

# ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ ПРИ ИССЛЕДОВАНИИ КРИМИНОЛОГИЧЕСКИХ ДАННЫХ (на примере Японии)\*



**Евгения Юрьевна НИКИТИНА**, старший преподаватель Института математики и компьютерных наук ДВГУ



**Михаил Александрович ГУЗЕВ**, доктор физико-математических наук, чл. кор. РАН

## ПРОБЛЕМЫ МНОГОФАКТОРНОГО АНАЛИЗА В КРИМИНОЛОГИИ

Во многих слабоформализованных областях деятельности приходится иметь дело с многофакторными зависимостями, когда значение показателя или группы показателей определяется поведением нескольких факторов. Но использование многофакторных моделей на практике для человека, не обладающего профессиональной математической подготовкой, порождает ряд трудно разрешимых стандартными статистическими методами проблем: сколько и какие факторы необходимо учитывать, изменяется ли их влияние с течением времени и др. Обозначенные нами проблемы присутствуют в такой социально-значимой сфере жизни, как общественная безопасность.

Потребность в безопасности — объективный фактор социальной жизни, влияющий на поведение людей. Поэтому сбор криминальной статистики и её анализ являются весьма актуальной задачей криминологии. Развитие этой науки показало, что одной причины преступности не существует, поскольку этот многогранный феномен не может быть порождён каким-либо одним явлением, она — результат процесса функционирования и развития множества взаимосвязанных факторов.

Причины (детерминанты) преступности — это те социально-психологические факторы, от которых непосредственно зависит совершение преступлений. В числе наиболее распространённых называют такие причины преступности, как корыстолюбие, стяжательство, агрессивность, национализм, неуважительное отношение к общественным правилам и нормам, гедонизм, правовой нигилизм [1].

Нельзя свести задачу анализа детерминант преступности к составлению «каталога причин»: в разных сочетаниях, в неодинаковых исходных условиях (жизнедеятельности людей, их изменения условий и предшествующее

\* Работа выполнена при поддержке гранта НШ 2810.2008.1.

состояние преступности) даже наиболее распространённые факторы, оказывающие прямое влияние на преступность, по-разному влияют на количественные и качественные показатели преступности [2].

В криминологии принято выделять количественные и качественные показатели преступности: объём (состояние), коэффициенты преступности и преступной активности, абсолютный рост и темп роста (снижение) преступности, цена преступности и пр. Каждый из них играет свою особую роль в оценке ситуации, но взятый в отдельности не может дать объективного представления об описываемой картине, да и существующие методы моделирования и анализа криминологической ситуации предполагают рассмотрение этих показателей только в их взаимосвязи друг с другом. В теории и на практике юристы замечают несовершенство стандартных методов параметрического корреляционного и регрессионного анализа, дающих ответ на вопрос, как изменится резульативная переменная при изменении факторной переменной, их низкую объясняющую способность [2].

Существуют проблемы достоверности данных о количестве совершённых преступлений, так называемой латентной преступности, не позволяющей объективно оценить криминологическую картину, а также определить направленность основных усилий в борьбе с преступностью. Для объективного анализа важно иметь чёткое представление об объёме и полноте анализируемых данных, так как официальные статистические данные не могут воспроизводить целостную картину, тем более что проблема латентности весьма актуальна, так как является, например, существенным криминогенным фактором.

При исследовании сложных объектов и систем часто мы не можем непосредственно измерить величины, определяющие свойства этих объектов, а иногда нам неизвестны даже число и содержательный смысл факторов. Для измерений могут быть доступны иные величины, тем или иным способом зависящие от этих факторов. К сожалению, построение методов анализа для многомерных статистических данных является весьма трудной и ресурсоёмкой. Поэтому во многих случаях для аккуратного статистического анализа имеющихся данных не находится адекватных статистических средств. В связи с этим приходится применять стандартные, хорошо зарекомендовавшие себя методы даже тогда, когда для их применения не находится достаточно оснований. Конечные результаты такого анализа нелегко интерпретировать, они не имеют ясного вероятностного обоснования [3].

Из вышесказанного следует, что неразработанность методов многофакторного анализа является одной из основных проблем криминологии. Несомненно, учёт многофакторности представляет более глубокий уровень детализации рассматриваемого явления или процесса. Грубое описание предлагает редукцию к более простой модели, в которой число параметров описания сведено к минимуму.

После выбора одной описательной характеристики среди полного набора разнообразных статистических данных о состоянии преступности можно сформировать информационный массив данных по выбранной характери-

стике и упорядочить по специальному алгоритму. Затем становится возможным провести сравнительный анализ наборов эмпирических данных, получить теоретические функциональные зависимости и выявить тенденции в изменяющихся характеристиках и влияющие на них факторы.

## СЕМИОТИЧЕСКИЙ АНАЛИЗ ВРЕМЕННЫХ РЯДОВ И ПОСТРОЕНИЕ РАНГОВЫХ РАСПРЕДЕЛЕНИЙ

В соответствии с изложенной идеей для проведения редукции параметров модели предлагаем оставить для описания криминологической ситуации такую лишённую субъективности характеристику, как количество совершённых преступлений. Информация по ним доступна, достаточна (полна) и репрезентативна, так как правоохранительные органы ведут регулярный статистический учёт всех аспектов состояния преступности. Благодаря этому мы имеем возможность зафиксировать интересующий нас промежуток времени, сформировать конечное множество отдельных составов преступлений, указанных в УК, и составить некоторое ограниченное множество наборов конкретных значений вышеуказанной характеристики по разным составам в разные моменты времени, получив наборы статистических временных рядов.

Проведя эту процедуру, мы можем посмотреть на информационный массив, описывающий картину преступности, как на знаковую (семиотическую) систему, что даст возможность использовать для исследования методы, применяемые в лингвистике. Совокупность сведений о совершённых преступлениях можно рассматривать как единый текст, объединив вместе все виды отчётности правоохранительных органов о зафиксированных преступлениях. Тогда семиотическое понятие «знак» для нашего случая будет иметь смысл конкретного преступного эпизода. А Уголовный кодекс (УК) с формулировками различных составов преступлений будет выступать в роли словаря знаков (здесь «знак» — конкретная статья УК), т.е. множества слов, из которых построен упомянутый выше текст [4, с. 56—57].

Количество знаков в словаре ограничено сверху некоторым конечным числом, в нашем случае это количество статей УК. Стандартным приёмом лингвистического анализа является указание для каждого слова (знака) из составленного словаря одной очевидной характеристики — количества повторений этого слова в анализируемом тексте. Пользуясь этим лингвистическим приёмом, проводим анализ текстовых источников информации по криминальной статистике для создания наборов значений частот встречаемости совершённых преступлений. Имея в своём распоряжении многолетнюю статистику преступности, можно для каждого года составить такой набор для всех составов преступлений, имеющих в УК, указать их частоту появления (встречаемости) —  $\omega$  в сообщениях по криминальной статистике, называя эту характеристику реальной мощностью знака в наборе.

Наряду с реальной мощностью знака введём понятие теоретической (виртуальной) мощности —  $\tilde{\omega}$ . Одним из факторов, определяющих её, являются представления создателей законов о примерном количестве совершаемых конкретных правонарушений, на предотвращение которых направлен создаваемый закон. Эти представления связаны с уровнем компетенции законотворцев, антропо- и социогенными предпосылками создания закона. С точки зрения криминологии, понятие виртуальной мощности преступления связано с ценой преступления, за которой стоит не только реальный вред, причинённый общественным отношениям, но и различные последствия совершённого преступления в политической, духовной, нравственной, трудовой и семейной сферах жизни человека.

Очевидно, что для идеально работающего закона виртуальная мощность совпадает с реальной, так как в нём учитываются все частные случаи и особенности условий применения закона, а также цена преступления, которому он противодействует [4, с. 52—53]. Если бы в распоряжении судебных органов был такой свод законов, то осуждённый по статье этого идеального закона своим наказанием полностью бы исчерпывал меру вины перед обществом. О том, насколько далеки некоторые статьи нашего УК от идеальных, свидетельствуют сообщения о многих криминальных в недавнем прошлом личностях, являющихся в настоящее время депутатами в различных органах законодательной власти, представителями органов государственной власти, преуспевающими бизнесменами и топ-менеджерами ведущих компаний страны. Это означает, что для ныне действующего закона разница между реальной и виртуальной мощностью очень велика. Понятно, что идеал недостижим, усилия законодателей должны быть направлены на минимизацию этой разницы в цене.

Мы предполагаем, что в условиях устойчивого экономического развития общества криминальные взаимоотношения сложились и представляют собой устойчивую структуру. В качестве объекта для исследования рассмотрим криминологическую ситуацию в Японии, где правовая система уникальна и исключительно эффективна. Предпринимаемые японской правоохранительной системой меры воздействия на преступность заслужили общественное признание в стране и за рубежом. Преступность удерживается на весьма низком по международно согласованным меркам уровне, что выражается в основных криминологических характеристиках:

- коэффициент преступности в несколько раз меньше аналогичных показателей в странах Западной Европы и США;
- уровень преступности наиболее стабилен, рост минимален;
- коэффициент раскрываемости в 2—3 раза выше, чем в США, Великобритании, Франции, ФРГ [5].

Для рассмотрения были взяты данные по количеству совершённых преступлений по 33 составам УК Японии с 1996 по 2005 г. [6] Формулировки составов преступлений порождают словарь знаков, суммарная мощность которого ограничена и равна 33. Количество зарегистрированных преступлений по одному составу рассматривалось как частота встречаемости зна-

ка (состава), или реальная мощность знака. За десятилетний период сформировались наборы значений реальных мощностей знаков. Упорядочим их по возрастанию, присваивая каждому знаку номер. Эта процедура называется ранжированием, а введённая новая характеристика знака называется рангом. Таким образом, мы получаем экспериментальную зависимость ранга знака от его реальной мощности.

Следует отметить, что используемый нами подход был предложен В.П. Масловым при решении задачи выбора оптимальной цены на автомобильном рынке США. Помимо номинальной цены, объявленной продавцом, для покупателя существует ещё шлейф дополнительных расходов, возникающих после приобретения автомобиля. К таковым можно отнести стоимость запасных частей в случае поломок, покупка страховок от угона и ДТП, расходы на приобретение и содержание гаражей или нахождение автомобиля на автостоянках и парковках и др. Чтобы избежать рисков при покупке автомобиля, необходимо выбрать товар оптимальной цены, исходя из бюджетных ограничений покупателя. В.П. Маслов рассмотрел ранговые распределения марок автомобиля и числа проданных автомобилей и предложил свой метод нахождения оптимальной цены на автомобильном рынке [4, с. 46—48]. Кроме того, им было сформулировано утверждение, описывающее ситуацию общего положения безотносительно к природе рассматриваемых величин [4, с. 32—34], которым можно воспользоваться для нахождения теоретической зависимости мощности знака от его ранга в наборе.

Используя результаты В.П. Маслова, мы построили для эмпирических диаграмм ранговых распределений криминологических наборов данных (рис. 1) теоретическую двухпараметрическую функцию распределения

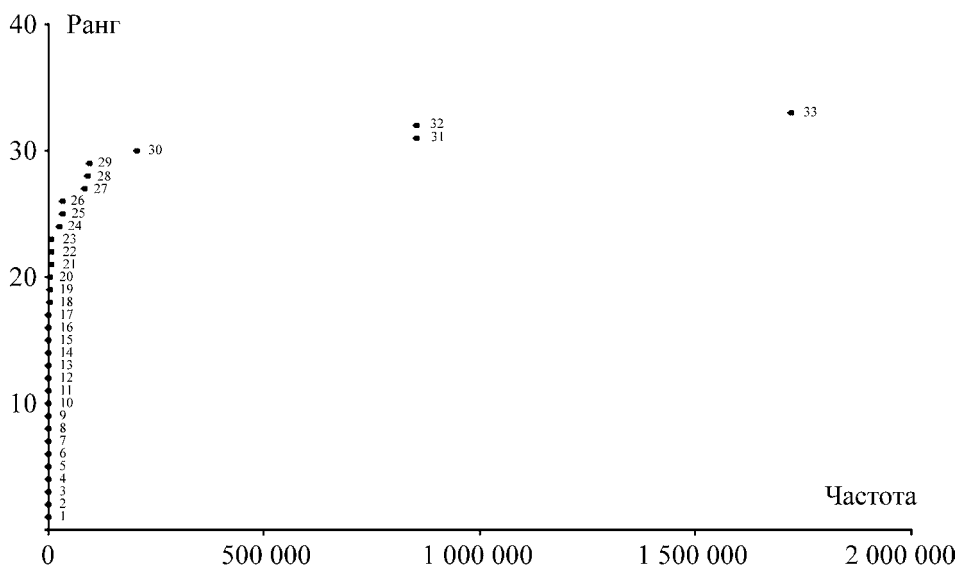


Рис. 1. Эмпирическая зависимость частоты от ранга.

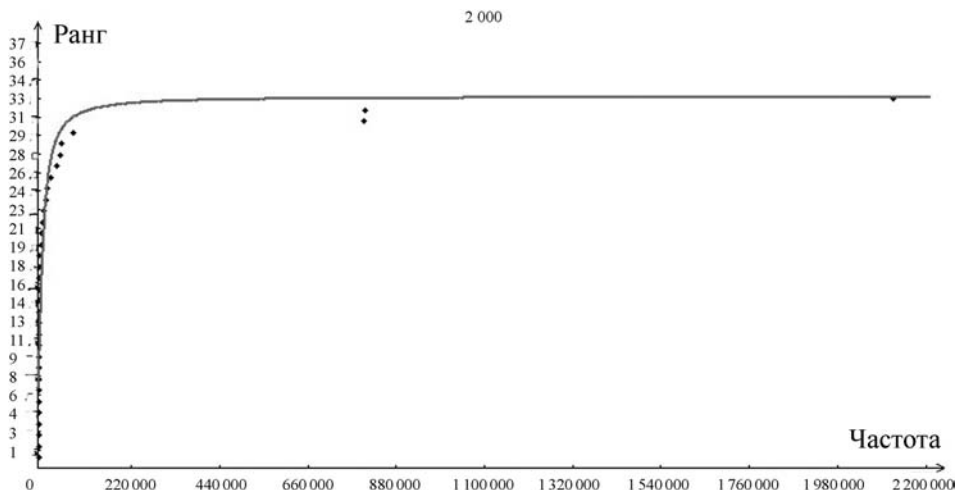


Рис. 2. Зависимость количества совершенных преступлений от их ранга.

в виде арктангенса от дробнорациональной функции (рис. 2), аппроксимирующую экспериментальную зависимость со среднеквадратичной дисперсией равной 0,011 [7].

### АНАЛИЗ РАНГОВЫХ РАСПРЕДЕЛЕНИЙ

Полученные диаграммы эмпирических зависимостей частоты совершенных преступлений от ранга и семейство теоретических кривых для разных лет демонстрируют одинаковое качественное поведение. Установлено, что для ранговых распределений криминологических данных характерно «постоянство ранга» — преступления, соответствовавшие при первичном ранжировании некоторому рангу, в последующие годы изменяют этот ранг не более чем на 2—3 значения (насильственные преступления) (рис. 3), а некоторые виды преступлений вовсе не меняют свой ранг с годами (кражи, халатность) или единожды незначительно изменяют его (мошенничество) (рис. 4).

Такая стабильность связана, во-первых, со «специализацией» преступников и преступных групп на определённых видах преступлений, и смена «рода занятий» криминалитета происходит редко и всегда под влиянием объективных факторов. Вторая причина такой стабильности поведения может быть связана с эффективностью японской правоохранительной системы и в первую очередь УК Японии, при создании которого законодатели нашли оптимальный компромисс между реальной и виртуальной ценой преступления, что позволяет государству контролировать преступность и удерживать её рост на минимальном уровне.

На рис. 2 видна область «провала», где хорошо заметно расхождение виртуальной и реальной мощности законов — теоретическая кривая проходит выше эмпирических данных. Они соответствуют преступлениям, свя-

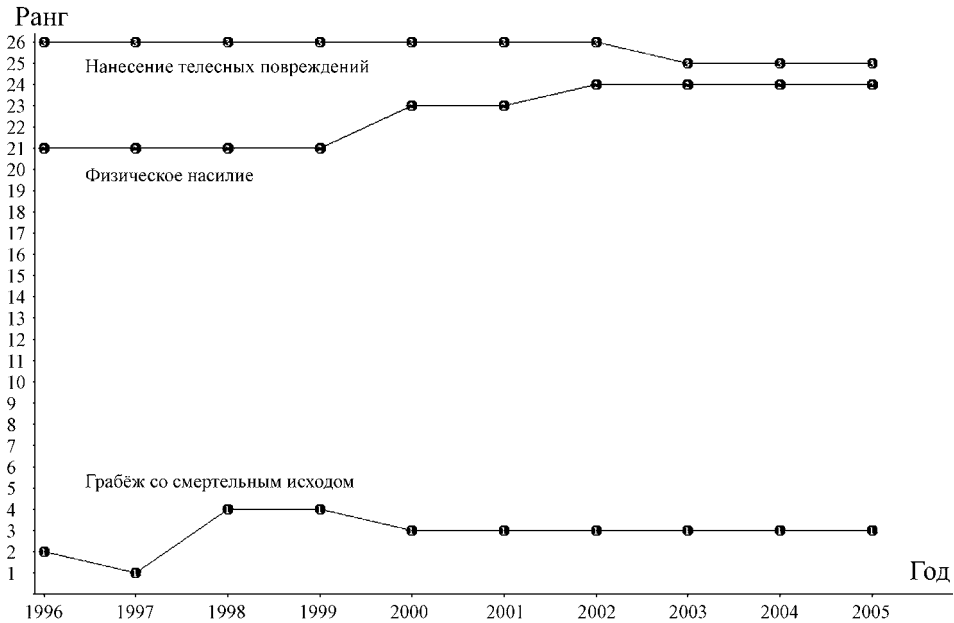


Рис. 3. Изменения ранга для насильственных преступлений.

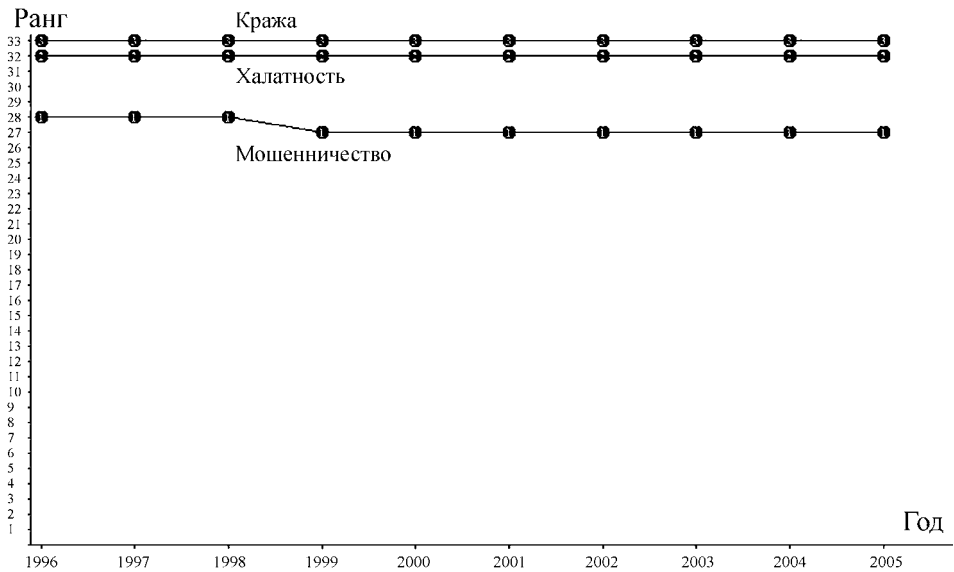


Рис. 4. Изменения ранга для имущественных преступлений.

занным с мошенничеством, порчей имущества, присвоением утерянных вещей. Положение теоретической кривой демонстрирует латентную преступность по этим составам, так как не все указанные преступления регистрируются полицией по разным причинам — например, небольшой причиняемый ущерб, боязнь японцев «потерять лицо». Поэтому можно сделать вывод о несоответствии предполагаемой законодателями «цены преступления» этих составов УК своей виртуальной мощности.

В связи с применением лингвистических методов для анализа криминологических данных необходимо упомянуть, что ещё в конце 40-х гг. XX столетия гарвардский лингвист Джордж Ципф (George Kingsley Zipf), собрав огромный статистический материал, предположил, что распределение слов естественного языка подчиняется одному простому закону: частота появления слова в тексте обратно пропорциональна его порядковому номеру (рангу) слова в частотном словаре. Исследовав разнообразные тексты и языки, в том числе языки тысячелетней давности, Ципф для каждого из них построил указанные зависимости, при этом все кривые имели одинаковую форму — форму «гиперболической лестницы», т.е. при замене одного текста другим общий характер распределения не изменялся.

Закон Ципфа был открыт экспериментально. Позднее французский математик Бенуа Мандельброт предложил его теоретическое обоснование. Он полагал, что можно сравнивать письменный язык с кодированием, причём все знаки должны иметь определённую «стоимость». Исходя из требований минимальной стоимости сообщений, Мандельброт математическим путём пришёл к аналогичной закону Ципфа зависимости — частота появления знака в сообщении обратно пропорциональна его рангу в некотором словаре:

$$\omega = \frac{c}{r^\alpha},$$

где  $\alpha$  — величина (близкая к единице), которая может изменяться в зависимости от свойств текста,  $c$  — числовая константа [8].

Дж. Ципфом и другими исследователями было установлено, что такому распределению подчиняются не только все естественные языки мира, но и другие явления социального и биологического характера: распределения учёных по числу опубликованных ими статей, городов США по численности населения, населения по размерам дохода в капиталистических странах, биологических родов по численности видов и др. [9].

В описанном выше методе В.П. Маслова решения экономической задачи выбора оптимальной цены с помощью рангового распределения реальная мощность  $\omega$  и виртуальная мощность  $\tilde{\omega}$  знака, связаны формулой:

$$\tilde{\omega} = \omega (1 + \alpha \omega^\gamma + \alpha^{-1} \omega^{-\gamma}),$$

где  $\alpha$  и  $\gamma$  — уточняются в процессе исследования [4, с. 61]. При анализе криминологических данных и построении по методу В.П. Маслова рангового распределения в виде арктангенса мы использовали это соотношение. В предельном случае больших частот получили следующее соотношение между рангом преступления и его частотой встречаемости:

$$(r - r_0) \omega^\gamma = c,$$

для которого характерный параметр  $\gamma$  изменяется в пределах от 1,1 до 1,2. Полученная формула напоминает закон Ципфа-Мандельброта, но имеет отклонения от него.



Тем не менее, найденное значение характерного параметра может быть использовано для проведения глобальной аппроксимации на всём массиве криминологических данных и позволит получить значения второго параметра и вычислить виртуальную мощность для каждого закона. Знание точной величины виртуальной мощности даёт возможность составить эффективно действующие законы, руководствуясь объективными предпосылками минимизации риска дополнительных затрат, как это было сделано в экономике.

## ЗАКЛЮЧЕНИЕ

Нас всех волнует качество жизни. Легко указать условия, его определяющие, а вот выбрать факторы, воздействие на которые даст желаемый результат, не так просто. Снова возникает проблема многофакторного анализа. В этом исследовании затронут только один аспект качества жизни — общественная безопасность.

Описанный метод рангового анализа криминологических данных может быть полезен так как указывает измеримые и легко анализируемые во времени оценочные параметры преступности даёт новое понимание причинности этого социального явления.

## ЛИТЕРАТУРА

1. Криминология: Учебник / под ред. Н.Ф. Кузнецовой, В.В. Лунёва. М., 2004.
2. Криминология: Учебник для вузов / под общ. ред. А.И. Долговой. М., 2001.
3. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере / под ред. В.Э. Фигурнова. М.: ИНФРА-М, Финансы и статистика, 1995. 384 с., ил.
4. Маслов В.П. Квантовая экономика. 2-е изд., доп. М.: Наука, 2006. 92 с.
5. Морозов Н.А. Преступность и борьба с ней в Японии. СПб.: Изд-во «Юридический центр Пресс», 2003. 215 с.
6. Источник статистической информации (White Paper on Crime) // Research and Training Institute Ministry of Justice. Japan, 2005.
7. Гузев М.А., Никитина Е.Ю., Морозов Н.А. Применение метода квантовой статистики в криминологии // XXXII Дальневосточная математическая школа-семинар им. академика Е.В. Золотова. 30 авг.— 3 сен. 2007. Владивосток: Дальнаука, 2007. С. 62—63.
8. Чурсин Н.Н. Популярная информатика. К.: «Техника», 1982.
9. Clauset A., Shalizi C.R. and Newman M.E.J. Power-law distributions in empirical data. E-print (2007). <http://arxiv.org/abs/0706.1062v1> (28 нояб. 2008).

**SUMMARY:** The article acquaints readers with a new approach in processing criminological data based on constructing rank distribution according to V.P. Maslov's method. The authors demonstrate some results of its application for describing criminal situation in Japan in 1996—2005.