>2</sup>	y <sup>2</sup>
1	2	3	4	5	6
1980	0,290	50	14,5	0,0841	2500
1981	0,207	48,7	10,08	0,0428	2371,69
1982	0,210	47,6	9,99	0,0441	2265,76
1983	0,260	50,9	13,23	0,0676	2590,81
1984	0,240	49	11,7	0,0576	2401
1985	0,284	43	12,21	0,0807	1849
1986	0,221	33,9	7,49	0,0488	1149,21
1987	0,240	31,6	7,58	0,0576	998,56
1988	0,264	38,8	10,24	0,0697	1505,44
1989	0,275	54,1	14,87	0,0756	2926,81
1990	0,281	59,6	16,74	0,0790	3552,16
1991	0,260	60,7	15,78	0,0676	3684,49

Закінчення таблиці 1

1	2	3	4	5	6
1992	0,289	81,5	23,55	0,0835	6642,25
1993	0,407	101,6	41,35	0,1657	10322,56
1994	0,409	102,5	41,92	0,1673	10506,25
1995	0,387	97,1	37,57	0,1498	9428,41
1996	0,387	86,9	33,63	0,1498	7551,61
1997	0,390	80,9	31,55	0,1521	6544,81
1998	0,394	83,3	32,82	0,1552	6938,89
1999	0,400	87,7	35,08	0,1600	7691,29
2000	0,395	88,6	34,99	0,1560	7849,96
2001	0,397	98,3	39,02	0,1576	9662,89
2002	0,397	101,4	40,25	0,1576	10281,96
2003	0,403	100,4	40,46	0,1624	10080,16
2004	0,409	106,2	43,43	0,1673	11278,44
Итого	8,096	1784,3	620,17	65,545	3183726
Среднее значение	0,323	71,372	24,806	2,6218	127349,1
$\sigma$	2,517				
$\sigma^2$	122255,1				

Среднее значение определяется по формулам:

$$\begin{aligned} \bar{x} &= \frac{\sum x_i}{n} = \frac{8,096}{25} = 0,323, \\ \bar{y} &= \frac{\sum y_i}{n} = \frac{1784,3}{25} = 71,372, \\ \overline{xy} &= \frac{\sum x_i y_i}{n} = \frac{620,17}{25} = 24,806. \end{aligned}$$

По сути, среднее значение это сумма показателя  $x$ ,  $y$ ,  $xy$ , делённая на 25, то есть на количество измерений. В этом случае, это количество лет, за которое проходило обобщение с 1980 по 2004 гг.

Для вычисления коэффициента корреляции необходимо вычислить среднеквадратическое отклонение  $\sigma(x, y)$ , которое можно вычислить, только выявив дисперсию показателей  $x$  и  $y$ , то есть самих показателей неравенства в распределении доходов населения ( $x$ ) и уровня насильственной преступности ( $y$ ). Итак, сначала вычисляем дисперсию  $x$  и  $y$ :

$$\begin{aligned} D_{(x)} &= \frac{\sum x_i^2}{n} - \bar{x}^2 = \frac{65,545}{25} - 0,323^2 = 2,517 \\ D_{(y)} &= \frac{\sum y_i^2}{n} - \bar{y}^2 = \frac{3183726}{25} - 71,372^2 = 122255,1 \end{aligned}$$

Теперь дело стало за вычислением среднеквадратического отклонения  $x$  и  $y$ :

$$\sigma_{(x)} = \sqrt{D_{(x)}} = \sqrt{2,517} = 1,586,$$

$$\sigma_{(y)} = \sqrt{D_{(y)}} = \sqrt{122255,1} = 349,7.$$

Следующим шагом будет определение коэффициента корреляции между  $x$  и  $y$ , что, в сущности, очень просто, поскольку только и надо, что подставить уже вычисленные выше результаты:

$$r_{xy} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sigma_{(x)} \sigma_{(y)}} = \frac{24,806 - 0,323 \cdot 71,372}{1,586 \cdot 349,7} = 0,003.$$

Связь между признаком  $y$  и фактором  $x$  настолько слаба, что её не стоит принимать во внимание. Выходит такое уравнение регрессии:

$$y_x = r_{xy} \frac{x - \bar{x}}{\sigma_{(x)}} \sigma_{(y)} + \bar{y} = 0,003 \frac{x - 0,323}{1,586} 349,7 + 71,372 = 347,79x + 71.$$

Коэффициент регрессии:  $k = a = 347,79$

Коэффициент детерминации:  $R^2 = 0,71^2 = 0,50$ , то есть в 50% случаев изменения  $x$  приводят к изменению  $y$ . Другими словами, точность подбора уравнения регрессии умеренная. Остаточная дисперсия: 50%.

Полученный результат отличается от того, который получил И.С. Скифский, на 134 единицы ( $483 - 348 = 134$ ) меньше, а коэффициент детерминации тоже заметно отличается.

Параметр регрессии позволяет сделать вывод, что увеличение коэффициента Джини на 0,1 или на 1% может увеличить насильственную преступность не на 48 300 преступлений, а на 34 преступления на 100 000 населения, хотя индекс относительного учёта преступлений на 100 000 населения, как искусственная аналитическая единица не пригодна для сравнения.

Нужно проверить значимость коэффициента корреляции. Для этого требуется проверить гипотезу, не равна ли  $r_{xy}$  нулю. Если да, то корреляция статистически не значима. Для того чтобы проверить гипотезу о равенстве нулю генерального коэффициента корреляции нормальной двумерной случайной величины, надо вычислить наблюдаемое значение критерия (величина случайной ошибки):

$$T_{набл} = r_{xy} \frac{\sqrt{n-2}}{1-r_{xy}^2} = 0,003 \frac{\sqrt{10}}{\sqrt{1-0,00009}} = 3,16$$

Вычисляем степени свободы для того, чтобы определить значимость коэффициента корреляции:  $df = 25 + 25 - 2 = 48$ , то есть это 48-й номер ряда в таблице Стьюдента, которую легко найти в Интернете. Число 25 в этом уравнении указывает на количество показателей, а это 25 лет. Таким образом, в таблице Стьюдента  $3,16 < 3,505$ , что означает значимость коэффициента корреляции  $< 0,001$ , но  $> 0,01$ , то есть в банк с такой значимостью я бы не пошёл.

Теперь пришло время построить уравнение линейной корреляции между двумя переменными, которые будут представлять проверочную группу. Результат должен показать влияние количества произведённого молока в России за тот же период, который изучил И.С. Скифский, то есть за 25 лет, с 1980 по 2004 гг., на насильственную преступность. Необходимые данные

сосредоточены в табл. 2. Как можно было увидеть раньше, на результат влияет масштаб показателей. Если взять цифры индекса Джини, то в сравнении с абсолютными показателями насильственных преступлений, измеряемых тысячами, получались отрицательные величины в промежуточных исчислениях, говорившие о невозможности установить среднее значение показателя, а это уже маркер ошибки. Поскольку данные о количестве произведённого молока представлены в тысячах тонн, то разумно было бы использовать абсолютные показатели насильственной преступности, чтобы сравниваемые данные были паритетными. Итак, количество произведённого молока, как влияющий фактор, у нас –  $x$ , а количество насильственных преступлений –  $y$ . Нужно выяснить, как может изменяться насильственная преступность, в зависимости от изменения количества произведённого молока в России за 25-летний период.

Таблица 2

год	$x$	$y$	$xy$	$x^2$	$y^2$
1	2	3	4	5	6
1980	46823	69176	3239027848	2192393329	4785318976
1981	45453	67741	3079031673	2065975209	4588843081
1982	47403	66621	3158035263	2247044409	4438357641
1983	50211	71824	3606354864	2521144521	5158686976
1984	50421	69614	3510007494	2542277241	4846108996
1985	50169	61699	3095377131	2516928561	3806766601
1986	52217	48805	2548450685	2726615089	2381928025
1987	52880	45857	2424918160	2796294400	2102864449
1988	54534	56889	3102384726	2973957156	3236358321
1989	55742	79563	4435000746	3107170564	6330270969
1990	55715	88051	4905761465	3104161225	7752978601
1991	51886	89931	4666159866	2692156996	8087584761
1992	47236	120948	5713099728	2231239696	14628418704
1993	46524	158735	7384987140	2164482576	25196800225
1994	42176	151852	6404509952	1778814976	23059029904
1995	39241	143603	5635125323	1539856081	20621821609
1996	35819	128295	4595398605	1283000761	16459607025
1997	34136	119041	4063583576	1165266496	14170759681
1998	33255	122288	4066687440	1105895025	14954354944
1999	32274	128293	4140528282	1041611076	16459093849
2000	32259	128951	4159830309	1040643081	16628360401
2001	32874	142324	4678759176	1080699876	20256120976
2002	33462	145923	4882875426	1119705444	21293521929
2003	33316	145475	4846645100	1109955856	21162975625
2004	31861	153148	4879448428	1015123321	23454309904
Итого	1087887	2604647	107221988406	49162412965	6784185994609
Среднее значение	43515,5	104186	4288879536,24	1966496519	271367439784
$\sigma$					
$\sigma^2$					

Среднее значение определяется по формулам:

$$\begin{aligned}\bar{x} &= \frac{\sum x_i}{n} = \frac{1087887}{25} = 43515,5 \\ \bar{y} &= \frac{\sum y_i}{n} = \frac{2604647}{25} = 104186 \\ \overline{xy} &= \frac{\sum x_i y_i}{n} = \frac{107221988406}{25} = 4288879536,24\end{aligned}$$

Для вычисления коэффициента корреляции необходимо вычислить среднеквадратическое отклонение  $\sigma(x, y)$ , которое можно вычислить, только выявив дисперсию показателей  $x$  и  $y$ , то есть показателей производства молока ( $x$ ) и уровня насильственной преступности ( $y$ ). Вычисляем дисперсию  $x$  и  $y$ :

$$\begin{aligned}D_{(x)} &= \frac{\sum x_i^2}{n} - \bar{x}^2 = \frac{49162412965}{25} - 43515,5^2 = 72897778,35 \\ D_{(y)} &= \frac{\sum y_i^2}{n} - \bar{y}^2 = \frac{271367439784}{25} - 104186^2 = -25004,64\end{aligned}$$

Теперь надо вычислить среднеквадратическое отклонение  $x$  и  $y$ :

$$\begin{aligned}\sigma_{(x)} &= \sqrt{D_{(x)}} = \sqrt{72897778,35} = 8538 \\ \sigma_{(y)} &= \sqrt{D_{(y)}} = \sqrt{-25004,64} = -158\end{aligned}$$

Следующий шаг – определение коэффициента корреляции между  $x$  и  $y$ :

$$r_{xy} = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\sigma_{(x)} \sigma_{(y)}} = \frac{4288879536,24 - 43515,5 \cdot 104186}{8538 \cdot (-158)} = \frac{-244826346,76}{-1349004} = 182$$

Выходит такое уравнение регрессии:

$$y_x = r_{xy} \frac{x - \bar{x}}{\sigma_{(x)}} + \bar{y} = 182 \frac{x - 43515,5}{8538} (-158) + 104186 = -51895,5x + 250746,3$$

Коэффициент регрессии:  $k = a = -51895,5$

Коэффициент детерминации:  $R^2 = 0,250746^2 = 0,06$ , то есть в 0,06% случаев изменения  $x$  приводят к изменению  $y$ , причём отрицательным. Иными словами, увеличение производства молока на 0,1% должно уменьшать насильственную преступность на 518 единиц. Ух ты! Пейте больше молока и кушайте больше молочных продуктов – это уменьшает количество насильственных преступлений! Вряд ли. Очевидно, тут имеет место ложная обратная корреляция. Хотя пользу чистых молочных продуктов я отрицать не стану. Но то, что их производство способствует снижению насильственной преступности нужно рассматривать, как типичный пример ложной корреляции, поскольку утверждение обратного было бы полнейшим вздором.

Нужно проверить значимость коэффициента корреляции. Для этого требуется проверить гипотезу, не равна ли  $r_{xy}$  нулю. Если да, то корреляция статистически не значима. Для того чтобы проверить гипотезу о равенстве



нулю генерального коэффициента корреляции нормальной двумерной случайной величины, надо вычислить наблюдаемое значение критерия (величина случайной ошибки):

$$T_{набл} = r_{xy} \frac{\sqrt{n-2}}{1-r_{xy}^2} = 182 \frac{\sqrt{10}}{\sqrt{1-33124}} = 0,017$$

Вычисляем степени свободы для того, чтобы определить значимость коэффициента корреляции:  $df = 25+25-2=48$ , то есть это 48-й номер ряда в таблице Стьюдента. Число 25 в этом уравнении указывает на количество показателей. Таким образом, в таблице Стьюдента  $0,017 < 1,047$ , что означает значимость коэффициента корреляции  $< 0,3$ .

Можно попытаться парировать, что увеличение производства молока связано с экономическим ростом. Но экономический рост не обязательно приводит к уменьшению экономического расслоения в обществе.

**Выводы.** Произведённые расчёты говорят о том, что проверяемость результата не сработала. Схожие или такие же результаты не были получены. Это означает, что воспроизводимости найти не удалось. Посему утверждение о том, что уровень распределения доходов населения влияет на уровень насильственной преступности, не подтверждается. Однако остаётся вопрос. Почему видна связь между динамикой индекса Джини и динамикой насильственной преступности, если посмотреть на соответствующие столбцы (x, y) в табл. 1? Вероятно, это говорит не о наличии корреляции между этими показателями, а о том, что методика расчёта индекса Джини основана на учёте количества населения, и как раз за период с 1980 по 2004 гг. в исследованном государстве количество населения в возрасте от 16 до 54 лет, то есть криминально активного населения, увеличивалось, и изменялся показатель распределения доходов. Но дело в том, что на динамику насильственной преступности влияет именно изменение количества народонаселения.

#### Литература

1. Юзиханова Э.Г. Преступность, обусловленная политико-экономическими факторами (региональный криминологический анализ на примере Тюменской области) : дисс. на соиск. уч. степени канд. юрид. наук : спец. 12.00.08 / Э.Г. Юзиханова ; Российская академия наук. Сибирское отделение – Тюменский научный центр. – М., 2002. – 193 с.
2. Ольков С.Г. О пользе и вреде неравенства (Криминологическое исследование) / С.Г. Ольков // Государство и право. – 2004. – № 8. – С. 73–78.
3. Ольков С.Г. Влияние степени неравенства в распределении доходов народонаселения на уровень умышленных убийств / С.Г. Ольков // Актуальные проблемы экономики и права. – 2011. – № 1. – С. 240–252.
4. Скифский И.С. Объяснение и прогнозирование насильственной преступности в Российской Федерации : дисс. на соиск. уч. степени канд. юрид. наук : спец. 12.00.08 / И.С. Скифский ; Тюменский государственный институт мировой экономики, управления и права. – Тюмень, 2006. – 303 с.
5. Доклад Я.И. Гилянского на конференции в Балтийском институте экологии, политики и права 15 мая 2013 года / Я.И. Доклад [Электронный ресурс]. – Режим доступа : <https://www.youtube.com/watch?v=sdH7-W2r4Jc>.

**А н н о т а ц і я**

***Сердюк П. П. Ложные корреляции в исследовании насильственной преступности.*** – Статья.

В статье на примере одного из криминологических исследований, в котором утверждается о наличии корреляции между динамикой уровня насильственной преступности и динамикой неравенства в распределении доходов населения, продемонстрировано, что использование метода линейной регрессии в криминологических исследованиях чаще всего неоправданно из-за того, что оно основано на ложной корреляции. Используя другой пример возможной корреляции показано, что корреляционные связи посредством этого метода можно установить и между совершенно не связанными явлениями.

*Ключевые слова:* индекс Джини, корреляция, линейная регрессия, насильственная преступность, фактор.

**А н о т а ц і я**

***Сердюк П. П. Хибні кореляції в дослідженні насильницької злочинності.*** – Стаття.

У статті на прикладі одного з криминологічних досліджень, в якому стверджується про наявність кореляції між динамікою рівня насильницької злочинності й динамікою нерівності в розподілі доходів населення, продемонстровано, що використання методу лінійної регресії в криминологічних дослідженнях найчастіше невиправдано через те, що воно засноване на помилковій кореляції. Використовуючи інший приклад можливої кореляції показано, що кореляційні зв'язки за допомогою цього методу можна встановити і між абсолютно не пов'язаними явищами.

*Ключові слова:* індекс Джині, кореляція, лінійна регресія, насильницька злочинність, фактор.

**S u m m a r y**

***Serdiuk P. P. The false correlations in the research of violent crime.*** – Article.

The article demonstrates on the example of one of the criminological studies in which it is asserted that there is a correlation between the dynamics of the level of violent crime and the dynamics of inequality in the distribution of the incomes of the population, it has been, that the use of the linear regression method in criminological researches is most often unreasonable, because it is based on a false correlation. It is used another example of a possible correlation, it is shown that correlation links can be established by this method between completely unrelated phenomena.

*Key words:* correlation, factor, Gini index, linear regression, violent crime.