

Санкт–Петербургский государственный университет

Антоненко Елизавета Павловна

Выпускная квалификационная работа

*Статистический анализ факторов, влияющих
на подростковый коэффициент преступности*

Уровень образования: бакалавриат

Направление 01.03.02 «Прикладная математика и информатика»

Основная образовательная программа «Прикладная математика,
фундаментальная информатика и программирование»

Кафедра «Теория игр и статистических решений»

Научный руководитель:

Панкратова Я.Б.

Рецензент:

Балыкина Ю.Е.

Санкт-Петербург

2021 г.

Содержание

Глава 1. Введение	4
1.1. Факторы	4
1.2. Актуальность	6
1.3. Обзор литературы	7
1.4. Задачи	8
Глава 2. Теоретические основы построения регрессионных моделей	9
2.1. Понятие и построение множественной регрессии	9
2.2. Спецификация модели множественной регрессии	9
2.3. Ограничения и условия классической регрессионной модели	12
2.4. Метод наименьших квадратов	13
2.4.1 Первая предпосылка	15
2.4.2 Вторая предпосылка	16
2.4.3 Третья предпосылка	18
2.4.4 Мультиколлинеарность	20
2.5. Верификация модели	21
Глава 3. Визуализация данных	23
Глава 4. Построение множественных регрессионных моделей оценки уровня подростковой преступности	24
4.1. Работа с данными. Построение модели I	24
4.2. Работа с выбросами	27
4.3. Модель II	28
4.4. Оценка параметров модели II. Построение модели III	32
4.5. Модель IV	36
4.6. Оценка параметров модели IV. Построение модели V	38
Глава 5. Верификация построенных моделей	41
Глава 6. Заключение	43
Список литературы	45
Глава 7. Приложения	46
7.1. Приложение 1	46

7.2.	Приложение 2	48
7.3.	Приложение 3	49
7.4.	Приложение 4	50
7.5.	Приложение 5	51
7.6.	Приложение 6	52
7.7.	Приложение 6	53
7.8.	Приложение 7	54

Глава 1. Введение

Преступность — это социально-правовое негативное массовое явление, которое складывается из совокупности совершаемых в тот или иной период в государстве преступлений^{*}, характеризующихся количественными (состоянием, динамикой) и качественными (структурой и характером преступности) показателями. Соответственно, подростковая преступность — совокупность преступлений, совершенных несовершеннолетними. Целью выпускной квалификационной работы является выявление и анализ наиболее значимых факторов, влияющих на степень подростковой преступности в регионах России, а также составление и проверка значимости многофакторной регрессионной модели.

1

Анализ факторов поведения несовершеннолетних на протяжении многих десятилетий остается в областях научных исследований: юриспруденции, психиатрии, социальных наук, медицине, педагогике и др. Современные знания о расстройствах личности позволяют сделать вывод, что они представляют собой сложные формы поведения, на которые влияет система взаимосвязанных факторов. Эти факторы можно разделить на биологические и социальные.

1.1 Факторы

Ввиду сложности интерпритации биологических факторов, в данной работе исследуются только социальные ([1] - [5]):

1. Число разводов в регионе.
2. Уровень занятости населения в возрасте от 14 до 18 лет.
3. Уровень жизни в регионе.
4. Число организаций начального и среднего образования.

^{1*}На протяжении всей работы будет использоваться именно понятие преступления как отдельной группы правонарушений, нарушающих Уголовный кодекс.

5. Число молодежных и спортивных центров.
6. Уровень безработицы в регионе.
7. Численность населения в регионе.
8. Численность подростков в регионе.

Давайте разберем их более подробно.

1. Разводы.

Одним из важнейших факторов, по мнению исследователей, является развод родителей: ситуация развода не только увеличивает вероятность асоциального поведения подростка, но и приводит к снижению уровня родительского контроля, что повышает вероятность появления у подростков криминальных связей и криминального референтного сообщества.

2. Уровень занятости населения в возрасте от 14 до 18 лет.

Социальная незанятость подростков также влияет на их противоправное поведение. Большинство подростков на момент совершения преступления нигде не учатся и не работают.

3. Уровень жизни в регионе.

Низкий уровень жизни большей части населения, сильный контраст распределения доходов в обществе. Среди малообеспеченных слоев наиболее ярко выражен уровень наркотизации и алкоголизации, которые во многом являются причинами преступности и среди подростков.

4. Число организаций начального и среднего образования.

Именно в таких организациях ребенок обретает базовые знания, которые формируют мышление и способствуют его дальнейшему интеллектуальному развитию. Также он учится вливаться в коллективе, определяя свою роль в социуме, получает нравственный опыт и

осознание собственной ответственности. Здесь ребенок встречается с четко сформированной системой моральных принципов, учится понимать общие ценности и задачи, а также управлять своим поведением в соответствии с ними. Принято считать, что именно в организациях начального и среднего образования во многом формируется его личность.

5. Число молодежных и спортивных центров.

Слабая организация сети клубов, кружков, спортивных секций, отсутствие заботы о вовлечении и закреплении в них несовершеннолетних, находящихся в неблагоприятных условиях жизни и воспитания, тоже пагубно влияют на мировоззрения подростков.

6. Уровень безработицы в регионе.

Много исследований показывает, что у большинства подростков, совершивших преступление, родители либо не имели работу, либо имели работу с низкой заработной платой.

7. Численность населения и численность подростков в регионе.

Численность населения и численность подростков нужны ввиду того, что в работе исследуются разные регионы с разной численностью.

1.2 Актуальность

Актуальность и значимость изучения причин девиантного поведения подростков выражается в том, что состояние преступности среди несовершеннолетних, ее особенности и тенденции – это прогноз будущего состояния преступности в целом. Специалисты из разных областей: юристы, педагоги, психологи, социологи высказывают тревогу по поводу перспектив молодого поколения. Рост среди молодежи преступности, аморальных проявлений, то есть девиантное поведение подростков, вызывают тревогу в социуме. В научных кругах, в обществе постоянно идет обсуждение необходимости оказания социальной поддержки и помощи молодым людям, имеющим проблемы с законом, а не только применение репрессивных

мер в отношении несовершеннолетних.

Сейчас официальные статистические данные показывают значимое снижение преступности несовершеннолетних. Число преступлений, совершенных несовершеннолетними (а также с их участием), сократилось с 195,4 тыс. в 2000 году, до 61,8 тыс. в 2015 году.

Однако другие факты следует трактовать как вызывающие беспокойство. Во-первых, при оценке динамики преступности несовершеннолетних необходимо учитывать сокращение численности данной демографической группы, произошедшее за рассматриваемый промежуток времени, а также высокую долю латентной преступности, не отраженной в официальной статистической информации. Во-вторых, происходят изменения в структуре подростковой преступности. Основными видами преступлений, совершаемых несовершеннолетними, являются кражи и грабежи. После снижения доли тяжких и особо тяжких преступлений в 2000-2008 годах произошел их значительный рост в последние годы. По данным МВД, в 2016 году эта доля составила 21,5% от всех преступлений несовершеннолетних.

Неоднозначность тенденций в сфере подростковой преступности требует анализа основных причин и факторов, лежащих в основе преступности несовершеннолетних в современный период, и статистического анализа их влияния на число совершаемых преступлений. Чему и будет посвящена моя научная работа.

1.3 Обзор литературы

Преступления, совершаемые несовершеннолетними, занимают весьма существенное место как на общенациональном, так и региональном уровнях, что говорит о том, что это важная тема для исследования. Для отбора факторов, влияющих на уровень подростковой преступности, были прочитаны следующие статьи по криминологии и психологии несовершеннолетних [1] - [6]. Все данные взяты с официальных сайтов rosstat.gov.ru и fedstat.ru. Визуализация данных была проведена посредством сайта app.datawrapper.de. В работе много внимания уделяется теоретической составляющей. Основными источниками для изучения регрессионного анализа были следующие

учебники и методические пособия: [7] - [10]. Для изучения дополнительных вспомогательных характеристик регрессионного анализа были изучены следующие материалы [11] - [13]. Для успешной работы в Excel, где производились основные вычисления и анализ, были прочитаны [9],[14]

1.4 Задачи

1. Изучить теоретические основы построения регрессионных моделей: методы отбора параметров, оценка параметров регрессионного уравнения и способы оценки качества модели.
2. Собрать необходимые данные и сформировать выборку.
3. Визуализировать данные о подростковой преступности по регионам России за 2018 год.
4. Провести факторный и регрессионный анализ, построить несколько моделей.
5. Провести верификацию и сравнение построенных моделей.

Глава 2. Теоретические основы построения регрессионных моделей

2.1 Понятие и построение множественной регрессии

Регрессионный метод анализа можно применить к измерению связей между двумя признаками - парная регрессия, или к измерению связей между тремя и большим числом факторов - множественная регрессия.

Множественная регрессионная модель имеет вид:

$$y = f(x_1, x_2, \dots, x_m) + \varepsilon,$$

где y – зависимая переменная (результативный признак), x_1, x_2, \dots, x_m – независимые переменные (факторы), а ε – случайная ошибка.

Множественные регрессионные модели широко используются при решении проблем спроса и предложения, при описании потребительских и производственных функций, для исследования многих макро- и микроэкономических проблем.

Общая постановка задачи множественной регрессии:

По имеющимся данным n наблюдений за изменением признака y в зависимости от наборов значений факторов x_1, x_2, \dots, x_m выбрать экономическую модель $y = f(x_1, x_2, \dots, x_m) + \varepsilon$, оценить ее параметры и статистически обосновать, что факторы x_1, x_2, \dots, x_m существенны, а построенная модель $f(x_1, x_2, \dots, x_m)$ наиболее точно соответствует данным наблюдений.

2.2 Спецификация модели множественной регрессии

Спецификация модели множественной регрессии включает решение двух задач:

1. Выбрать независимые переменные x_1, x_2, \dots, x_m ;
2. Выбрать форму $f(x_1, x_2, \dots, x_m)$ зависимости y от переменных x_1, x_2, \dots, x_m .

При выборе независимых x_1, x_2, \dots, x_m , следует учесть, что они:

1. Должны быть количественно измеримыми. Если возникает необходимость включить в модель качественный признак, то ему придается количественная форма.
2. Не должны коррелировать между собой и тем более не находиться в функциональной связи.

Второй пункт определяет индивидуальное влияние отдельных регрессоров x_1, x_2, \dots, x_m на результат y , что является весьма актуальным для осуществления прогнозов и принятия управляющих решений.

Отбор факторов, как правило, осуществляется в несколько этапов. Сначала отбираются факторы, связанные с изучаемым явлением на основе данных теоретического исследования. Затем можно провести предварительную обработку данных, удалить выбросы. Далее отобранные и обработанные данные подвергаются проверке существенности их влияния на изучаемый показатель с использованием методов математической статистики, малозначимые факторы при этом исключаются из модели.

Один из методов отбора факторов базируется на анализе матрицы парных коэффициентов корреляции, элементами которой являются линейные коэффициенты парной корреляции факторов x_1, x_2, \dots, x_m , как в зависимости переменной y , так и между собой. Считается, что две переменные явно коллинеарны, то есть находятся друг с другом в высокой линейной зависимости, если коэффициент интеркорреляции (то есть парный коэффициент корреляции) превышает 0.7. В таком случае одна из них исключается из модели. Предпочтение отдается тому фактору, который достаточно тесно связан с результативным фактором, но имеет при этом наименьшую тесноту связи с другими объясняющими факторами.

Отбор факторов, включаемых в регрессию, является одним из важнейших этапов практического использования методов регрессии. Подходы к отбору факторов на основе показателей корреляции могут быть разными. Наиболее широкое распространение получили следующие методы построения модели множественной регрессии:

- метод исключения;

- метод включения.

Метод исключения заключается в том, что на первом шаге строится уравнение регрессии с полным набором факторов, а затем после исключения коллинеарных факторов отбираются факторы, имеющие наибольшее влияние на изменение результативного признака, менее значимые факторы при этом исключаются.

Для отбора факторов, исключаемых из модели, применяется t -критерий Стьюдента. Незначимыми коэффициентами считаются те, для которых расчетное значение критерий Стьюдента ниже табличного или величина p - значения (доверительной вероятности) превышает выбранный уровень значимости (обычно он равен 0,05).

Процедуру исключения незначимых факторов продолжают до тех пор, пока не будет получено значимое уравнение регрессии со значимыми коэффициентами.

Рассмотрим модель множественной линейной регрессии:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m + \varepsilon$$

Ввиду отмеченной четкой интерпретации коэффициентов a_j уравнения линейные множественные модели широко представлены в эконометрическом анализе. Однако реальное соотношение между социально-экономическими явлениями и процессами далеко не всегда можно выразить линейными функциями.

В таком случае прибегают к нелинейным моделям, среди которых наиболее часто применяются:

- степенная $y = a_0x_1^{a_1}x_2^{a_2}\dots x_m^{a_m}\varepsilon$

Здесь $a_j, j = 1, 2, \dots, m$, - коэффициенты эластичности, которые показывают, на сколько процентов изменится в среднем результат y с изменением фактора x_j на 1% при неизменности действия других факторов;

- экспоненциальная $y = e^{a_0+a_1x_1+a_2x_2+\dots+a_mx_m}\varepsilon$;

- гиперболическая $y = a_0 + \frac{a_1}{x_1} + \frac{a_2}{x_2} + \dots + \frac{a_m}{x_m} + \varepsilon$

В данной работе будет рассматриваться линейная модель множественной регрессии $y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m + \varepsilon$

2.3 Ограничения и условия классической регрессионной модели

Рассмотрим линейную множественную регрессию $y = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n + \varepsilon$.

Тогда в матричном виде модель наблюдений примет вид: $y = Xa + \varepsilon$, где y - вектор результата, X - матрица наблюдений размерности $[n, m]$, a - вектор параметров, ε - случайный вектор возмущений.

Задача состоит в следующем: по имеющейся выборке значений X и Y построить уравнение регрессии $\hat{y} = \hat{a}_0 + \hat{a}_1x_1 + \hat{a}_2x_2 + \dots + \hat{a}_nx_n$, коэффициенты которой $\hat{a}_0, \hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ являются оценками неизвестных параметров регрессии $a_0, a_1, a_2, \dots, a_n$. Желательно, чтобы эти оценки обладали свойствами *несмещенности*, *эффективности* и *состоятельности*.

Оценка \hat{a}_i называется *несмещенной*, если $M(\hat{a}_i) = a_i$, то есть среднее значение оценки \hat{a}_i равно оцениваемому параметру. В противном случае оценка называется смещенной.

Оценка \hat{a}_i называется *эффективной*, если среди всех других несмещенных оценок она имеет наименьшую дисперсию, то есть в меньшей степени отклонена от a_i

Оценка \hat{a}_i называется *состоятельной*, если при увеличении объема выборки n вероятностью того, что оценка \hat{a}_i будет отличаться от a_i на сколь угодно малую величину $\hat{\varepsilon}$ будет равна нулю.

Существует несколько методов вычисления коэффициентов a_0, a_1, \dots, a_m . На практике в большинстве случаев используется метод наименьших квадратов (МНК).

Теорема Гаусса-Маркова. Если выполнены условия:

1. Математическое ожидание случайной величины ε_i равно нулю: $M(\varepsilon_i) = 0$;

2. Дисперсия случайной величины ε_i постоянна для всех наблюдений:
 $D(\varepsilon_i) = \sigma^2 = \text{const}$;
3. Отсутствует систематическая связь между значениями случайных величин: $(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$,

тогда оценки \hat{a}_i метода наименьших квадратов являются наилучшими линейными несмещенными оценками.

При построении классических линейных множественных регрессионных моделей необходимо выполнение и таких предположений, как:

1. Отсутствие мультиколлинеарности (нет зависимости между факторами);
2. Число наблюдений существенно больше числа объясняющих переменных (по крайней мере, в три раза);
3. Отсутствуют ошибки спецификации.

2.4 Метод наименьших квадратов

Оценки неизвестных параметров $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_m$ линейной модели множественной регрессии находятся с помощью метода наименьших квадратов из условия оптимизации функции:

$$S = S(\hat{a}_0, \hat{a}_1, \dots, \hat{a}_m) = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

то есть из условия:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min$$

Для нахождения параметров $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_m$ на основании необходимого условия экстремума приравниваются к нулю частные производные функции $S(\hat{a}_0, \hat{a}_1, \dots, \hat{a}_m)$ по переменным $\hat{a}_0, \hat{a}_1, \dots, \hat{a}_m$. В итоге получается система, содержащая $m+1$ линейным уравнением (по числу параметров) с $m+1$ переменными:

$$\left\{ \begin{array}{l} n\hat{a}_0 + \hat{a}_1 \sum_{i=1}^n (x_1)_i + \hat{a}_2 \sum_{i=1}^n (x_2)_i + \dots + \hat{a}_m \sum_{i=1}^n (x_m)_i = \sum_{i=1}^n y_i, \\ \hat{a}_0 \sum_{i=1}^n (x_1)_i + \hat{a}_1 \sum_{i=1}^n (x_2)_i^2 + \hat{a}_2 \sum_{i=1}^n (x_1)_i (x_2)_i + \dots + \hat{a}_m \sum_{i=1}^n (x_1)_i (x_m)_i = \sum_{i=1}^n (x_1)_i y_i, \\ \dots, \\ \hat{a}_0 \sum_{i=1}^n (x_m)_i + \hat{a}_1 \sum_{i=1}^n (x_m)_i (x_1)_i + \hat{a}_2 \sum_{i=1}^n (x_m)_i (x_2)_i + \dots + \hat{a}_m \sum_{i=1}^n (x_m)_i^2 = \sum_{i=1}^n (x_m)_i y_i. \end{array} \right.$$

Здесь n - число наблюдений y_1, y_2, \dots, y_n зависимой переменной y , а $(x_j)_1, (x_j)_2, \dots, (x_j)_n$ - наблюдаемые значения j -ого фактора ($j = 1, 2, \dots, m$).

Для решения системы может быть применен метод Крамера, метод Гаусса, матричный метод или любой другой метод решения систем линейных уравнений.

Введем величину $e_i = y_i - \hat{y}_i$, характеризующую отклонение выборочного значения y_i от предсказанного \hat{y}_i . Эту величину назовем невязкой (или остатком) регрессии в i -ой точке. Тогда измеренные значения y_i можно записать выражением:

$$y_i = a_0 + a_1 x_{1i} + a_2 x_{2i} + \dots + a_n x_{ni} + e_i$$

Метод наименьших квадратов применительно к множественной линейной регрессионной модели дает хорошие результаты (несмещенные, эффективные и состоятельные оценки параметров регрессии) при выполнении определенных требований к случайной величине e (Теорема Гаусса-Маркова).

Линейная модель, оценки параметров которой несмещенные, эффективные и состоятельные, называется классической линейной моделью множественной регрессии. Если же в дополнение выполняется предположение о нормальном распределении случайной величины e , то классическая линейная модель называется нормальной.

Если игнорировать проверку выполнимости модельных предположений, то регрессионная модель может оказаться статистически незначимой, а значит, прогнозы по ней будут подвергаться сомнению. В связи с этим и возникает необходимость рассмотрения методов обнаружения и устранения

нарушений предпосылок МНК. Проверка их является важным и неотъемлемым этапом верификации регрессионной модели.

2.4.1 Первая предпосылка

Первая предпосылка говорит о том, что во всех наблюдениях среднее значение (математическое ожидание) случайного возмущения равно нулю. Эта предпосылка отвечает за несмещенность параметров линейной модели множественной регрессии.

Если в регрессионное уравнение включен свободный член, то условие 1 теоремы Гаусса-Маркова никогда не нарушается. В тех же случаях, когда возникает необходимость рассмотрения уравнения регрессии с нулевым свободным членом, первое модельное предположение может быть нарушено. Поэтому для устранения проблем с первой предпосылкой в модель регрессии следует включать свободный член.

Как известно, несмещенной оценкой математического ожидания случайной величины является выборочное среднее. Поэтому для проверки предположения о равенстве нулю математического ожидания случайной величины достаточно оценить величину среднего выборочного остатков регрессии $e_i = y_i - \bar{y}_i, i = 1, 2, \dots, n$, т.е. величину $\bar{e} = \frac{1}{n} \sum_{i=1}^n e_i$.

Если есть сомнения при оценке выборочного среднего, можно рассмотреть статистический критерий о значении математического ожидания нормально распределенной генеральной совокупности.

В нашем случае дисперсия генеральной совокупности неизвестна, поэтому рассмотрим *алгоритм критерия о равенстве математического ожидания нормально распределенной генеральной совокупности нулю* при условии, что дисперсия неизвестна:

1. Выдвигаем нулевую $H_0 : a = 0$ и альтернативную $H_1 : a \neq 0$ гипотезы.
2. Задаем уровень значимости α .
3. Вычисляем значение статистики критерия:

$$\tau = \frac{\bar{e}\sqrt{n}}{s},$$

где $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2}$ - стандартное отклонение.

4. Определяем критическую область, соответствующую гипотезе $H_1 : a \neq 0$. В данном случае рассматривается двусторонняя критическая область: $(-\infty; t_{\frac{\alpha}{2}, n-1}) \cup (t_{1-\frac{\alpha}{2}, n-1}; \infty)$, где $t_{\frac{\alpha}{2}, n-1}$ и $t_{1-\frac{\alpha}{2}, n-1}$ - квантили уровней $\frac{\alpha}{2}$ и $1 - \frac{\alpha}{2}$ распределения Стьюдента (однопараметрическое семейство абсолютно непрерывных распределений) с $n - 1$ степенью свободы. В силу симметрии плотности распределения Стьюдента имеет место равенство $t_{\frac{\alpha}{2}, n-1} = t_{1-\frac{\alpha}{2}, n-1} = t_{cr}$. Тогда, если численное значение статистики τ не попадает в критическую область, т.е. $|\tau| < t_{cr}$, то нулевая гипотеза H_0 принимается, в противном случае отвергается.

2.4.2 Вторая предпосылка

Вторая предпосылка теоремы требует, чтобы во всех наблюдениях дисперсия случайного возмущения была одинаковой. Это свойство получило название свойства гомоскедастичности или однородности. В случае невыполнения данного условия говорят, что случайные возмущения в уравнениях наблюдения гетероскедастичные или неоднородные. Если случайные возмущения гетероскедастичные, то оценки параметров модели остаются несмещенными, но теряется эффективность оценки дисперсий параметров.

В настоящее время для оценки нарушения гомоскедастичности предложено большое число тестов. Чаще всего используются графический анализ отклонений, тест ранговой корреляции Спирмена и тест Голдфелда-Квандта. Наиболее простым и наглядным является графический метод, основанный на графическом представлении отклонений. В этом случае по оси абсцисс откладывается объясняющая переменная (либо линейная комбинация объясняющих переменных), а по оси ординат — остатки e_i либо квадраты остатков e_i^2 . В случае гомоскедастичности все остатки распо-

ложены внутри некоторой полосы, параллельной оси абсцисс. Однако, во многих случаях гетероскедастичность не столь очевидна визуально. Наиболее популярным тестом обнаружения гетероскедастичности является тест Голдфелда-Квандта. Тест применяется в том случае, если ошибки регрессии можно считать нормально распределенными случайными величинами и отсутствует автокорреляция возмущений (глава 2.4.3).

Тест состоит в следующем:

1. На основе выборочных данных строится линейная модель множественной регрессии с m объясняющими переменными x_1, x_2, \dots, x_m .
2. В модели множественной регрессии (например, на основе графического анализа) выбирается факторная переменная, от которой предположительно могут зависеть остатки. Значения этой переменной ранжируются, располагаются по возрастанию и делятся на три части объемами $k, (n - 2k), k$ (обычно принимают $k \approx \frac{n}{3}$).
3. Для первой и третьей частей строятся две независимые модели регрессии.
4. По каждой из построенных моделей рассчитывают суммы квадратов остатков S_1 и S_3 .
5. Осуществляется проверка основной гипотезы об отсутствии гетероскедастичности с помощью F -критерия Фишера. Значение F -критерия рассчитывается следующим образом:

$$F_{obs} = \frac{S_3}{S_1}, \text{ если } S_3 > S_1$$

или

$$F_{obs} = \frac{S_1}{S_3}, \text{ если } S_1 > S_3$$

6. Если $F_{obs} > F_{cr}$, то в основной модели присутствует гетероскедастичность, зависящая от выбранной объясняющей переменной (число степеней свободы определяется значениями $k_1 = k - m - 1$ и

$$k_2 = k - m - 1).$$

Если нет уверенности относительно выбора объясняющей переменной, вызывающей гетероскедастичность, то тест осуществляется для каждой из объясняющих переменных x_1, x_2, \dots, x_m .

При установлении гетероскедастичности возникает необходимость ее устранения, для чего необходимо преобразование модели. На практике чаще всего значения дисперсий возмущений σ_i^2 неизвестны, поэтому, прежде всего, пытаются найти оценки этих дисперсий. После этого используется взвешенный метод наименьших квадратов, согласно которому для нахождения оценок коэффициентов регрессии минимизируется взвешенная сумма остатков:

$$\sum_{i=1}^n \left(\frac{y_i}{\sigma_i} - \frac{a_0}{\sigma_i} - a \frac{x_i}{\sigma_i} \right)^2$$

Тем самым наблюдениям с наименьшими дисперсиями придаются наибольшие веса, а наблюдениям с наибольшими дисперсиями – наименьшие веса.

Главная проблема взвешенного метода наименьших квадратов состоит в необходимости знания среднеквадратических отклонений σ_i случайных ошибок регрессии. На практике обычно дисперсии σ_i^2 неизвестны. В таком случае делаются реалистические предположения об их величине. В частности, принимаются предположения о том, что либо дисперсии σ_i^2 отклонений, либо сами среднеквадратичные отклонения σ_i пропорциональны значениям переменной x_i или значению линейной комбинации $x_i, i = 1, \dots, m$.

2.4.3 Третья предпосылка

Третья предпосылка теоремы требует независимости случайных возмущений. Другими словами, каким бы не оказалось значение случайного возмущения в первом наблюдении, оно никак не сказывается на значениях случайного возмущения в любом другом наблюдении. Это свойство получило название неавтокоррелируемости случайных возмущений. Если это

свойство не выполняется, то говорят, что случайный член подвержен автокорреляции.

Автокорреляция обычно встречается в регрессионном анализе при использовании данных временных рядов. Поэтому под автокорреляцией понимается в основном корреляционная зависимость между наблюдаемыми показателями во времени.

Существуют различные методы определения автокорреляции. Наиболее распространенными являются следующие два:

- Построение графика остатков в зависимости от их порядковых номеров и визуальное определение наличия или отсутствия автокорреляции. Считается, что автокорреляции нет, если остатки расположены возле линии регрессии хаотично без видимой закономерности
- Использование критерия Дарбина–Уотсона:

1. Вычисляется величина

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

2. Затем по таблицам критических точек Дарбина–Уотсона определяются пороговые значения d_1 и d_2 в зависимости от уровня значимости, количества наблюдений и числа объясняющих переменных.
3. Далее наличие или отсутствие автокорреляции определяется тем, на какой участок отрезка $[0;4]$ попадает значение DW :
- (а) если $d_2 < DW < (4 - d_2)$, то признается отсутствие автокорреляции;
 - (б) если $0 < DW < d_1$, то имеется положительная автокорреляция;
 - (с) если $(4 - d_1) < DW < 4$, то существует отрицательная автокорреляция;

- (d) если $d_1 < DW < d_2$ или $(4 - d_2) < DW < (4 - d_1)$, то имеет место зона неопределенности, когда нельзя ни отклонить, ни применить нулевую гипотезу об отсутствии автокорреляции.

Действия по устранению автокорреляции необходимо начинать с проверки спецификации модели, поскольку всегда существует вероятность того, что обнаруженная автокорреляция связана с пропущенной переменной или использованием неправильной функциональной формы уравнения. Если ошибки спецификации устранены, то, возможно, это связано с внутренними свойствами ряда отклонений. Тогда для устранения автокорреляции можно воспользоваться обобщенным методом наименьших квадратов (ОМНК).

2.4.4 Мультиколлинеарность

Мультиколлинеарность – это линейная зависимость между двумя или несколькими факторными переменными в уравнении множественной регрессии. Если такая зависимость является функциональной, то говорят о полной мультиколлинеарности.

Для оценки наличия мультиколлинеарности в модели используется определитель матрицы парных коэффициентов корреляции между факторами (определитель матрицы межфакторной корреляции). Чем ближе определитель матрицы межфакторной корреляции к 0, тем выше мультиколлинеарность, и наоборот, чем ближе определитель к 1, тем ниже мультиколлинеарность.

Статистическая значимость мультиколлинеарности факторов определяется по следующему алгоритму:

1. Задается нулевая $H_0 : \Delta r = 1$ и альтернативная $H_0 : \Delta r \neq 1$ гипотезы.
2. Задается уровень значимости α .
3. Вычисляется значение статистики: $\chi_{obs}^2 = n - 1 - \frac{1}{6}(2m + 5)lg\Delta r$, где n – число наблюдений, m – число факторов.

4. Для заданного уровня значимости по таблице критических точек распределения Пирсона определяется критическое значение χ_{cr}^2 .
5. Если $\chi_{obs}^2 > \chi_{cr}^2$, то гипотеза отклоняется и считается, что в модели присутствует мультиколлинеарность факторов.

Простейшим методом устранения мультиколлинеарности является исключение из модели одной или ряда коррелированных переменных. В ряде случаев минимизировать либо совсем устранить мультиколлинеарность можно с помощью преобразования факторных переменных.

2.5 Верификация модели

Проверка качества оцененной множественной регрессионной модели проводится по следующим направлениям:

1. Оценка тесноты связи рассматриваемого набора факторов с исследуемым признаком с помощью коэффициента множественной корреляции.

$$R = \sqrt{1 - \frac{SSres}{SStot}}, \text{ где}$$

где $SSres = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ - сумма квадратов остатков регрессии, $SStot = \sum_{i=1}^n (y_i - \bar{y}_i)^2$ - общая сумма квадратов.

Чем ближе к 1 индекс множественной корреляции, тем теснее связь результативного признака со всем набором исследуемых факторов.

2. Проверка общего качества уравнения регрессии.

Одной из наиболее эффективных оценок общего качества множественной модели и характеристической силы является скорректированный коэффициент детерминации $\bar{R}^2 = 1 - \frac{n-1}{n-m-1}(1 - R^2)$, где m - число объясняющих переменных в модели, n - число наблюдений, а R^2 - коэффициент детерминации (квадрат коэффициента множественной корреляции).

В отличие от R^2 скорректированный коэффициент детерминации \bar{R}^2 может уменьшаться при введении в модель новых объясняющих переменных, не оказывающих существенного влияния на зависимую переменную.

Общее качество множественной модели также может быть оценено с помощью стандартной ошибки регрессии $s = \sqrt{\frac{1}{n-(m+1)} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$.

Величина стандартной ошибки регрессии характеризует среднюю величину рассеивания наблюдаемых значений переменной y относительно теоретических.

Для оценки адекватности уравнения регрессии может быть применена средняя ошибка аппроксимации:

$$\bar{A} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} * 100\%$$

Ошибка аппроксимации не более 8–12% свидетельствует о хорошем качестве модели.

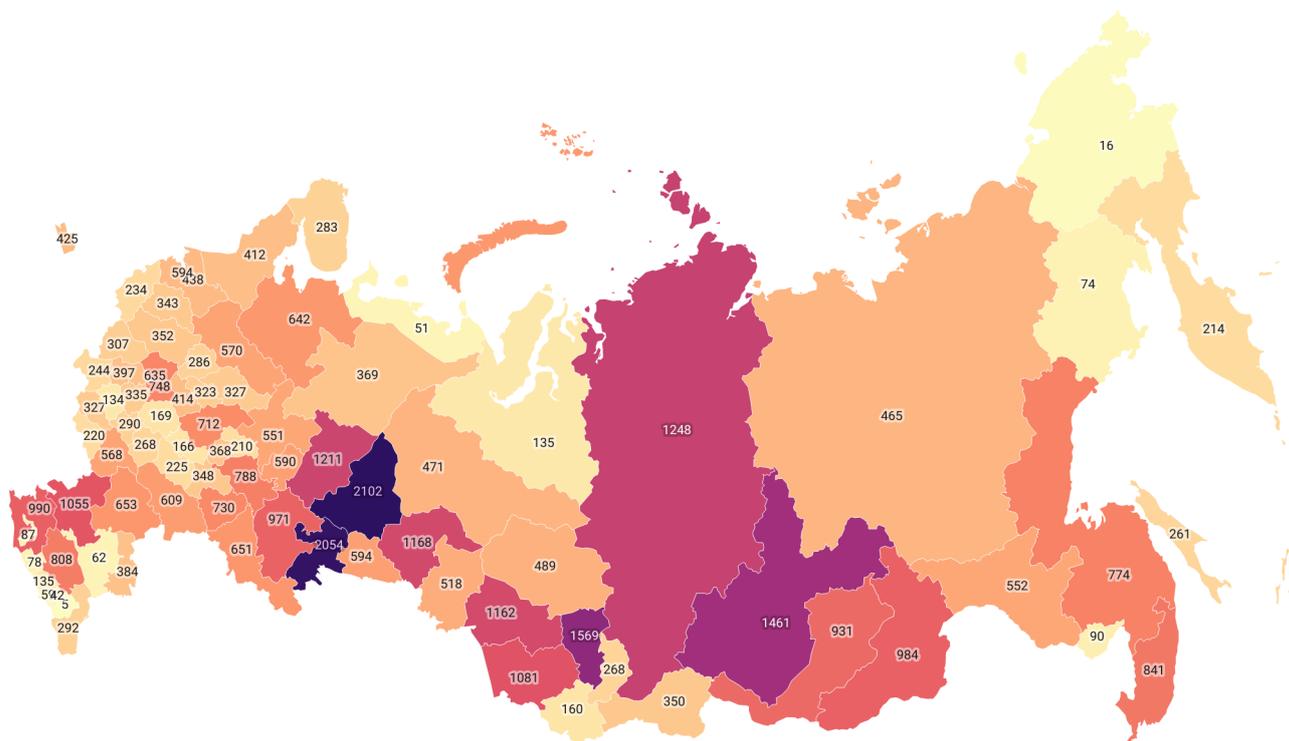
3. Проверка выполнимости предпосылок МНК (Теорема Гаусса-Маркова).

Глава 3. Визуализация данных

Визуализация данных была проведена посредством сайта app.datawrapper.de

На карте представлены данные по уровню подростковой преступности в регионах России за 2018 год.

Teenage Crime 2018



Created with Datawrapper

На карте видно, что наибольший уровень подростковой преступности наблюдается в следующих регионах: Свердловская область, Челябинская область, Кемеровская область, Иркутская область. Сделаем предположение, что данные наблюдения будут удалены как "выбросы" при предварительной обработке данных в Главе 4.2.

Глава 4. Построение множественных регрессионных моделей оценки уровня подростковой преступности

4.1 Работа с данными. Построение модели I

Для построения множественной регрессионной модели была собрана информация о числе подростковых преступлений для каждого региона России за 2018 год (CRIME) в зависимости от следующих факторов:

- Численность населения в регионе (POP).
- Число разводов в регионе (DIV).
- Численность подростков (POPTEEN).
- Уровень занятости населения в возрасте от 14 до 18 лет (EMPRATE).
- Уровень безработицы в регионе (UNEMP).
- Число спортивных центров (SPORT).
- Уровень жизни населения в регионе (STLIV).
- Число организаций начального и среднего образования (SCIENCE).
- Число молодежных центров (YCENTRE).

Все переменные оказались количественными (Приложение 1).

Для оценки коллинеарности факторов, включенных в модель, рассчитаем коэффициент парной корреляции и построим корреляционную матрицу.

	CRIME	POP	DIV	POPTEEN	EMPRATE	UNEMP	SPORT	STLIV	SCIENCE	YCENTRE
CRIME	1									
POP	0,483566387	1								
DIV	0,215751304	-0,0368204	1							
POPTEEN	0,572396076	0,96688302	-0,0956586	1						
EMPRATE	0,253854331	0,86728751	-0,0243533	0,79452472	1					
UNEMP	-0,088656597	-0,0764811	-0,0101693	-0,0265595	-0,0904917	1				
SPORT	0,483452995	0,86416152	-0,0147007	0,84133081	0,80829327	-0,1432601	1			
STLIV	0,206941037	0,6397085	0,32631733	0,58037853	0,53894795	-0,1492032	0,55608386	1		
SCIENCE	0,657534328	0,72779107	-0,1044754	0,80649915	0,41955696	0,0400615	0,64447621	0,44158028	1	
YCENTRE	0,708611147	0,81125617	0,05058984	0,82247123	0,48911872	-0,0463525	0,7074083	0,51833474	0,85239337	1

Из расчетов видно, что между факторами POP и POPTEEN существует тесная корреляционная связь (коэффициент парной корреляции между переменными равен 0,966). Это означает, что факторы коллинеарны и один из них должен быть исключен из модели.

Коэффициент корреляции между переменными POP и CRIME равен 0,4835, а между переменными POPTEEN и CRIME - 0,5723. Следовательно, фактор POP должен быть исключен из регрессионной модели.

Далее, еще раз строим корреляционную матрицу и продолжаем аналогичный процесс. В конечном итоге остаются только 4 фактора, между которыми нет тесной корреляционной связи: POPTEEN, DIV, UNEMP, STLIV.

	<i>CRIME</i>	<i>DIV</i>	<i>POPTEEN</i>	<i>UNEMP</i>	<i>STLIV</i>
<i>CRIME</i>	1				
<i>DIV</i>	0,2157513	1			
<i>POPTEEN</i>	0,57239608	-0,0956586	1		
<i>UNEMP</i>	-0,0886566	-0,0101693	-0,0265595	1	
<i>STLIV</i>	0,20694104	0,32631733	0,58037853	-0,1492032	1

Построим регрессионную модель:

<i>Регрессионная статистика</i>						
Множественный R	0,721336802					
R-квадрат	0,520326782					
Нормированный R	0,49603953					
Стандартная ошибка	294,037809					
Наблюдения	84					
<i>Дисперсионный анализ</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>	
Регрессия	4	7409077,82	1852269,455	21,423864	5,38086E-12	
Остаток	79	6830200,418	86458,23314			
Итого	83	14239278,24				
	<i>Коэффициенты</i>	<i>стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
У-пересечение	179,2811818	195,3696321	0,917651223	0,3615946	-209,592344	568,1547071
DIV	212,4354117	41,51285814	5,117340054	2,126E-06	129,8061345	295,064689
POPTEEN	0,009310502	0,00108307	8,596398447	5,977E-13	0,007154705	0,011466299
UNEMP	-28,17655238	16,95825042	-1,661524726	0,1005718	-61,9311055	5,578000733
STLIV	-17,98689503	4,15222204	-4,331872154	4,305E-05	-26,2516863	-9,7221038

Рисунок 2 показывает, что свободный член имеет P-значение 0,76, то есть с вероятностью 76% является статистически незначимым. Проведем анализ данных, учитывая нулевое значение свободного члена.

Результат будет следующим:

Регрессионная статистика						
Множ R	0,893107946					
R-квадрат	0,797641804					
Нормир R-квадрат	0,777553371					
Станд ошибка	293,7474573					
Наблюдения	84					
Дисперсионный анализ						
	df	SS	MS	F	Значимость F	
Регрессия	4	27209798,51	6802449,627	78,834642	8,56003E-27	
Остаток	80	6903005,493	86287,56866			
Итого	84	34112804				
	Коэффициенты	Станд ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
Y-пересечение	0	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д
DIV	229,4667396	37,09627038	6,185709163	2,473E-08	155,6428088	303,2906703
POPTEEN	0,009190502	0,001074085	8,556587158	6,536E-13	0,007053005	0,011327999
UNEMP	-21,5787235	15,34346224	-1,406379027	0,1634842	-52,1131865	8,955739467
STLIV	-16,13465639	3,625023106	-4,450911324	2,738E-05	-23,3486823	-8,920630505

Анализ качества параметров модели выявил незначимый фактор - уровень безработицы в регионе UNEMP. Результат регрессионного анализа после удаления данного фактора:

Регрессионная статистика						
Множественный R	0,890302614					
R-квадрат	0,792638745					
Нормированный F	0,775173035					
Стандартная ошиб	295,515324					
Наблюдения	84					
Дисперсионный анализ						
	df	SS	MS	F	Значимость F	
Регрессия	3	27039130,16	9013043,385	103,2075454	2,0268E-27	
Остаток	81	7073673,845	87329,30673			
Итого	84	34112804				
	Коэффициенты	стандартная ошиб	t-статистика	P-Значение	Нижние 95%	Верхние 95%
Y-пересечение	0	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д
DIV	212,4752039	35,28478321	6,021723377	4,82067E-08	142,269553	282,6808544
POPTEEN	0,009113916	0,00107916	8,445383826	9,92735E-13	0,00696673	0,011261105
STLIV	-16,30597045	3,644780127	-4,47378714	2,48212E-05	-23,55794	-9,054001281

Рисунок 3 показывает, что полученное уравнение регрессии имеет вид:

$$Y_I = 212,5DIV + 0,009POPTEEN - 16,3STLIV$$

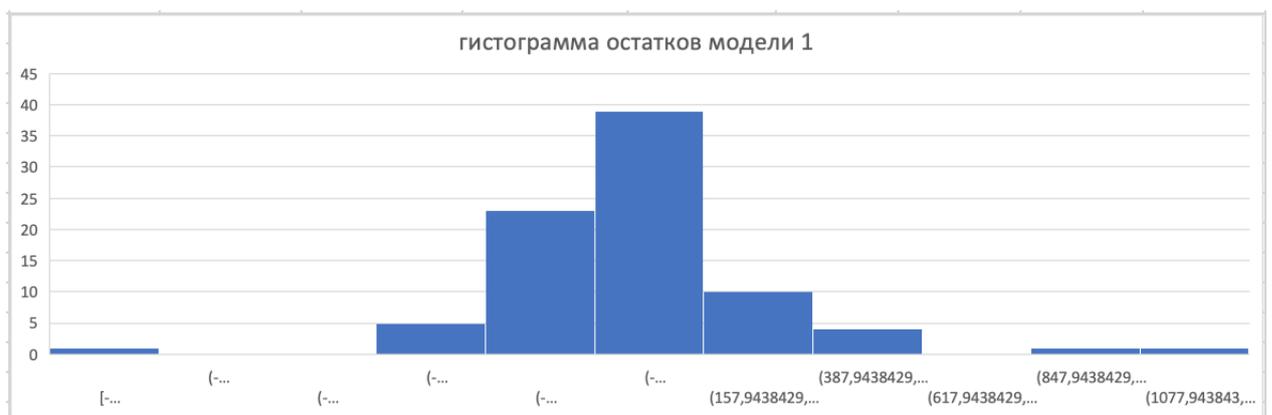
Можно сделать вывод о том, что уровень подростковой преступности в России в большей степени зависит от числа разводов, численности подростков и уровня жизни.

Верификация модели:

1. Коэффициент множественной корреляции $R = 0,71$ говорит о том, что исследуемый признак достаточно тесно связан с рассматриваемым набором факторов.
2. Коэффициент детерминации $R^2 = 0,5$ и исправленный коэффициент детерминации $\bar{R}^2 = 0,48$ говорят о том, что построенное уравнение регрессии объясняет 50% общей вариации результативного признака, а другие 50% приходится на другие факторы или причины.
3. Оценка адекватности уравнения регрессии $\bar{A} = 66\%$ (что намного выше 12%) говорит о том, что в среднем расчетных значений отклоняются от фактических 66%.

2 и 3 пункты говорят о построении некачественной модели.

Посмотрим на гистограмму остатков:



Анализируя график, был сделан вывод о том, что в модели есть выбросы, которые влияют на ее качество.

4.2 Работа с выбросами

Так как регрессионная модель имеет большую чувствительность к выбросам, было принято решение удалить все наблюдения, которые выходят за 1.5 межквартильных размаха. Если подтвердится, что резко выделяющееся значение действительно не может относиться к объектам данной группы и попало в записи вследствие ошибок внимания, следует такой

выброс исключить из обработки. Построим таблицу базовых данных для каждого фактора:

Q1	212,5	3,8	17197,25	3,7	2,4	657,5	39,875	202,75	31
Q3	619,25	4,425	59428	15,075	4,925	2297,75	53	700,25	66,5
IQR	406,75	0,625	42230,75	11,375	2,525	1640,25	13,125	497,5	35,5
L BOUND	-397,625	2,8625	-46148,875	-13,3625	-1,3875	-1802,875	20,1875	-543,5	-22,25
R BOUND	1229,375	5,3625	122774,125	32,1375	8,7125	4758,125	72,6875	1446,5	119,75

Q1, Q3 - первый и третий квартиль соответственно, IQR - межквартильный диапазон (средние 50% значений данных), L BOUND, R BOUND - нижняя и верхняя границы соответственно (минимальное и максимальное значения диапазона данных) соответственно. В результирующей таблице будем выводить ЛОЖЬ, если значение входит в диапазон (а значит не является выбросом), и ИСТИНА, если значение не входит в диапазон (а значит является выбросом). Уберем все выбросы из выборки (Приложение 2). Тем самым остается 60 данных для исследования регрессии (Приложение 3)

4.3 Модель II

Снова составим корреляционную матрицу:

	CRIME	DIV	POPTeen	EMPRATE	UNEMP	SPORT	STLIV	SCIENCE	YCENTRE
CRIME	1								
DIV	0,1456665	1							
POPTeen	0,8393559	0,0502074	1						
EMPRATE	0,5859362	-0,0729521	0,8177121	1					
UNEMP	-0,0577538	0,1324026	-0,0869817	-0,0431636	1				
SPORT	0,4574371	-0,0779063	0,6975037	0,7269123	-0,09633	1			
STLIV	0,0377302	0,2933953	0,3696677	0,3919006	-0,1985749	0,4731152	1		
SCIENCE	0,7273253	-0,0584315	0,8568913	0,8263263	-0,0321199	0,6909724	0,2579166	1	
YCENTRE	0,5765694	-0,0224257	0,7426074	0,7196474	-0,0687708	0,6647631	0,3092834	0,8271196	1

Выбедем те факторы, которые хорошо коррелируют с результирующим и при этом коррелируют друг с другом с коэффициентом корреляции менее 0,7. По таблице видно, что лучше всего с результирующим фактором коррелируют POPTeen (0.839) и SCIENCE (0.727), но при этом их нельзя рассматривать в одной группе факторов, потому что коэффициент корреляции между ними выше 0.7 (а именно 0.74). Поэтому рассмотрим 2 группы факторов:

- POPTeen, DIV, UNEMP, SPORT, STLIV.

- SCIENCE, SPORT, DIV, UNEMP, STLIV.

Рассмотрим сначала первую группу факторов и построим модель:

<i>Регрессионная статистика</i>						
Множественный R	0,915201509					
R-квадрат	0,837593802					
Нормир R-квадрат	0,822556191					
Стандарт ошибка	120,6513954					
Наблюдения	60					
<i>Дисперсионный анализ</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>	
Регрессия	5	4054051,986	810810,3972	55,6999251	4,24788E-20	
Остаток	54	786064,9972	14556,75921			
Итого	59	4840116,983				
	<i>Коэффициенты</i>	<i>Станд ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	82,08408181	141,9127784	0,578412196	0,56539051	-202,433908	366,602072
DIV	123,3272173	34,7873542	3,545173818	0,00081936	53,58277138	193,0716632
POPTEEN	0,01468182	0,00114173	12,85927058	4,5283E-18	0,012392788	0,016970852
UNEMP	-14,2942139	10,04770606	-1,422634562	0,16059345	-34,4386517	5,850223861
SPORT	-0,013819109	0,026651009	-0,51852104	0,60621314	-0,06725117	0,039612947
STLIV	-12,55983218	2,231148711	-5,629311984	6,6529E-07	-17,033016	-8,08664834

Полученное уравнение регрессии имеет вид:

$$Y_{II} = 82,08 + 123,33DIV + 0,01POPTEEN - \\ -14,29UNEMP - 0,01SPORT - 12,56STLIV$$

Исправленный коэффициент детерминации $\bar{R}^2 = 0,82$, то есть построенное уравнение регрессии объясняет 82% общей вариации результативного признака, а другие 18% приходятся на другие факторы, не учтенные в модели. Значения t-статистик показывают, что не все переменные, включенные в модель, являются статистически значимыми при уровне значимости 0,05. Так, например, переменная SPORT имеет P-значение 0,606, т.е. с вероятностью 60,6% является статистически незначимой. Удалим ее из модели. Параметры регрессии после исключения фактора SPORT представлены на следующем рисунке:

<i>Регрессионная статистика</i>					
Множественный R	0,914759634				
R-квадрат	0,837851873				
Нормиров R-квадрат	0,824915019				
Стандартная ошибка	119,84678				
Наблюдения	60				
Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	4	4050138,197	1012534,549	70,4948046	5,38653E-21
Остаток	55	789978,7867	14363,25067		
Итого	59	4840116,983			
	<i>Коэффициенты</i>	<i>Станд ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>
Y-пересечение	75,53736268	140,407363	0,53798719	0,592756359	-205,8452806
DIV	128,5277951	33,08801153	3,884421854	0,000277662	62,21793818
POPTEN	0,014299638	0,00086613	16,50980403	5,71677E-23	0,012563875
UNEMP	-14,65852096	9,956268187	-1,47229069	0,14664201	-34,61132828
STLIV	-13,01876627	2,034426636	-6,399231137	3,61983E-08	-17,09584836

Рисунок показывает, что уравнение регрессии имеет вид:

$$Y_{II} = 75,54 + 128,53DIV + 0,01POPTEN - 14,66UNEMP - 13,02STLIV$$

P-значение для свободного члена составил 0,54, что говорит о незначимости данного фактора. Рассмотрим модель с константой-ноль:

<i>Регрессионная статистика</i>					
Множественный R	0,972515798	SSrez	794135,9455		
R-квадрат	0,945786977	SStot	4840116,983		
Нормиров R-квадрат	0,925025565	R-квадрат	0,835926291		
Стандартная ошибка	119,0839999				
Наблюдения	60				
Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	4	13854299,05	3463574,764	244,2405332	2,60697E-34
Остаток	56	794135,9455	14180,99903		
Итого	60	14648435			
	<i>Коэффициенты</i>	<i>Станд ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>
Y-пересечение	0	#Н/Д	#Н/Д	#Н/Д	#Н/Д
DIV	141,5290966	22,45690061	6,302254219	4,87865E-08	96,54251892
POPTEN	0,014305797	0,000860542	16,62416521	2,56203E-23	0,012581924
UNEMP	-13,55007354	9,678752805	-1,399981363	0,167034815	-32,93894527
STLIV	-12,66455136	1,912687702	-6,621337787	1,45879E-08	-16,49612524

После исключения свободного члена регрессионная модель приобрела

ВИД:

$$Y_{II} = 141,53DIV + 0,01POPTEEN - 13,55UNEMP - 12,66STLIV$$

Исправленный коэффициент детерминации после удаления свободного члена хоть и незначительно, но увеличился и составил 0,821.

Анализ качества параметров выявил еще один статистически незначимый фактор - UNEMP. Исключим его из уравнения регрессии:

Регрессионная статистика		SSres	821929,9635		
Множественный R	0,971539794	SStot	4840116,983		
R-квадрат	0,943889572	R - квадрат	0,830183864		
Нормиров R-квадрат	0,924376925				
Стандартная ошибка	120,0825713				
Наблюдения	60				
Дисперсионный анализ					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>
Регрессия	3	13826505,04	4608835,012	319,6179813	3,46624E-35
Остаток	57	821929,9635	14419,82392		
Итого	60	14648435			
	<i>Коэффициенты</i>	<i>Станд ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>
Y-пересечение	0	#Н/Д	#Н/Д	#Н/Д	#Н/Д
DIV	124,8691607	19,20437448	6,502120696	2,13902E-08	86,41306419
POPTEEN	0,014304398	0,000867758	16,48432164	2,36767E-23	0,012566743
STLIV	-12,18409148	1,897425246	-6,421381558	2,91003E-08	-15,98362

После исключения фактора UNEMP регрессионная модель приобрела вид:

$$Y_{II} = 124,87DIV + 0,01POPTEEN - 12,18STLIV$$

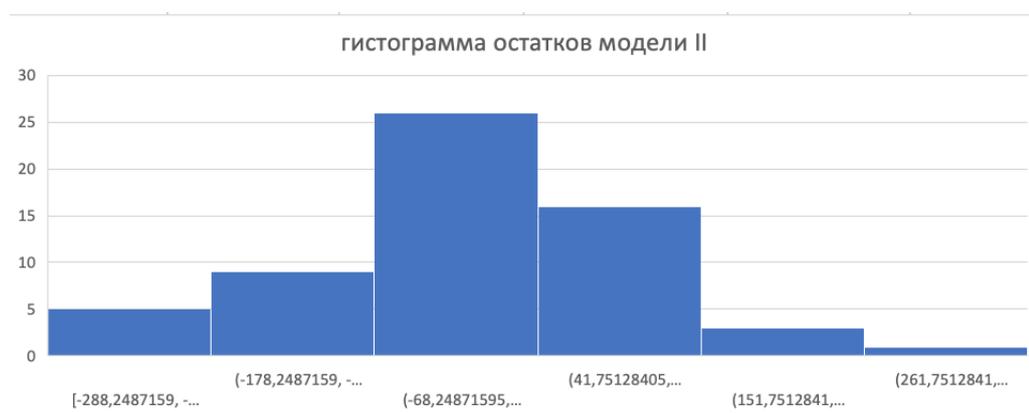
Коэффициент детерминации после удаления UNEMP члена хоть и незначительно, но увеличился и составил 0,818.

Верификация модели:

1. Коэффициент множественной корреляции $R = 0,71$ говорит о том, что исследуемый признак достаточно тесно связан с рассматриваемым набором факторов.

- Исправленный коэффициент детерминации $\bar{R}^2 = 0,82$ говорят о том, что построенное уравнение регрессии объясняет 82% общей вариации результативного признака, а другие 18% приходится на другие факторы или причины.
- Оценка адекватности уравнения регрессии $\bar{A} = 12\%$ говорит о том, что в среднем расчетных значений отклоняются от фактических 12%.

Гистограмма остатков выглядит следующим образом:



Из рисунка видно, что в модели была проведена предварительная обработка данных (удалены выбросы). Так же видно ярковыраженное нормальное распределение остатков. Верификация модели и гистограмма остатком говорят о построении качественной модели.

4.4 Оценка параметров модели II. Построение модели III

Проверим выполнимость предпосылком МНК для остатков модели:

- Первая предпосылка - $E(\varepsilon_i) = 0$ Как было написано в главе 1: "Для проверки предположения о равенстве нулю математического ожидания случайной величины достаточно оценить величину среднего выборочного остатков регрессии $e_i = y_i - \tilde{y}_i, i = 1, 2, \dots, n$ "

$$\bar{e} = \frac{1}{n} \sum_{i=1}^n e_i = 0.42$$

Среднее значение близко к нулю, но для убедительности проверим гипотезу о равенстве нулю математического ожидания случайной величины e_i с помощью t -статистики.

(a) Зададим нулевую $H_0 : E(e_i) = 0$ и альтернативную $H_1 : E(e_i) \neq 0$ гипотезы.

(b) Уровень значимости $\alpha = 0,05$.

(c) Вычислим t -статистику:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2} = 117,04, \quad \tau = \frac{\bar{e}\sqrt{n}}{s} = 0,03$$

(d) Найдем критическое значение статистики Стьюдента с числом степеней свободы, равным $n - 1$ (Приложение 4):

$$t_{cr} = 2,001$$

(e) $\tau < t_{cr}$. Это значит, что нулевая гипотеза $H_0 : E(e_i) = 0$ принимается.

2. Проверим вторую предпосылку - $D(e_i) = \sigma^2 = const$ с помощью теста Голдфелда-Квандта. Проведем тест для каждого фактора.

В процессе теста было выявлено, что в модели присутствует гетероскедастичность, зависящая от переменной PORTEEN:

(a) Выборка, содержащая результирующую CRIME и фактор PORTEEN, была упорядочена по возрастанию значений наблюдений фактора PORTEEN.

(b) Затем разделена на три равные части (по 20 наблюдений в каждой).

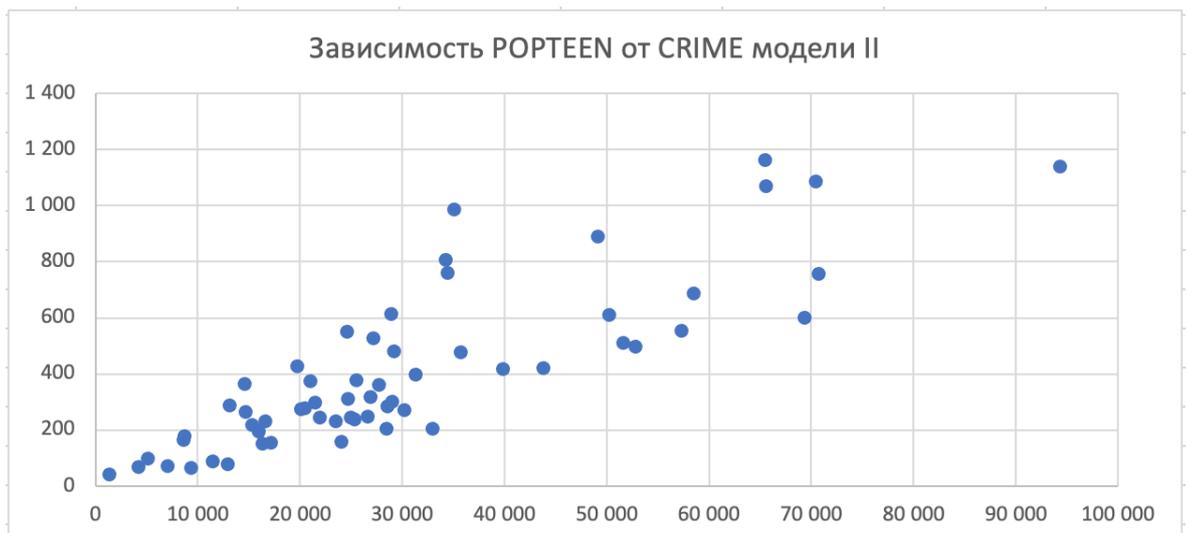
(c) Для первой и третьей частей были построены две независимые модели регрессии.

(d) По каждой из построенных моделей были рассчитаны суммы квадратов остатков $S_1 = 71921,46$ и $S_3 = 420668,14$.

(e) Вычислено наблюдаемое значение F -критерия: $F_{obs} = 5,85$.

- (f) Найдено критическое значение $F_{cr} = F_{cr}(\alpha, k_1 = 20 - 3 - 1, k_2 = 20 - 3 - 1) = F_{cr}(0.05, 16, 16) = 2,4$ (Приложение 5).
- (g) Так как $F_{obs} > F_{cr}$, то был слан вывод о том, что в модели присутствует гетероскедастичность.

Для устранения гетероскедастичности был применен взвешенный метод наименьших квадратов. Все наблюдаемые значения были разделены на корень из значения фактора РОРТЕЕН. И построена новая регрессионная модель с новыми переменными. В полученной модели выполнены 1 (нулевое математическое ожидание случайной величины) и 2 (гомоскедастичность) условия. Сравнение модели II с исправленной моделью II (модель III):



На первом графике явно выражена гетероскедастичность.

3. Проверим модель III на отсутствии автокорреляции двумя способами:

- Построение графика остатков в зависимости от их порядковых номеров и визуальное определение наличия или отсутствия автокорреляции.



Остатки расположены хаотично, без видимой закономерности, из чего можно сделать вывод, что автокорреляции нет.

- Для статистического анализа автокорреляции воспользуемся критерием Дарбина-Уотсона.
 - Рассчитаем величину $DW = 1,598$.
 - Определим пороговые значения (при уровне значимости $\alpha = 0,05$, количестве наблюдений $n = 60$ и числе объясняющих переменных $k = 3$) $d_1 = 1,48$ и $d_2 = 1,69$ (Приложение 6).
 - $d_2 < DW < 4 - d_2$, значит автокорреляция отсутствует.

4. Определитель матрицы межфакторной корреляции $\Delta r = 0,78$, что достаточно близко к 1, поэтому можно сделать вывод о ботсутствии мультиколлинеарности в модели.

В итоге, модель III имеет уравнение регрессии

$$Y_{III} = 92,71DIV + 0,014POPTTEEN - 9,04STLIV,$$

коэффициенты которой являются несмещенными, эффективными и состоятельными. Исправленный коэффициент детерминации \bar{R}^2 составил 0,85, средняя ошибка аппроксимации $\bar{A} = 12\%$.

Учитывая все вышесказанное, можно сделать вывод о том, что Модель III является классической нормальной линейной качественной моделью, оценки параметров регрессии которой являются несмещенными, состоятельными и эффективными.

Итог: Модели II и III также показывают, что уровень подростковой преступности в России в большей степени зависит от числа разводов, численности подростков и уровня жизни.

Другим методом избавления от гетероскедастичности является удаление фактора, из-за которого гетероскедастичность присутствует. Поэтому далее будет рассмотрена модель с факторами SCIENCE, SPORT, DIV, UNEMP, STLIV (модель IV).

4.5 Модель IV

Рассмотрим модель, включающую следующие факторы: SCIENCE, SPORT, DIV, UNEMP, STLIV. Построим множественную регрессию. Её параметры представлены на следующем рисунке:

<i>Регрессионная статистика</i>						
Множественный R	0,796777934					
R-квадрат	0,634855076					
Нормир R-квадрат	0,60104536					
Стандарт ошибка	180,9104623					
Наблюдения	60					
<i>Дисперсионный анализ</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>	
Регрессия	5	3072772,833	614554,5667	18,77729734	9,0517E-11	
Остаток	54	1767344,15	32728,59537			
Итого	59	4840116,983				
	<i>Коэффициенты</i>	<i>Станд ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	-189,6514029	214,3972621	-0,884579407	0,380305324	-619,4920331	240,1892273
DIV	174,8406997	51,6786256	3,38323045	0,001339633	71,23129362	278,4501058
UNEMP	-22,49755875	15,05354721	-1,494502155	0,140865491	-52,67810377	7,682986276
SPORT	0,030874995	0,040581725	0,760810314	0,450079898	-0,050486465	0,112236455
STLIV	-10,29569735	3,364335683	-3,060246754	0,003440716	-17,04078428	-3,550610425
SCIENCE	0,937785383	0,142079668	6,600419327	1,83353E-08	0,652932801	1,222637966

Модель выявила 2 явно незначимых параметра - SPORT и свободный член. Удалим их (поочередно) из модели и построим регрессию снова:

<i>Регрессионная статистика</i>						
Множественный R	0,936064711					
R-квадрат	0,876217143					
Нормир R-квадрат	0,851728776					
Стандарт ошибка	179,9417133					
Наблюдения	60					
<i>Дисперсионный анализ</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>	
Регрессия	4	12835209,87	3208802,467	99,10128379	1,81937E-24	
Остаток	56	1813225,13	32379,02018			
Итого	60	14648435				
	<i>Коэффициенты</i>	<i>Станд ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>
Y-пересечение	0	#Н/Д	#Н/Д	#Н/Д	#Н/Д	#Н/Д
DIV	132,5372474	33,8966383	3,910041054	0,000251667	64,63412128	
UNEMP	-24,90041386	14,67577752	-1,696701509	0,045305512	-54,29952897	
STLIV	-9,915089743	2,849794014	-3,479230321	0,000980855	-15,62391315	
SCIENCE	0,994635376	0,105098333	9,463854955	3,23972E-13	0,784098117	

Все факторы оказались значимыми. В итоге модель имеет вид:

$$Y_{IV} = 132,54DIV - 24,9UNEMP - 9,92STLIV + 0,99SCIENCE$$

Исправленный коэффициент детерминации R^2 составил 0,59, средняя ошибка аппроксимации $\bar{A} = 49,5\%$ - эти значения говорят о построении некачественной модели. Оценим параметры Модели IV.

4.6 Оценка параметров модели IV. Построение модели V

1. Проверим гипотезу о равенстве нулю математического ожидания $E(e_i)$ с помощью t -статистики:

(а) Зададим нулевую $H_0 : E(e_i) = 0$ и альтернативную $H_1 : E(e_i) \neq 0$ гипотезы.

(б) Уровень значимости $\alpha = 0,05$.

(с) Вычислим t -статистику:

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2} = 173,82, \quad \tau = \frac{\bar{e}\sqrt{n}}{s} = -0,1.$$

(d) Найдем критическое значение статистики Стьюдента с числом степеней свободы, равным $n - 1$ (Приложение 4):

$$t_{cr} = 2,003.$$

(е) $\tau < t_{cr}$. Это значит, что нулевая гипотеза $H_0 : E(e_i) = 0$ принимается.

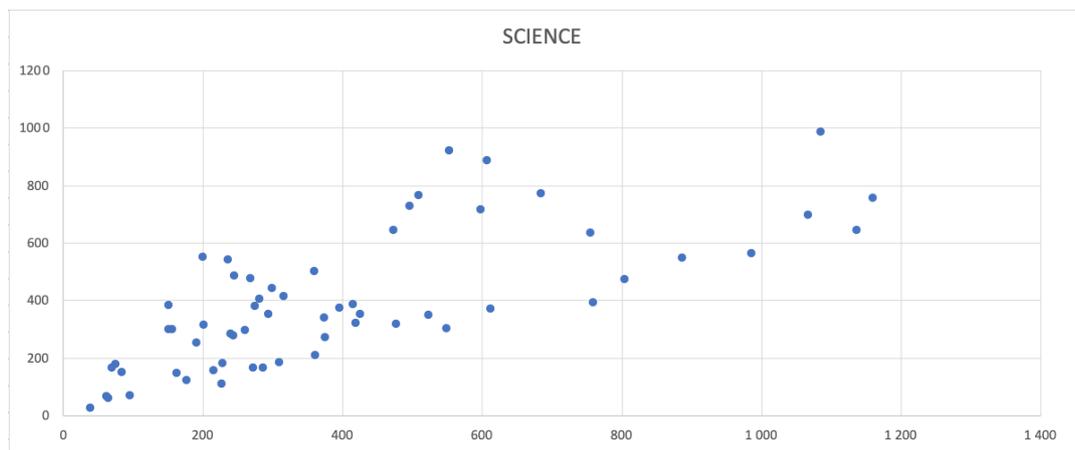
2. Проверим вторую предпосылку: $D(e_i) = \sigma^2 = const$ с помощью теста Голдфелда-Квандта. Проведем тест для каждого фактора.

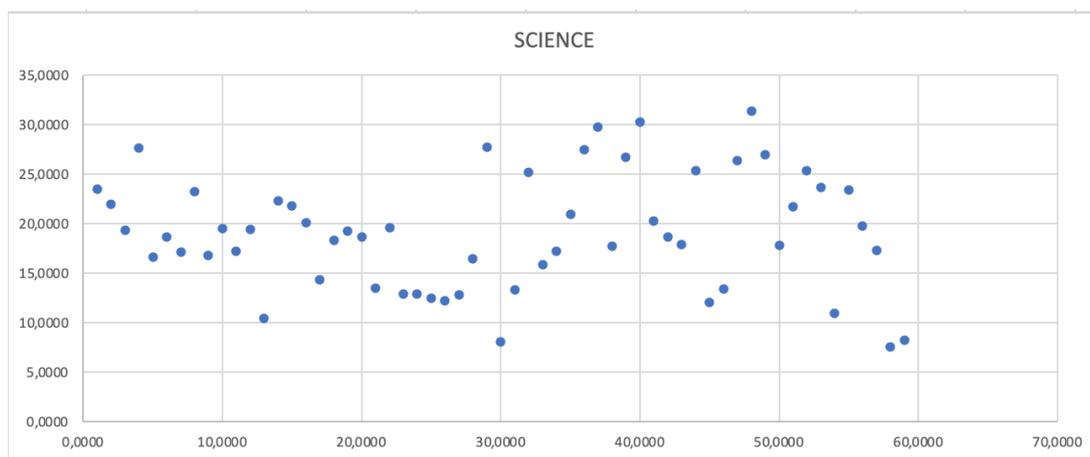
В процессе теста было выявлено, что в модели присутствует гетероскедастичность, зависящая от переменной SCIENCE:

(а) Выборка, содержащая результирующую CRIME и фактор SCIENCE, была упорядочена по возрастанию значений наблюдений фактора SCIENCE.

- (b) Затем разделена на три равные части (по 20 наблюдений в каждой).
- (c) Для первой и третьей частей были построены две независимые модели регрессии.
- (d) По каждой из построенных модели были рассчитаны суммы квадратов остатков $S_1 = 85328,71$ и $S_3 = 532451,14$.
- (e) Вычислено наблюдаемое значение F -критерия: $F_{obs} = 6,24$.
- (f) Найдено критическое значение $F_{cr} = F_{cr}(\alpha, k_1 = 20 - 3 - 1, k_2 = 20 - 3 - 1) = F_{cr}(0.05, 16, 16) = 2,4$ (Приложение 5).
- (g) Так как $F_{obs} > F_{cr}$, то был сделан вывод о том, что в модели присутствует гетероскедастичность.

Для устранения гетероскедастичности был применен взвешенный метод наименьших квадратов. Все наблюдаемые значения были разделены на корень из значения фактора SCIENCE (Приложение 7). В полученной модели выполнены 1 (нулевое математическое ожидание случайной величины) и 2 (гомоскедастичность) условия. Сравнение модели IV с исправленной моделью IV (модель V):





На первом графике явно выражена гетероскедастичность. На втором заметно, что она отсутствует.

3. Далее модель V не проверяется, потому что исправленный коэффициент детерминации $R^2 = 0,28$, оценки адекватности уравнения регрессии $\bar{A} = 44,6\%$, что говорит о построении некачественной модели.

Из полученных результатов Моделей IV и V можно сделать вывод о том, что уровень подростковой преступности сильно зависит от числа подростков.

Глава 5. Верификация построенных моделей

Проверим качество построенных в главе 3 моделей на данных 2019 года. Для начала сделаем предварительную обработку данных 2019 года (Приложение 7).

Далее, построим графики предсказанных и истинных значений для всех четырех моделей:

1. Модель I имеет уравнение регрессии вида:

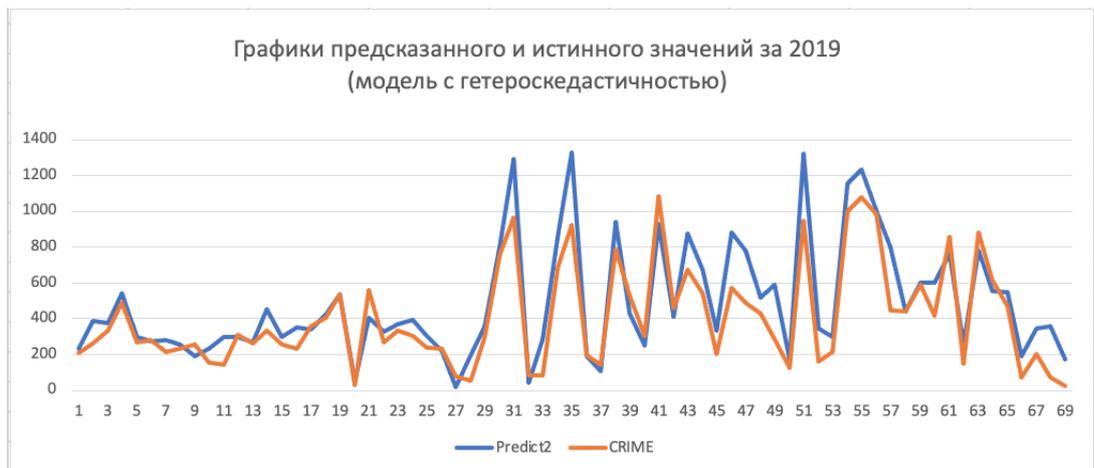
$$Y_I = 212,5DIV + 0,009POPTTEEN - 16,3STLIV$$



Видно, что модель не во всех местах повторяет динамику истинных значений. В большинстве случаев предсказанные значения выше истинных (это особенно заметно в последних пяти наблюдениях).

2. Модель II имеет уравнение регрессии вида:

$$Y_{II} = 124,87DIV + 0,01POPTTEEN - 12,18STLIV$$



Модель II точно повторяет динамику истинных значений, но в некоторых местах предсказанные значения выше истинных.

3. Модель III имеет уравнение регрессии вида:

$$Y_{III} = 92,71DIV + 0,014POPT EEN - 9,04STLIV$$



По графику видно, что это самая качественная модель. Она наиболее точно прогнозирует уровень подростковой преступности за 2019 год.

Глава 6. Заключение

Для статистического анализа уровня подростковой преступности в регионах России были собраны необходимые данные и собрана выборка. Проведена визуализация данных по числу подростковой преступности в регионах России за 2018 год. Для построения множественной регрессионной модели был изучен и представлен необходимый теоретический материал и собрана выборка из следующих факторы:

1. Число разводов в регионе.
2. Уровень занятости населения в возрасте от 14 до 18 лет.
3. Уровень жизни в регионе.
4. Число организаций начального и среднего образования.
5. Число молодежных и спортивных центров.
6. Уровень безработицы в регионе.
7. Численность населения в регионе.
8. Численность подростков в регионе.

По результатам анализа было выявлено, что наиболее значимыми факторами, влияющими на уровень подростковой преступности, являются: число разводов, уровень безработицы, уровень жизни в регионе.

Построены 5 моделей множественной регрессии. Первая модель была построена на основе необработанных данных, что сказалось на её качестве. Модель II была построена уже после отчистки данных от "выбросов". При верификации в Модели II была выявлена гетероскедастичность, что повлекло за собой замену переменных и построение Модели III, которая удовлетворяла всем условиям качественной модели. Далее были рассмотрены другие факторы (была исключена численность подростков и добавлены число организаций начального и среднего образования и число спортивных центров) и на основе них построена Модель IV, в которой также была

выявлена гетероскедастичность, что стало причиной построения Модели V. Последние две модели оказались некачественными, что доказало, что уровень подростковой преступности сильно зависит от числа подростков в регионе. В итоге Модель I, Модель II и Модель III верифицировались на новых данных (за 2019 год) и были представлены графически для визуального анализа и сравнения.

Самой качественной моделью оказалась Модель III:

$$Y_{III} = 92,71DIV + 0,014POPTTEEN - 9,04STLIV,$$

которая показала, что уровень подростковой преступности в регионах России в большей степени зависит от количества разводов, численности подростков и уровня жизни в регионе.

Список литературы

- [1] А.М.Великоцкая «Значимые факторы социальной ситуации развития подростка-правонарушителя »
- [2] Официальный сайт Правительства Камчатского края «Подростковая преступность, её причины и пути преодоления »
- [3] «Профилактика преступлений и правонарушений среди несовершеннолетних, воспитание правового сознания обучающихся »
- [4] «Психологические и социальные проблемы несовершеннолетних »
- [5] МВД по республике Северная Осетия-Алания «Факторы, влияющие на рост преступности »
- [6] Шапкина И.М. «Причины девиантного поведения несовершеннолетних»
- [7] С. Ф. Каморников, С.С. Каморников «Эконометрика (учебное пособие)».
- [8] В. М. БУРЕ, Е. М. ПАРИЛИНА, А. А. СЕДАКОВ «МЕТОДЫ ПРИКЛАДНОЙ СТАТИСТИКИ в R и Excel»
- [9] Воскобойников Ю.Е. «ЭКОНОМЕТРИКА в EXCEL »
- [10] Евсеев Е.А., Кузютин Д.В. «Эконометрика»(электронное пособие)
- [11] Айвазян С.А., Мхитарян В.С. «Прикладная статистика и основы эконометрики».
- [12] Благовещенский Ю.Н. «Основы математико-статистического анализа региональных данных».
- [13] Дубров А.М., Мхитарян В.С., Трошин Л.П. «Многомерные статистические методы: для экономистов и менеджеров».
- [14] Лялин В.С. Статистика: теория и практика в Excel

Глава 7. Приложения

7.1 Приложение 1

Data_2018

REG	CRIME	POP	DIV	POPTEN	EMPRATE	UNEMP	SPORT	STLIV	SCIENCE	YCENRE
Белгородская область	220	1548	4.4	33,044	13.1	4.4	3201	64.4	551	63
Брянская область	244	1200	4.5	26,711	6.1	2.6	1459	45.3	483	40
Владимирская область	414	1366	4	31,327	9	5.1	1804	47.6	373	47
Воронежская область	568	2328	4.1	51,678	23.4	2.7	4392	60.8	765	69
Ивановская область	323	1004	3.9	21,952	6.7	2	1040	45.4	277	28
Калужская область	397	1009	4.4	21,507	5.3	1	958	53	350	51
Костромская область	327	637	4	14,719	2.7	2.3	666	40.7	295	55
Курская область	327	1107	4.4	25,340	9.5	2.7	1139	55.4	542	47
Липецкая область	290	1144	4.1	24,974	5.6	1.8	2002	59.2	283	37
Московская область	748	7599	3.9	163,893	20.5	4.6	2066	72.4	1512	203
Орловская область	134	740	3.8	16,382	8.5	3.1	548	46.9	382	36
Рязанская область	169	1114	4.2	24,106	7.6	1.6	1291	49.9	297	59
Смоленская область	307	942	4.4	20,549	5.7	6.6	1259	47.2	379	48
Тамбовская область	268	1016	3.8	23,537	7.8	1.1	2845	46.9	109	36
Тверская область	352	1270	4.3	27,774	6.4	5.4	2135	41.5	499	64
Тульская область	335	1479	4.1	30,271	9.4	2.2	1296	54.8	474	38
Ярославская область	286	1260	4	28,573	8	3.4	1231	49.6	403	35
г. Москва	635	12615	3.5	250,219	213.2	2.8	12017	77.4	808	142
Республика Карелия	412	618	4	14,623	2.7	3.8	607	36.3	207	25
Республика Коми	369	830	4.6	21,059	3.7	4.7	877	39.9	337	41
Архангельская область	642	1144	4.3	28,949	5.4	3	1118	39.3	370	38
Ненецкий автономный округ	51	44	3.7	1,445	3	7.7	31	35.7	26	0
Вологодская область	570	1168	3.9	27,222	4.8	2.5	899	40.4	349	38
Калининградская область	425	1002	5.2	24,767	5.9	3.9	1090	59	183	39
Ленинградская область	438	1848	4.1	39,946	1.4	2.8	1679	58.1	385	59
Мурманская область	283	748	5.2	20,155	1.8	6.4	317	47.3	166	49
Новгородская область	343	600	4.4	13,152	2.1	2.4	895	40.5	166	30
Псковская область	234	630	4.3	15,377	3.1	3.3	509	41.6	155	28
г. Санкт-Петербург	594	5384	4.8	103,679	91.8	3.6	2824	75.7	767	63
Республика Адыгея	87	455	3.6	11,499	3.7	2.2	632	47.9	150	21
Республика Калмыкия	62	272	3.7	7,131	2.2	3.7	303	28.8	165	4
Краснодарский край	990	5648	3.8	132,272	30.2	2	5337	64	1243	175
Астраханская область	384	1014	4.1	25,570	7.9	4	705	41.8	271	24
Волгоградская область	653	2508	3.9	58,573	16.1	1.9	2477	47.3	770	63
Ростовская область	1055	4203	4.1	97,100	37	1.9	6002	54.1	1169	112
г. Севастополь	47	443	4.4	9,426	3.4	2	219	53.5	66	9
Республика Дагестан	292	3086	1.4	111,410	13.7	10.1	2094	39.1	1463	65
Республика Ингушетия	42	498	0.8	17,144	1.8	2.6	188	28.5	129	7
Кабардино-Балкарская Республика	135	866	2.4	24,319	4.1	3.6	943	34.1	266	29
Карачаево-Черкесская Республика	78	466	3.2	13,009	3.1	1.9	299	25.3	178	28
Республика Северная Осетия – Алания	59	699	2.5	20,609	5.2	0.5	502	39.8	190	24
Чеченская Республика	5	1457	0.5	61,993	7.7	6.8	783	38.8	527	39

Ставропольский край	808	2795	3.5	70,791	18.6	3.9	2788	53	633	92
Республика Башкортостан	971	4051	3.8	102,304	26.4	3	6474	50.2	1318	120
Республика Марий Эл	210	681	3.8	16,027	5.4	2.3	999	39.5	252	39
Республика Мордовия	166	795	3.3	17,215	6.8	3	1147	46.3	297	40
Республика Татарстан	788	3899	3.5	92,684	41.8	3.6	3809	66.1	1409	102
Удмуртская Республика	590	1507	2.6	37,800	12.8	2.9	1558	46.5	582	55
Чувашская Республика	368	1223	3.2	29,073	9	5.2	2306	45.9	440	44
Пермский край	1211	2611	4.1	65,510	15.6	2.7	2234	46.7	756	66
Кировская область	551	1272	3.9	27,917	7.9	9.2	1148	40.2	514	71
Нижегородская область	712	3215	3.9	70,553	22.5	5.6	2855	55.1	835	121
Оренбургская область	651	1963	4.3	50,285	12	7.2	2610	47.8	886	68
Пензенская область	225	1318	3.8	28,482	8.4	3.9	2295	48.5	315	49
Самарская область	730	3183	4.4	69,374	27.4	2.9	2464	54.2	713	109
Саратовская область	609	2441	4	57,345	20.2	2.3	1664	47.2	919	89
Ульяновская область	348	1238	3.7	26,957	9.4	2.2	1031	48.8	413	52
Курганская область	594	835	4	19,812	4.2	3.2	1338	28.9	350	41
Свердловская область	2102	4316	4.4	99,457	33.8	5.1	4136	56.8	1081	160
Ханты-Мансийский автономный округ – Югра	471	1664	5.3	43,857	6	5.1	1021	61	320	53
Ямало-Ненецкий автономный округ	135	541	5.4	15,360	0.1	7.4	279	57.9	131	31
Тюменская область	1168	1518	4.5	94,415	14.9	2.4	1903	56.3	643	28
Челябинская область	2054	3476	4.5	82,567	24.2	2.1	3576	52.4	841	130
Республика Алтай	160	219	4	8,691	0.7	4.1	175	28.8	146	14
Республика Тыва	350	324	1.9	15,377	1.4	3.8	418	16.2	176	31
Республика Хакасия	268	537	3.9	16,677	2.3	2.8	568	42.2	180	34
Алтайский край	1081	2333	4.3	65,640	13.2	3.7	2865	37.6	696	89
Красноярский край	1248	2874	4.8	79,507	19.3	3.9	2389	46.3	990	121
Иркутская область	1461	2398	4.6	73,555	18.6	3.7	1870	39.2	894	100
Кемеровская область	1569	2674	4.3	79,446	11.7	4.9	4484	43.3	674	106
Новосибирская область	1162	2793	4.5	70,485	27.5	2.5	2148	51.2	985	90
Омская область	518	1944	4.3	52,878	22.2	6.3	2708	41.7	727	66
Томская область	489	1077	3.7	29,224	16.4	7.3	808	44.5	318	29
Республика Бурятия	931	983	3.6	34,257	5.7	3.8	1055	30.5	472	47
Республика Саха (Якутия)	465	967	3.4	35,764	7.1	5.2	569	37.6	643	82
Забайкальский край	984	1066	4.5	35,111	5.7	3.9	685	27.2	561	58
Камчатский край	214	315	5.2	8,791	1.2	3.1	272	47.9	120	24
Приморский край	841	1902	4.8	49,223	11.9	4.3	1526	45.2	547	52
Хабаровский край	774	1321	5	34,454	10.9	2.2	1517	48.8	392	40
Амурская область	552	794	5.1	24,685	4.1	2.3	1220	43.8	301	34
Магаданская область	74	141	5.1	4,279	0.9	5.7	151	47.5	58	13
Сахалинская область	261	490	5.6	14,085	1.5	5	418	46	161	33
Еврейская автономная область	90	160	4.8	5,212	0.5	7	225	29.9	68	9
Чукотский автономный округ	16	50	4.8	1,848	0.1	9	10	38.9	41	6

7.2 Приложение 2

REG	CRIME	DIV	POPTEEN	EMPRATE	UNEMP	SPORT	STLIV	SCIENCE	YCENTRE
Белгородская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Брянская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Владимирская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Воронежская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ивановская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Калужская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Костромская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Курская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Липецкая область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Московская область	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ИСТИНА
Орловская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Рязанская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Смоленская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Тамбовская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Тверская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Тульская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ярославская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
г. Москва	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ИСТИНА	ЛОЖЬ	ИСТИНА	ИСТИНА	ЛОЖЬ	ИСТИНА
Республика Карелия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Коми	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Архангельская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ненецкий автономный округ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Вологодская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Калининградская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ленинградская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Мурманская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Новгородская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Псковская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
г. Санкт-Петербург	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ
Республика Адыгея	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Калмыкия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Краснодарский край	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ИСТИНА
Астраханская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Волгоградская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ростовская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
г. Севастополь	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Дагестан	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ
Республика Ингушетия	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Кабардино-Балкарская Республика	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Карачаево-Черкесская Республика	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Северная Осетия – Алания	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Чеченская Республика	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ставропольский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Башкортостан	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ИСТИНА
Республика Марий Эл	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Мордовия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Татарстан	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Удмуртская Республика	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Чувашская Республика	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Пермский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Кировская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Нижегородская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА
Оренбургская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Пензенская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Самарская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Саратовская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ульяновская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Курганская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Свердловская область	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА
Ханты-Мансийский автономный округ – Югра	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ямало-Ненецкий автономный округ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Тюменская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Челябинская область	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА
Республика Алтай	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Тыва	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ
Республика Хакасия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Алтайский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Красноярский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА
Иркутская область	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Кемеровская область	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Новосибирская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Омская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Томская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Бурятия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Саха (Якутия)	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Забайкальский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Камчатский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Приморский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Хабаровский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Амурская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Магаданская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Сахалинская область	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ерейская автономная область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Чукотский автономный округ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ

7.3 Приложение 3

REG	CRIME	DIV	POPTEEN	EMPRATE	UNEMP	SPORT	STLIV	SCIENCE	YCENRE
Белгородская область	201	4,4	33 044	13,1	4,4	3201	64,4	551	63
Брянская область	246	4,5	26 711	6,1	2,6	1459	45,3	483	40
Владимирская область	396	4	31 327	9	5,1	1804	47,6	373	47
Воронежская область	510	4,1	51 678	23,4	2,7	4392	60,8	765	69
Ивановская область	244	3,9	21 952	6,7	2	1040	45,4	277	28
Калужская область	295	4,4	21 507	5,3	1	958	53	350	51
Костромская область	261	4	14 719	2,7	2,3	666	40,7	295	55
Курская область	236	4,4	25 340	9,5	2,7	1139	55,4	542	47
Липецкая область	241	4,1	24 974	5,6	1,8	2002	59,2	283	37
Орловская область	151	3,8	16 382	8,5	3,1	548	46,9	382	36
Рязанская область	157	4,2	24 106	7,6	1,6	1291	49,9	297	59
Смоленская область	275	4,4	20 549	5,7	6,6	1259	47,2	379	48
Тамбовская область	228	3,8	23 537	7,8	1,1	2845	46,9	109	36
Тверская область	360	4,3	27 774	6,4	5,4	2135	41,5	499	64
Тюльская область	269	4,1	30 271	9,4	2,2	1296	54,8	474	38
Ярославская область	282	4	28 573	8	3,4	1231	49,6	403	35
Республика Карелия	362	4	14 623	2,7	3,8	607	36,3	207	25
Республика Коми	374	4,6	21 059	3,7	4,7	877	39,9	337	41
Архангельская область	613	4,3	28 949	5,4	3	1118	39,3	370	38
Ненецкий автономный округ	39	3,7	1 445	3	7,7	31	35,7	26	0
Вологодская область	524	3,9	27 222	4,8	2,5	899	40,4	349	38
Калининградская область	310	5,2	24 767	5,9	3,9	1090	59	183	39
Ленинградская область	416	4,1	39 946	1,4	2,8	1679	58,1	385	59
Мурманская область	273	5,2	20 155	1,8	6,4	317	47,3	166	49
Новгородская область	287	4,4	13 152	2,1	2,4	895	40,5	166	30
Псковская область	216	4,3	15 377	3,1	3,3	509	41,6	155	28
Республика Адыгея	85	3,6	11 499	3,7	2,2	632	47,9	150	21
Республика Калмыкия	70	3,7	7 131	2,2	3,7	303	28,8	165	4
Астраханская область	376	4,1	25 570	7,9	4	705	41,8	271	24
Волгоградская область	685	3,9	58 573	16,1	1,9	2477	47,3	770	63
г. Севастополь	63	4,4	9 426	3,4	2	219	53,5	66	9
Карачаево-Черкесская Республика	76	3,2	13 009	3,1	1,9	299	25,3	178	28
Ставропольский край	756	3,5	70 791	18,6	3,9	2788	53	633	92
Республика Марий Эл	191	3,8	16 027	5,4	2,3	999	39,5	252	39
Республика Мордовия	152	3,3	17 215	6,8	3	1147	46,3	297	40
Чувашская Республика	300	3,2	29 073	9	5,2	2306	45,9	440	44
Пермский край	1 160	4,1	65 510	15,6	2,7	2234	46,7	756	66
Оренбургская область	607	4,3	50 285	12	7,2	2610	47,8	886	68
Пензенская область	202	3,8	28 482	8,4	3,9	2295	48,5	315	49
Самарская область	599	4,4	69 374	27,4	2,9	2464	54,2	713	109
Саратовская область	553	4	57 345	20,2	2,3	1664	47,2	919	89
Ульяновская область	317	3,7	26 957	9,4	2,2	1031	48,8	413	52
Курганская область	426	4	19 812	4,2	3,2	1338	28,9	350	41
Ханты-Мансийский автономный округ	419	5,3	43 857	6	5,1	1021	61	320	53
Тюменская область	1 137	4,5	94 415	14,9	2,4	1903	56,3	643	28
Республика Алтай	163	4	8 691	0,7	4,1	175	28,8	146	14
Республика Хакасия	229	3,9	16 677	2,3	2,8	568	42,2	180	34
Алтайский край	1 067	4,3	65 640	13,2	3,7	2865	37,6	696	89
Новосибирская область	1 085	4,5	70 485	27,5	2,5	2148	51,2	985	90
Омская область	497	4,3	52 878	22,2	6,3	2708	41,7	727	66
Томская область	478	3,7	29 224	16,4	7,3	808	44,5	318	29
Республика Бурятия	805	3,6	34 257	5,7	3,8	1055	30,5	472	47
Республика Саха (Якутия)	474	3,4	35 764	7,1	5,2	569	37,6	643	82
Забайкальский край	986	4,5	35 111	5,7	3,9	685	27,2	561	58
Камчатский край	177	5,2	8 791	1,2	3,1	272	47,9	120	24
Приморский край	887	4,8	49 223	11,9	4,3	1526	45,2	547	52
Хабаровский край	760	5	34 454	10,9	2,2	1517	48,8	392	40
Амурская область	550	5,1	24 685	4,1	2,3	1220	43,8	301	34
Магаданская область	65	5,1	4 279	0,9	5,7	151	47,5	58	13
Еврейская автономная область	96	4,8	5 212	0,5	7	225	29,9	68	9

7.4 Приложение 4

n	P - 0,01	P - 0,05	P - 0,1	P - 0,15	P - 0,2	P - 0,25	P - 0,3
51	2,6757222	2,0075838	1,6752850	1,4615468	1,2983727	1,1636032	1,0470798
52	2,6737336	2,0066468	1,6746892	1,4611170	1,2980450	1,1633454	1,0468730
53	2,6718226	2,0057460	1,6741162	1,4607037	1,2977298	1,1630975	1,0466741
54	2,6699848	2,0048793	1,6735649	1,4603059	1,2974265	1,1628588	1,0464826
55	2,6682160	2,0040448	1,6730340	1,4599228	1,2971343	1,1626289	1,0462982
56	2,6665124	2,0032407	1,6725223	1,4595535	1,2968527	1,1624073	1,0461204
57	2,6648705	2,0024655	1,6720289	1,4591974	1,2965810	1,1621936	1,0459489
58	2,6632870	2,0017175	1,6715528	1,4588538	1,2963189	1,1619873	1,0457833
59	2,6617588	2,0009954	1,6710930	1,4585219	1,2960657	1,1617881	1,0456234
60	2,6602830	2,0002978	1,6706489	1,4582013	1,2958211	1,1615955	1,0454689
61	2,6588571	1,9996236	1,6702195	1,4578913	1,2955846	1,1614094	1,0453196
62	2,6574786	1,9989715	1,6698042	1,4575914	1,2953558	1,1612293	1,0451750
63	2,6561450	1,9983405	1,6694022	1,4573011	1,2951343	1,1610550	1,0450351
64	2,6548543	1,9977297	1,6690130	1,4570201	1,2949198	1,1608861	1,0448996
65	2,6536045	1,9971379	1,6686360	1,4567478	1,2947120	1,1607226	1,0447683
66	2,6523935	1,9965644	1,6682705	1,4564838	1,2945106	1,1605640	1,0446410
67	2,6512197	1,9960084	1,6679161	1,4562278	1,2943152	1,1604102	1,0445176
68	2,6500813	1,9954689	1,6675723	1,4559795	1,2941256	1,1602609	1,0443978
69	2,6489768	1,9949454	1,6672385	1,4557384	1,2939416	1,1601161	1,0442815
70	2,6479046	1,9944371	1,6669145	1,4555042	1,2937629	1,1599754	1,0441685
71	2,6468634	1,9939434	1,6665997	1,4552768	1,2935893	1,1598387	1,0440588
72	2,6458519	1,9934636	1,6662937	1,4550557	1,2934205	1,1597058	1,0439521
73	2,6448688	1,9929971	1,6659962	1,4548408	1,2932564	1,1595766	1,0438484
74	2,6439129	1,9925435	1,6657069	1,4546317	1,2930968	1,1594509	1,0437475
75	2,6429831	1,9921022	1,6654254	1,4544282	1,2929415	1,1593286	1,0436493
76	2,6420783	1,9916726	1,6651514	1,4542302	1,2927903	1,1592095	1,0435537
77	2,6411976	1,9912544	1,6648845	1,4540374	1,2926430	1,1590936	1,0434606
78	2,6403400	1,9908471	1,6646246	1,4538495	1,2924996	1,1589806	1,0433699
79	2,6395046	1,9904502	1,6643714	1,4536665	1,2923598	1,1588705	1,0432815
80	2,6386906	1,9900634	1,6641246	1,4534881	1,2922236	1,1587632	1,0431953
81	2,6378971	1,9896863	1,6638839	1,4533141	1,2920907	1,1586586	1,0431113
82	2,6371234	1,9893186	1,6636492	1,4531444	1,2919611	1,1585565	1,0430294
83	2,6363688	1,9889598	1,6634202	1,4529788	1,2918347	1,1584569	1,0429494
84	2,6356325	1,9886097	1,6631967	1,4528173	1,2917113	1,1583597	1,0428713
85	2,6349139	1,9882679	1,6629785	1,4526595	1,2915908	1,1582648	1,0427951
86	2,6342123	1,9879342	1,6627654	1,4525055	1,2914732	1,1581722	1,0427207
87	2,6335272	1,9876083	1,6625573	1,4523550	1,2913582	1,1580816	1,0426480
88	2,6328580	1,9872899	1,6623540	1,4522080	1,2912459	1,1579932	1,0425770
89	2,6322042	1,9869787	1,6621553	1,4520643	1,2911362	1,1579067	1,0425075
90	2,6315652	1,9866745	1,6619611	1,4519238	1,2910289	1,1578222	1,0424397
91	2,6309405	1,9863772	1,6617712	1,4517865	1,2909240	1,1577396	1,0423733
92	2,6303296	1,9860863	1,6615854	1,4516521	1,2908214	1,1576587	1,0423083
93	2,6297321	1,9858018	1,6614037	1,4515207	1,2907210	1,1575796	1,0422448
94	2,6291476	1,9855234	1,6612259	1,4513921	1,2906227	1,1575022	1,0421827
95	2,6285757	1,9852510	1,6610518	1,4512662	1,2905265	1,1574265	1,0421218
96	2,6280158	1,9849843	1,6608814	1,4511430	1,2904324	1,1573523	1,0420622
97	2,6274678	1,9847232	1,6607146	1,4510223	1,2903402	1,1572796	1,0420039
98	2,6269311	1,9844675	1,6605512	1,4509041	1,2902499	1,1572085	1,0419467
99	2,6264055	1,9842170	1,6603912	1,4507883	1,2901614	1,1571388	1,0418908
100	2,6258905	1,9839715	1,6602343	1,4506749	1,2900748	1,1570705	1,0418359

7.5 Приложение 5

k1	1	2	3	4	5	6	8	12	24	∞
k2										
1	161,4	199,5	215,7	224,5	230,1	233,9	238,8	243,9	249,0	254,32
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48

7.6 Приложение 6

n	k = 1		k = 2		k = 3		k = 4		k = 5	
	d _L	d _U								
15	1,08	1,36	0,95	1,54	0,82	1,75	0,69	1,97	0,56	2,21
16	1,10	1,37	0,98	1,54	0,86	1,73	0,74	1,93	0,62	2,15
17	1,13	1,38	1,02	1,54	0,90	1,71	0,78	1,90	0,67	2,10
18	1,16	1,39	1,05	1,53	0,93	1,69	0,82	1,87	0,71	2,06
19	1,18	1,40	1,08	1,53	0,97	1,68	0,86	1,85	0,75	2,02
20	1,20	1,41	1,10	1,54	1,00	1,68	0,90	1,83	0,79	1,99
21	1,22	1,42	1,13	1,54	1,03	1,67	0,93	1,81	0,83	1,96
22	1,24	1,43	1,15	1,54	1,05	1,66	0,96	1,80	0,86	1,94
23	1,26	1,44	1,17	1,54	1,08	1,66	0,99	1,79	0,90	1,92
24	1,27	1,45	1,19	1,55	1,10	1,66	1,01	1,78	0,93	1,90
25	1,29	1,45	1,21	1,55	1,12	1,66	1,04	1,77	0,95	1,89
26	1,30	1,46	1,22	1,55	1,14	1,65	1,06	1,76	0,98	1,88
27	1,32	1,47	1,24	1,56	1,16	1,65	1,08	1,76	1,01	1,86
28	1,33	1,48	1,26	1,56	1,18	1,65	1,10	1,75	1,03	1,85
29	1,34	1,48	1,27	1,56	1,20	1,65	1,12	1,74	1,05	1,84
30	1,35	1,49	1,28	1,57	1,21	1,65	1,14	1,74	1,07	1,83
31	1,36	1,50	1,30	1,57	1,23	1,65	1,16	1,74	1,09	1,83
32	1,37	1,50	1,31	1,57	1,24	1,65	1,18	1,73	1,11	1,82
33	1,38	1,51	1,32	1,58	1,26	1,65	1,19	1,73	1,13	1,81
34	1,39	1,51	1,33	1,58	1,27	1,65	1,21	1,73	1,15	1,81
35	1,40	1,52	1,34	1,58	1,28	1,65	1,22	1,73	1,16	1,80
36	1,41	1,52	1,35	1,59	1,29	1,65	1,24	1,73	1,18	1,80
37	1,42	1,53	1,36	1,59	1,31	1,66	1,25	1,72	1,19	1,80
38	1,43	1,54	1,37	1,59	1,32	1,66	1,26	1,72	1,21	1,79
39	1,43	1,54	1,38	1,60	1,33	1,66	1,27	1,72	1,22	1,79
40	1,44	1,54	1,39	1,60	1,34	1,66	1,29	1,72	1,23	1,79
45	1,48	1,57	1,43	1,62	1,38	1,67	1,34	1,72	1,29	1,78
50	1,50	1,59	1,46	1,63	1,42	1,67	1,38	1,72	1,34	1,77
55	1,53	1,60	1,49	1,64	1,45	1,68	1,41	1,72	1,38	1,77
60	1,55	1,62	1,51	1,65	1,48	1,69	1,44	1,73	1,41	1,77
65	1,57	1,63	1,54	1,66	1,50	1,70	1,47	1,73	1,44	1,77
70	1,58	1,64	1,55	1,67	1,52	1,70	1,49	1,74	1,46	1,77
75	1,60	1,65	1,57	1,68	1,54	1,71	1,51	1,74	1,49	1,77
80	1,61	1,66	1,59	1,69	1,56	1,72	1,53	1,74	1,51	1,77
85	1,62	1,67	1,60	1,70	1,57	1,72	1,55	1,75	1,52	1,77
90	1,63	1,68	1,61	1,70	1,59	1,73	1,57	1,75	1,54	1,78
95	1,64	1,69	1,62	1,71	1,60	1,73	1,58	1,75	1,56	1,78
100	1,65	1,69	1,63	1,72	1,61	1,74	1,59	1,76	1,57	1,78

7.7 Приложение 6

CRIME	DIV	UNEMP	STLIV	SCIENCE
8,5629	0,1874	0,1874	2,7435	23,4734
11,1934	0,2048	0,1183	2,0612	21,9773
20,5041	0,2071	0,2641	2,4646	19,3132
18,4391	0,1482	0,0976	2,1982	27,6586
14,6605	0,2343	0,1202	2,7278	16,6433
15,7684	0,2352	0,0535	2,8330	18,7083
15,1960	0,2329	0,1339	2,3696	17,1756
10,1371	0,1890	0,1160	2,3796	23,2809
14,3260	0,2437	0,1070	3,5191	16,8226
7,7258	0,1944	0,1586	2,3996	19,5448
9,1101	0,2437	0,0928	2,8955	17,2337
14,1258	0,2260	0,3390	2,4245	19,4679
21,8384	0,3640	0,1054	4,4922	10,4403
16,1158	0,1925	0,2417	1,8578	22,3383
12,3556	0,1883	0,1010	2,5170	21,7715
14,0474	0,1993	0,1694	2,4708	20,0749
25,1607	0,2780	0,2641	2,5230	14,3875
20,3731	0,2506	0,2560	2,1735	18,3576
31,8684	0,2235	0,1560	2,0431	19,2354
28,0491	0,2088	0,1338	2,1626	18,6815
22,9159	0,3844	0,2883	4,3614	13,5277
21,2013	0,2090	0,1427	2,9611	19,6214
21,1889	0,4036	0,4967	3,6712	12,8841
22,2755	0,3415	0,1863	3,1434	12,8841
17,3495	0,3454	0,2651	3,3414	12,4499
6,9402	0,2939	0,1796	3,9110	12,2474
5,4495	0,2880	0,2880	2,2421	12,8452
22,8404	0,2491	0,2430	2,5392	16,4621
24,6857	0,1405	0,0685	1,7046	27,7489
7,7548	0,5416	0,2462	6,5854	8,1240
5,6964	0,2399	0,1424	1,8963	13,3417
30,0483	0,1391	0,1550	2,1066	25,1595
12,0319	0,2394	0,1449	2,4883	15,8745
8,8199	0,1915	0,1741	2,6866	17,2337
14,3019	0,1526	0,2479	2,1882	20,9762
42,1888	0,1491	0,0982	1,6985	27,4955
20,3926	0,1445	0,2419	1,6059	29,7658
11,3814	0,2141	0,2197	2,7327	17,7482
22,4327	0,1648	0,1086	2,0298	26,7021
18,2418	0,1319	0,0759	1,5570	30,3150
15,5986	0,1821	0,1083	2,4013	20,3224
22,7707	0,2138	0,1710	1,5448	18,7083
23,4228	0,2963	0,2851	3,4100	17,8885
44,8389	0,1775	0,0946	2,2203	25,3574
13,4900	0,3310	0,3393	2,3835	12,0830
17,0687	0,2907	0,2087	3,1454	13,4164
40,4445	0,1630	0,1402	1,4252	26,3818
34,5710	0,1434	0,0797	1,6314	31,3847
18,4327	0,1595	0,2337	1,5466	26,9629
26,8049	0,2075	0,4094	2,4954	17,8326
37,0531	0,1657	0,1749	1,4039	21,7256
18,6927	0,1341	0,2051	1,4828	25,3574
41,6290	0,1900	0,1647	1,1484	23,6854
16,1578	0,4747	0,2830	4,3727	10,9545
37,9254	0,2052	0,1839	1,9326	23,3880
38,3858	0,2525	0,1111	2,4648	19,7990
31,7015	0,2940	0,1326	2,5246	17,3494
8,5349	0,6697	0,7484	6,2371	7,6158
11,6417	0,5821	0,8489	3,6259	8,2462

7.8 Приложение 7

REG	CRIME	DIV	UNEMP	STLIV	SCIENCE	YCENTRE	POPTEEN
Белгородская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Брянская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Владимирская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Воронежская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ивановская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Калужская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Костромская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Курская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Липецкая область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Московская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ИСТИНА	ИСТИНА	ИСТИНА
Орловская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Рязанская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Смоленская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Тамбовская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Тверская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Тульская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ярославская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
г. Москва	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ИСТИНА	ИСТИНА
Республика Карелия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Коми	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Архангельская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ненецкий автономный округ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Вологодская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Калининградская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ленинградская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Мурманская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Новгородская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Псковская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
г. Санкт-Петербург	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Адыгея	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Калмыкия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Краснодарский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ИСТИНА
Астраханская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Волгоградская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ростовская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
г. Севастополь	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Дагестан	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ
Республика Ингушетия	ЛОЖЬ	ИСТИНА	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Кабардино-Балкарская Республика	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Карачаево-Черкесская Республика	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Северная Осетия – Алания	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Чеченская Республика	ЛОЖЬ	ИСТИНА	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ставропольский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Башкортостан	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Марий Эл	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ

Республика Мордовия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Татарстан	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Удмуртская Республика	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Чувашская Республика	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Пермский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Кировская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Нижегородская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Оренбургская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Пензенская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Самарская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Саратовская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ульяновская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Курганская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Свердловская область	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ
Ханты-Мансийский автономный округ – Югра	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Ямало-Ненецкий автономный округ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Тюменская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Челябинская область	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ИСТИНА	ЛОЖЬ
Республика Алтай	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Тыва	ЛОЖЬ	ИСТИНА	ИСТИНА	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Хакасия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Алтайский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Красноярский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Иркутская область	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Кемеровская область	ИСТИНА	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Новосибирская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Омская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Томская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Бурятия	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Республика Саха (Якутия)	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Забайкальский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Камчатский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Приморский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Хабаровский край	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Амурская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Магаданская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Сахалинская область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Еврейская автономная область	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ
Чукотский автономный округ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ	ЛОЖЬ