

А.Б. Мухамедханова^{1*}, А.А. Шадиева², В. Сейтова³, С.Т. Койбагарова⁴, Б.С. Кулбай⁵

^{1,3,5}Южно-Казахстанский университет имени М. Ауэзова, Шымкент, Казахстан;

^{2,4}Центрально-Азиатский инновационный университет, Шымкент, Казахстан

¹Dia-2808@mail.ru, ²ms.shadieva@mail.ru, ³vilena_11@mail.ru,

⁴sksaltanat@mail.ru, ⁵bauka.ks@mail.ru

¹<https://orcid.org/0000-0003-2685-6125>, ²<https://orcid.org/0000-0003-0498-1945>,

³<https://orcid.org/0000-0002-4404-4916>, ⁵<https://orcid.org/0000-0002-3638-5745>

Снижение уровня преступности с помощью социально-экономической политики: регрессионный анализ

Аннотация:

Цель: Построить линейно регрессионную модель для прогнозирования количества зарегистрированных преступлений с использованием 10 социально-политических показателей.

Методы: Для данного исследования были применены методы корреляционного и регрессионного анализа. Все вычисления были выполнены в программе Microsoft Excel. За основу были взяты 1 зависимый и 10 независимых показателей.

Результаты: Определены 10 переменных, которые могут влиять на количество зарегистрированных преступлений, и были найдены статистические данные по этим показателям в Бюро национальной статистики РК с 2010 по 2022 гг. В итоге данной проверки были оставлены только два независимых переменных, на основе которых была построена линейная регрессионная модель. Модель была высоко оценена общепринятыми оценками, в том числе значимость f , p -value, нормированный R -квадрат.

Выводы: Наш анализ показывает влияние получателей адресной социальной помощи и зарегистрированных разводов на уровень преступности, что позволяет сформулировать высокоэффективную модель линейной регрессии. Благодаря точной оценке и надежной прогностической способности наши выводы проливают свет на направления политических вмешательств, направленных на борьбу с преступностью и повышение благосостояния и развития общества в Казахстане.

Ключевые слова: преступность, город, зависимость, социальные, экономические, факторы, статистика, эконометрика.

Введение

Понимание и снижение уровня преступности имеют первостепенное значение для благополучия и прогресса нации. В контексте Казахстана, страны, занимающей 48-е место из 142 стран по индексу преступности на начало 2023 г., снижение уровня преступности является обязательным условием для содействия развитию и приведения в соответствие с мировыми стандартами. Чтобы решить эту проблему, мы приступили к исследованию с целью прогнозирования количества зарегистрированных преступлений с использованием 10 социально-экономических факторов. Наше исследование нацелено на выявление влияния социально-экономических факторов на уровень преступности, чтобы сформулировать высокоэффективную модель линейной регрессии, благодаря которой можно будет прогнозировать направления политических вмешательств, направленных на борьбу с преступностью и повышение благосостояния и развития общества в Казахстане.

Обзор литературы

При проведении данного исследования за основу были взяты труды ученых, которые были проведены с учетом аналогичных количественных анализов зависимости преступности от определенных социальных и экономических факторов с использованием статистики и эконометрического моделирования. Например, Ботос выявил, что существуют определенные социальные и экономические характеристики, которые влияют на формирование уровня преступности в США как положительно, так и отрицательно (J. Bothos, 2016).

Мэтью Мелик (2003) обнаружил, что предоставление пособий по социальному обеспечению безработным служит сдерживающим фактором для совершения преступных действий (M. Melick, 2003). Кроме того, он подчеркнул важность как мотивации потенциальных преступников, так и нали-

* Автор-корреспондент. E-mail: Dia-2808@mail.ru

чие подходящих жертв для оказания влияния на преступную деятельность. Исаак Эрлих подошел к проблеме преступности с экономической точки зрения, построив теоретическую модель, напоминающую рыночную структуру (I. Ehrlich, 1996). В рамках этой модели отдельные лица взвешивают потенциальную выгоду от незаконной деятельности с учетом риска обнаружения и последующих юридических последствий. Честер Бритт наблюдал отрицательную корреляцию между безработицей среди молодежи и уровнем насильственных и имущественных преступлений в Соединенных Штатах в период с 1958 по 1990 год, что указывает на то, что возможности трудоустройства для молодежи могут смягчать преступное поведение (C. Britt, 1994). Дэвид Кантор и Кеннет Лэнд (C. David, 1985) подчеркнули влияние экономических условий на уровень преступности в Соединенных Штатах после Второй мировой войны. В частности они отметили, что периоды высокой безработицы совпадают с увеличением возможностей для преступной деятельности. Кроме того, Кантор и Лэнд выделили продолжительность безработицы в качестве фактора, отметив, что люди, страдающие от длительной безработицы, с большей вероятностью станут преступниками по сравнению с теми, кто сталкивается с более короткими периодами безработицы.

Данн подчеркивает глубокое влияние бедности на склонность людей к преступному поведению (A. Dunn, 2023). Тейлор предполагает, что бедность постепенно разрушает восприятие людьми рисков, связанных с преступными деяниями, поскольку они понимают, что им нечего терять, и чувствуют себя загнанными в ловушку обстоятельств с ограниченными возможностями для побега. Йенс Людвиг и другие приводят данные, свидетельствующие о том, что в районах с высоким уровнем бедности, как правило, наблюдается более высокий уровень насильственных преступлений (J. Ludwig, 2001). Кроме того, они обнаружили, что подростки, живущие в таких районах, более склонны бросать школу и вовлекаться в бандитскую деятельность. Чинг-Чи Се и М.Д. Пью провели метаанализ недавних исследований и пришли к выводу, что бедность и неравенство доходов неизменно коррелируют с более высоким уровнем преступности в различных контекстах (C. Hsieh, 1993). Э. Бритт Паттерсон вносит свой вклад в это обсуждение, подчеркивая тесную связь между абсолютной бедностью и уровнем преступности в стране (E. Patterson, 2006).

Методы

Для данного исследования были применены методы корреляционного и регрессионного анализа. Все вычисления были выполнены в программе Microsoft Excel. В качестве зависимой переменной было взято количество зарегистрированных преступлений. В качестве независимых были взяты следующие 10 социально-политических показателей: Безработное население, численность получателей адресной социальной помощи, число семей назначенной государственной жилищной помощи, численность получателей назначенных государственных социальных пособий, доходы населения, использованные на потребление, доля пользователей сети Интернет в возрасте 16–74 лет, численность населения, количество человек, численность городского населения, количество человек, число зарегистрированных разводов, объем розничной торговли алкоголи. Данные показатели были выбраны авторами из числа других на основе литературного анализа данной тематики. Модель была высоко оценена общепринятыми оценками в том числе значимость f , p -value, нормированный R -квадрат.

Результаты

Исследовательская работа состоит из следующих шагов:

- определить переменные, которые могут влиять на количество зарегистрированных преступлений, и найти статистические данные по этим показателям в Бюро национальной статистики РК с 2010 по 2022 гг.;
- сделать предварительную фильтрацию и обработку данных;
- определить корреляцию между независимыми переменными и, при выявлении сильной связи, исключить наименее важные из них;
- построить линейную регрессионную модель с количеством зарегистрированных преступлений и оставшимися зависимыми переменными;
- оценить и произвести корректировку модели;
- получить регрессионную модель и сравнить ее с результатами других исследователей.

Для составления уравнения и построения модели линейной регрессии была использована программа Microsoft Excel, с ее пакетом расширения «анализ данных».

Для данной статьи авторы в качестве независимых переменных выбрали следующие 10 показателей:

X1 — безработное население, количество человек;

X2 — численность получателей адресной социальной помощи, количество человек;

X3 — число семей назначенной государственной жилищной помощи, количество семей;

X4 — численность получателей назначенных государственных социальных пособий, количество человек;

X5 — доходы населения, использованные на потребление, в тенге;

X6 — доля пользователей сети Интернет в возрасте 16–74 лет (данный показатель был изменен на количество пользователей путем выявления процента от населения соответствующего года), количество человек;

X7 — численность населения, количество человек;

X8 — численность городского населения, количество человек;

X9 — число зарегистрированных разводов, количество единиц;

X10 — объем розничной торговли алкоголя, в млн тенге.

В качестве зависимой переменной было взято число зарегистрированных преступлений (количество единиц, отмечено как Y).

Таблица 1. Статистические данные по количеству зарегистрированных преступлений и 10 независимых показателей в Республике Казахстан в период с 2010 по 2022 гг.

Год	Число зарегистрированных преступлений, единиц	Безработное население	Численность получателей адресной социальной помощи	Число семей назначенной государственной жилищной помощи	Численность получателей назначенных государственных социальных пособий	Доходы населения, использованные на потребление по регионам	Доля пользователей сети Интернет в возрасте 16–74 лет	Численность населения	Численность городского населения	Число зарегистрированных разводов в Республике Казахстан	Объем розничной торговли алкоголя
	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
2010	131896	496500	196266	184544	767239	26152	5122973	16203274	8819620	41617	725622
2011	206801	473000	141621	144940	778504	30618	8139565	16440470	8973922	44862	733642
2012	287681	474781	97280	124380	768747	33745	11321600	16673933	9127543	48513	720234
2013	359844	470696	60705	123315	791631	36760	11431326	16910246	9277871	51482	740149
2014	341291	451891	56108	110002	774149	39256	11686542	17160855	9433575	52673	721476
2015	386718	454242	38244	93868	767458	40675	13444931	17415715	9837025	53293	714165
2016	361689	445467	28786	97074	785353	44197	14171256	17669896	10035577	51993	733360
2017	316418	442279	23267	76623	781692	48618	14603344	17918214	10331492	54626	727066
2018	292286	443644	571584	68389	779587	53224	15143219	18157337	10509796	54797	724790
2019	243462	440652	2177176	54476	787092	57426	15930561	18395567	10698208	59796	727296
2020	162783	448805	936189	37368	790385	61358	16582283	18631779	10938652	48002	742383
2021	157884	449644	990539	32237	804167	69038	17614622	18879552	11151376	48239	755928
2022	157473	458270	775388	28170	844314	79222	18430485	19503159	11991238	44517	799797

Примечание – Составлена авторами на основе Бюро национальной статистики Агентства РК по стратегическому планированию и реформам, <https://stat.gov.kz>

Сначала следует провести фильтрацию данных, то есть определить и убрать из таблицы выбросы данных. Для этого по каждой из переменных был применен инструмент «условное форматирование» из программы Microsoft Excel. Из таблицы 2 видно, что показатели X2 (численность получателей адресной социальной помощи) и X9 (число зарегистрированных разводов) имеют резко отличное значение в 2019 году. По остальным годам из рассматриваемого периода резких изменений не было замечено. Соответственно, для дальнейшего анализа из таблицы были исключены показатели для 2019 года.

Таблица 2. Визуальное представление выбросов

	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
2010	131896	496500	196266	184544	767239	26152	5122973	16203274	8819620	41617	725622
2011	206801	473000	141621	144940	778504	30618	8139565	16440470	8973922	44862	733642
2012	287681	474781	97280	124380	768747	33745	11321600	16673933	9127543	48513	720234
2013	359844	470696	60705	123315	791631	36761	11431326	16910246	9277871	51482	740149

Продолжение таблицы 2

	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
2014	341291	451891	56108	110002	774149	39256	11686542	17160855	9433575	52673	721476
2015	386718	454242	38244	93868	767458	40675	13444931	17415715	9837025	53293	714165
2016	361689	445467	28786	97074	785353	44198	14171256	17669896	10035577	51993	733360
2017	316418	442279	23267	76623	781692	48618	14603344	17918214	10331492	54626	727066
2018	292286	443644	571584	68389	779587	53224	15143219	18157337	10509796	54797	724790
2019	243462	440652	2177176	54476	787092	57426	15930561	18395567	10698208	59796	727296
2020	162783	448805	936189	37368	790385	61358	16582283	18631779	10938652	48002	742383
2021	157884	449644	990539	32237	804167	69038	17614622	18879552	11151376	48239	755928
2022	157473	458270	775388	28170	844314	79222	18430485	19503159	11991238	44517	799797

Примечание – Составлена авторами на основе Бюро национальной статистики Агентства РК по стратегическому планированию и реформам, <https://stat.gov.kz>

Следующим шагом после получения отфильтрованных данных является проверка взаимосвязи между зависимым и независимыми переменными. Для этого авторы построили корреляционную матрицу и определили, насколько крепка связь между каждым из них (табл. 3). Таблица является диагонально симметричной, по этой причине верхняя часть полностью повторяет нижнюю. Из таблицы 3 мы видим, насколько зависимые переменные имеют корреляционную связь с зависимым переменным. Чем крепче эта связь, тем больше влияния имеют исследуемые показатели. Обычно исследователи используют число 0,7 и выше ($r > 0,7$) как показатель достаточного влияния. Таблица показывает, что переменные X2 и X9 показывают наиболее высокие коэффициенты корреляции r , которые равны -0,71 и 0,83 соответственно. Авторы решили убрать из модели остальные переменные для создания более точной модели.

Таблица 3. Корреляционная матрица

	Y	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
Y	1,00										
X1	-0,35	1,00									
X2	-0,71	-0,27	1,00								
X3	0,16	0,77	-0,74	1,00							
X4	-0,42	-0,24	0,62	-0,68	1,00						
X5	-0,37	-0,59	0,82	-0,95	0,84	1,00					
X6	-0,02	-0,80	0,64	-0,98	0,67	0,92	1,00				
X7	-0,28	-0,68	0,77	-0,97	0,78	0,99	0,95	1,00			
X8	-0,33	-0,63	0,78	-0,95	0,81	0,99	0,92	1,00	1,00		
X9	0,83	-0,73	-0,30	-0,27	-0,27	0,01	0,36	0,12	0,06	1,00	
X10	-0,37	-0,59	0,82	-0,95	0,84	1,00	0,92	0,99	0,99	0,01	1,00

Примечание – Составлена авторами на основе Бюро национальной статистики Агентства РК по стратегическому планированию и реформам, <https://stat.gov.kz>

Также часто бывают случаи, когда существует большая связь между независимыми переменными. В науке это называют мультиколлинеарность. Так, можно наблюдать высокую корреляционную связь переменных X3, X4, X5, X6 и X7 и X10 с другими переменными. Для того чтобы модель выдавала меньше ошибок и показывала больший коэффициент надежности следуют исключить мультиколлинеарность. Из таблицы ниже мы можем наблюдать, что выбранные ранее переменные (X2 и X9) не имеют сильной связи с друг другом ($r = -0,3$), поэтому их можно использовать для дальнейшего создания модели. Связь остальных переменных не влияет для модели данного исследования, однако отсюда можно сделать следующие интересные выводы:

А. Доходы населения, использованные на потребление, растут с такой же скоростью, как и численность городского населения и население всей страны.

Б. Рост дохода населения, использованного на потребление, идентичен росту объема розничной торговли алкоголем.

С. Также можно заметить сильную корреляцию численности городского населения и населения страны с ростом объема розничной торговли алкоголем. В результате отбора количества переменных сократилось до 2 показателей: X2 (численность получателей адресной социальной помощи) и X9 (число зарегистрированных разводов).

Таблица 4. Корреляционная матрица, где $r > 0,7$ по отношению к Y

	Y	$X1$	$X2$
Y	1		
$X2$	-0,705865496	1	
$X9$	0,825127646	-0,2990153	1

Примечание – Составлена авторами на основе Бюро национальной статистики Агентства РК по стратегическому планированию и реформам, <https://stat.gov.kz>

Коэффициент корреляции может показать только силу связи переменных, однако она не может предсказывать, как один параметр может изменить другой. Для этого ученые используют другой анализ, который заключается в построении линейной регрессионной модели. Суть этой модели заключается в том, чтобы построить модель на основе линейного уравнения, который может предсказать изменение зависимой переменной Y , при изменении независимой переменной X (или нескольких переменных). Формула линейного уравнения регрессионной модели выглядит следующим образом:

$$Y = a_0 + a_1x_1 + \dots + a_nx_n,$$

где Y — зависимая переменная. В нашем случае — число зарегистрированных преступлений;
 a_0, a_1, a_n — коэффициенты регрессии;
 x_1, x_n — независимые переменные.

Это уравнение отражает равномерное изменение показателя эффективности при изменении факторных показателей. Прогнозируемый расчет числа зарегистрированных преступлений производится путем подстановки значений соответствующих факторов в планируемое уравнение. Значения параметров коэффициентов регрессии (a_0, a_1 и т.д.) можно найти различными способами. Наиболее распространенным является метод наименьших квадратов. При использовании этого метода линия, выравнивающая эмпирические данные, должна проходить таким образом, чтобы сумма квадратов отклонений от этой линии была наименьшей. Используя инструмент анализа данных из MS Excel, мы можем получить следующие данные (табл. 5).

Следующий шаг — это оценка модели. Далее будут выделены основные показатели, представляющие интерес в этой таблице. Одним из них является нормированный коэффициент детерминации (R -квадрат), определяющий долю дисперсии зависимой переменной (Y), которая может быть объяснена изменениями независимой переменной (переменных) (X), тем самым оценивая адекватность регрессионной модели. В данном случае значение нормированного R -квадрата, равно 0,89, указывает на достаточно высокий уровень объяснительной способности. Это означает что изменение числа зарегистрированных преступлений (Y) на 89 процентов может быть объяснено изменениями численности получателей адресной социальной помощи ($X2$) и числом зарегистрированных разводов ($X9$), и на 11 процентов изменениями других случайных показателей.

Таблица 5. Корреляционная матрица, где $r > 0,7$ по отношению к Y

Регрессионная статистика	
Множественный R	0,955167058
R-квадрат	0,912344108
Нормированный R-квадрат	0,892865021
Стандартная ошибка	30807,09536
Наблюдения	12

Дисперсионный анализ					
	df	SS	MS	F	Значимость F
Регрессия	2	88904055472	44452027736	46,83710795	1,74789E-05
Остаток	9	8541694120	949077124,5		
Итого	11	97445749593			

	Коэффициенты	Стандартная ошибка	t-статистика	P-значение	Нижние 95 %	Верхние 95 %	Нижние 95,0 %	Верхние 95,0 %
Y -пересечение	-429215,088	115623,5349	-3,7121775	0,004828	-690773,6	-167656,4	-690773,69	-167656
$X2$	-0,12500882	0,025640501	-4,8754436	0,000876	-0,183011	-0,067005	-0,1830116	-0,06701
$X9$	14,8043549	2,270425315	6,5205205	0,000108	9,668296	19,94041	9,6682960	19,94041

Примечание – Составлена авторами на основе Бюро национальной статистики Агентства РК по стратегическому планированию и реформам, <https://stat.gov.kz>

Далее следует обратить внимание на значимость F -статистики. Этот показатель оценивает общую значимость регрессионной модели, позволяя определить, точно ли модель отражает взаимосвязь между переменными, подтвержденную экспериментальными данными, и достаточно ли включенных объясняющих переменных для описания зависимой переменной. Обычно уровень значимости для F -критерия не должен превышать 0,05. В настоящем анализе уровень значимости определен как 1,75E-05, что подтверждает общую значимость модели.

Кроме того, были вычислены P -значения для дальнейшей оценки значимости отдельных коэффициентов. Значение P представляет собой наименьший уровень значимости, при котором нулевая гипотеза может быть отклонена на основе рассчитанной тестовой статистики. Как правило, значения P сравниваются со стандартными пороговыми значениями значимости, такими как 0,005 или 0,01. Например, значение P , равное 0,005, указывает на то, что вероятность того, что нулевая гипотеза верна, составляет всего 0,5 %. Следовательно, меньшие значения P указывают на более веские доказательства против нулевой гипотезы. Из таблицы мы видим, что P -значения для Y -пересечения, X_2 и X_9 (0,005, 0,0008 и 0,0001) подходят под эти требования и являются статистически значимыми.

Модель является статистически значимой по основным характеристикам оценки, поэтому формула линейного уравнения регрессионной модели будет выглядеть следующим образом:

$$Y_1 = a_0 + a_1 x_2 + a_2 x_9,$$

где Y_1 — число зарегистрированных преступлений;

a_0, a_1, a_2 — коэффициенты регрессии;

X_2 — численность получателей адресной социальной помощи;

X_9 — число зарегистрированных разводов.

Коэффициенты a_0, a_1, a_2 можно найти методом наименьших квадратов или посмотреть в таблице 5 в столбце коэффициенты. Уравнение в его окончательном виде с учетом найденных параметров a_0, a_1, a_2 примет следующий вид:

$$Y_1 = -429215 - 0,125x_2 + 14,804x_9.$$

Для проверки правильности вычисленных значений a_0, a_1, a_2 эти значения подставляются в исходные уравнения системы. Сравнив исходные уравнения с рассчитанными коэффициентами, можно определить точность вычисления.

Если мы подставим в уравнение регрессии различные значения численности получателей адресной социальной помощи и числа зарегистрированных разводов, то получим теоретические значения числа зарегистрированных преступлений (Y_1), соответствующие этим показателям (табл. 6).

Таблица 6. Теоретические значения числа зарегистрированных преступлений (Y_1) и процент ошибки модели

Год	X1	X2	Y	Y1	Ошибка, %
2010	196266	41617	131896	162362,7713	23,10
2011	141621	44862	206801	217234,0102	5,04
2012	97280	48513	287681	276827,7263	3,77
2013	60705	51482	359844	325354,0538	9,58
2014	56108	52673	341291	343560,7061	0,67
2015	38244	53293	386718	354972,5637	8,21
2016	28786	51993	361689	336909,2357	6,85
2017	23267	54626	316418	376579,026	19,01
2018	571584	54797	292286	310566,1092	6,25
2020	936189	48002	162783	164391,6762	0,99
2021	990539	48239	157884	161106,079	2,04
2022	775388	44517	157473	132900,0425	15,60
				Средняя ошибка	8,43

Примечание – Составлена авторами на основе Бюро национальной статистики Агентства РК по стратегическому планированию и реформам, <https://stat.gov.kz>

Коэффициенты уравнения регрессии a_0, a_1, a_2 обычно называются коэффициентами регрессии и служат ключевыми показателями в уравнении. Эти коэффициенты определяют среднее изменение переменной Y на единицу изменения переменных X на одну измеряемую единицу.

На основе перечисленных выше результатов можно сделать следующие выводы:

Увеличение численности получателей адресной социальной помощи на 1000 человек приведет к уменьшению количества зарегистрированных преступлений на 125 единиц.

Рост числа зарегистрированных разводов на 1 единицу приведет к росту числа зарегистрированных преступлений приблизительно на 15 единиц.

Свободный член уравнения a_0 (равен -429215) является определенной основой, которую необходимо учитывать при использовании коэффициента регрессии.

Используя $Y1$, мы можем оценить точность модели. В столбце 6 (ошибки) таблицы 6 был рассчитан процент ошибки модели. Средняя частота ошибок составляет 8,43 %. С учетом этой и других упомянутых выше проверок, в частности, R -квadrата (более 89 %), P -значений (0,005, 0,0008 и 0,0001 для Y , X_2 и X_9 соответственно) и значимости F (не превышает 0,05), можно сделать вывод, что модель дает вполне приемлемый результат, и с ее помощью можно составить прогноз числа зарегистрированных преступлений при заданных (планируемых) значениях численности получателей адресной социальной помощи и числа зарегистрированных разводов.

Обсуждение

Из литературного анализа было выявлено, что на уровень преступности влияет целый ряд факторов, включая социально-демографические характеристики, экономический климат, деятельность полиции и структуру, плотность населения и уровень грамотности (J. Saini, 2019), а также социально-экономические характеристики (K. Burdett, 2003). Увеличение числа круглосуточных магазинов также было связано с ростом преступности. Другие факторы, такие как генетика, динамика семьи, влияние сверстников, бедность, а также психологические, культурные, технологические и политические факторы, были определены в качестве факторов, способствующих росту уровня преступности (S. Al-lana, 2013).

В результате проведенного исследования было выявлено, что большинство выбранных факторов не показали достаточного показателя влияния на зависимую переменную. Однако авторам удалось выявить связь между количеством зарегистрированных преступлений и двумя из десяти факторов. В итоге данная модель была проверена широко употребляемыми критериями оценки, показав высокие показатели. Гипотезу исследования можно оценить как частично оправданной.

Данное исследование оценивало взаимосвязь показателей из одного и того же периода. То есть, например, оценивалось как рост числа зарегистрированных разводов за определенный год повлиял на рост количества зарегистрированных преступлений того же года. В одних случаях это взаимосвязь действительно важна и дает определенные выводы. Однако эффект некоторых социальных и экономических факторов может быть виден сразу, но через определенное время. Поэтому для дальнейшего ориентира авторы планируют исследовать эту тематику дальше, пытаясь определить наличие эффекта этих факторов, спустя 1, 3, 5 и 7 лет.

Результаты данного исследования имеют теоретическую значимость в виде ориентира для дальнейших исследований. Так как 8 из 10 факторов не показали свой эффект в казахстанских реалиях, встает вопрос будут ли иметь эти факторы эффект и спустя какое время это произойдет. Практическая значимость заключается в том, что, определив взаимосвязь между количеством преступлений и социально-экономическими факторами, можно будет снизить либо предотвратить рост количества преступлений.

Заключение

Для того чтобы ответить на вопрос данного исследования был выполнен следующий объем работ:

Определены 10 переменных, которые могут влиять на количество зарегистрированных преступлений, и найдены статистические данные по этим показателям в Бюро национальной статистики РК с 2010 по 2022 гг. Некоторые показатели были переделаны для корректности сравнения с другими. Например, доля пользователей сети Интернет в возрасте 16–74 лет был изменен на количество пользователей путем выявления процента от населения соответствующего года. Следующим шагом была проведена предварительная фильтрация и обработка данных. Это было сделано, чтобы убрать из подборки данных выбросные значения и сделать модель более точной. Были определены корреляция между зависимым и независимым переменными, а также была проведена проверка на мультиколлинеарность. В итоге данной проверки были оставлены только два независимых переменных, на основе которых была построена линейная регрессионная модель. Модель была высоко оценена общеприня-

тыми оценками, в том числе значимость f , p -value, нормированный R -квадрат. По результатам модели выявлено, что:

- увеличение численности получателей адресной социальной помощи на 1000 человек приведет к уменьшению количества зарегистрированных преступлений на 125 единиц;
- рост числа зарегистрированных разводов на 1 единицу приведет к росту числа зарегистрированных преступлений приблизительно на 15 единиц.

Свободный член уравнения a_0 (равен -429215) является определенной основой, которую необходимо учитывать при использовании коэффициента регрессии.

Используя YI , мы можем оценить точность модели. В столбце 6 (ошибки) таблицы 6 был рассчитан процент ошибки модели. Средняя частота ошибок составляет 8,43 %. С учетом этой и других упомянутых выше проверок, в частности, R -квадрата (более 89 %), P -значений (0,005, 0,0008 и 0,0001 для Y , X_2 и X_9 соответственно) и значимости F (не превышает 0,05), можно сделать вывод, что модель дает вполне приемлемый результат, и с ее помощью можно составить прогноз числа зарегистрированных преступлений при заданных (планируемых) значениях численности получателей адресной социальной помощи и числа зарегистрированных разводов.

Список литературы

- Allana, S., & Khokhar, S. (2013). Crimes: A Threat to Public Health. *International Journal of Nursing Education*, 5, 180. 10.5958/j.0974-9357.5.2.091.
- Bothos, J., & Thomopoulos, S. (2016). Factors influencing crime rates: an econometric analysis approach. 98421A. 10.1117/12.2223395.
- Britt, C. (1994). Crime and Unemployment Among Youths in the United States, 1958–1990: A Time Series Analysis. *American Journal of Economics and Sociology*, 53, 99–109. 10.1111/j.1536-7150.1994.tb02680.x.
- Burdett, K., Ricardo, L., & Randall, W. (2003). Crime, Inequality, and Unemployment. *American Economic Review*, 93(5), 1764–1777. DOI: 10.1257/000282803322655536
- Cantor, D., & Land, K. (1985). *Unemployment and crime rates in the post WWII United States — a theoretical and empirical analysis*.
- Dunn, A. (2023). The poverty-crime nexus revisited: absolute poverty, relative poverty, and crime rates in 105 countries. *International Journal of Comparative and Applied Criminal Justice*, 1–17. 10.1080/01924036.2023.2292044.
- Ehrlich, I. (1996). Crime, Punishment, and the Market for Offenses. *Journal of Economic Perspectives*, 10, 43–67. 10.1257/jep.10.1.43.
- Hsieh, C.-C., & Pugh, M. D. (1993). Poverty, Income Inequality, and Violent Crime: A Meta-Analysis of Recent Aggregate Data Studies. *Criminal Justice Review*, 18(2), 182–202. <https://doi.org/10.1177/073401689301800203>
- Ludwig, Jens & Duncan, Greg, & Hirschfield, Paul. (2001). Urban Poverty And Juvenile Crime: Evidence From A Randomized Housing-Mobility Experiment. *The Quarterly Journal of Economics*, 116, 655–679. 10.1162/00335530151144122.
- Melick, M. D. (2003). The relationship between crime and unemployment. *The Park Place Economist*, 11(1), 30–36.
- Patterson, E.. (2006). Poverty, income inequality, and community crime rates. *Criminology*, 29, 755–776. 10.1111/j.1745-9125.1991.tb01087.x.
- Saini, J., & Srivastava, V. (2019). Impact of population density and literacy levels on crime in India. In *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1–7). IEEE.
- The Bureau of National Statistics of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. Retrieved from <https://stat.gov.kz>.

А.Б. Мухамедханова¹, А.А. Шадиева², В. Сейтова³, С.Т. Койбагарова⁴, Б.С. Кулбай⁵

^{1,3,5}М. Әуезов атындағы Оңтүстік Қазақстан университеті, Шымкент, Қазақстан;

^{2,4}Орталық Азия Инновациялық университеті, Шымкент, Қазақстан

¹Dia-2808@mail.ru, ²ms.shadieva@mail.ru, ³vilena_11@mail.ru,

⁴sksaltanat@mail.ru, ⁵bauka.ks@mail.ru

¹<https://orcid.org/0000-0003-2685-6125>, ²<https://orcid.org/0000-0003-0498-1945>,

³<https://orcid.org/0000-0002-4404-4916>, ⁵<https://orcid.org/0000-0002-3638-5745>

Әлеуметтік экономикалық саясат арқылы қылмыс деңгейін төмендету: регрессиялық талдау

Аңдатпа:

Мақсаты: 10 әлеуметтік-саяси көрсеткіштерді қолдана отырып, тіркелген қылмыстардың санын болжау үшін сызықтық регрессиялық модель құру.

Әдісі: Зерттеуде корреляциялық және регрессиялық талдау әдістері қолданылды. Барлық есептеулер MS EXCEL бағдарламасында орындалды. 1 тәуелді және 10 тәуелсіз көрсеткіш негізге алынды.

Қорытынды: Тіркелген қылмыстардың санына әсер етуі мүмкін 10 айнымалы анықталды және 2010 жылдан 2022 жылға дейінгі ҚР Ұлттық статистика бюросында осы көрсеткіштер бойынша статистикалық деректер табылды. Модель жалпы қабылданған бағалаулармен жоғары бағаланды, соның ішінде: маңыздылығы f , p — мәні, нормаланған R квадрат.

Тұжырымдама: Біздің талдауымыз атаулы әлеуметтік көмек алушылар мен тіркелген ажырасулардың қылмыс деңгейіне әсерін көрсетеді, бұл жоғары тиімді желілік регрессия моделін құруға мүмкіндік береді. Дәл бағалау және сенімді болжау күші арқылы біздің қорытындыларымыз қылмыспен күресу және Қазақстанның әлеуметтік әл-ауқаты мен дамуын жақсарту үшін саяси араласулардың бағыттарын айқындайды.

Кілт сөздер: қылмыс, қала, тәуелділік, әлеуметтік, экономикалық, факторлар, статистика, эконометрия.

A.B. Mukhamedkhanova¹, A.A. Shadieva², V. Seitova³, S.T. Koibagarova⁴, B.S. Kulbai⁵

^{1,3,5}*M. Auezov South Kazakhstan University, Kazakhstan;*

^{2,4}*Central Asian Innovation University, Kazakhstan*

¹*Dia-2808@mail.ru, ²ms.shadieva@mail.ru, ³vilena_11@mail.ru,*

⁴sksaltanat@mail.ru, ⁵bauka.ks@mail.ru

¹<https://orcid.org/0000-0003-2685-6125>, ²<https://orcid.org/0000-0003-0498-1945>,

³<https://orcid.org/0000-0002-4404-4916>, ⁵<https://orcid.org/0000-0002-3638-5745>

Mitigating Crime through socio-economic policy: regression analysis

Abstract

Object: to build a linear regression model for predicting the number of reported crimes using 10 socio-political indicators.

Methods: methods of correlation and regression analysis were used for this study. All calculations were performed in the MS EXCEL program. 1 dependent and 10 independent indicators were used as a basis.

Findings: 10 variables have been identified that can affect the number of reported crimes and statistical data on these indicators were found in the Bureau of National Statistics of the Republic of Kazakhstan from 2010 to 2022. As a result of this check, only two independent variables were left, on the basis of which a linear regression model was built. The model was highly appreciated by generally accepted estimates, including: significance of f , p -value, normalized R square.

Conclusions: Our analysis shows the influence of recipients of targeted social assistance and registered divorces on the crime rate, which allows us to formulate a highly effective linear regression model. Thanks to accurate assessment and reliable predictive ability, our findings shed light on the directions of political interventions aimed at combating crime and improving the welfare and development of society in Kazakhstan.

Keywords: crime, city, dependence, social, economic, factors, statistics, econometrics.