**МИНИСТЕРСТВО ВНУТРЕННИХ ДЕЛ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

БЕЛГОРОДСКИЙ ЮРИДИЧЕСКИЙ ИНСТИТУТ

ОБЩИЕ ВОПРОСЫ СТАТИСТИКИ

Учебное пособие

**Белгород
Белгородский юридический институт МВД России
2014**

УДК 311
ББК 67.54
О-28

Печатается по решению
редакционно-издательского совета
Бел ЮИ МВД России

О-28 **Общие вопросы статистики** : учебное пособие / *Д. Н. Копонев, Д. Б. Лукьянов, А. Н. Прокопенко, Н. А. Чеканов*. – Белгород : Бел ЮИ МВД России, 2014. – 66 с.

ISBN 978-5-91776-025-9

Рецензенты:

Нудель С.Л., кандидат юридических наук, доцент (НИЦ № 5 ФГКУ «ВНИИ МВД России»);

Шахунов Р.Ю., начальник отдела оперативно-справочной, разыскной и криминалистической информации оперативных учетов (ИЦ УМВД России по Белгородской области).

Учебное пособие «Общие вопросы статистики» разработано в соответствии с рабочей программой учебной дисциплины «Правовая статистика» и предназначено для курсантов и слушателей образовательных организаций МВД России при изучении дисциплины «Правовая статистика», обучающихся по специальностям 030501.65 Юриспруденция и 030505.65 Правоохранительная деятельность, слушателям факультета заочного обучения и факультета повышения квалификации. Данное пособие выходит за рамки рабочей программы учебной дисциплины «Правовая статистика» в части использования математических методов анализа правовых явлений, поэтому может быть использована сотрудниками информационных и аналитических подразделений ОВД

УДК 311
ББК 67.54

ISBN 978-5-91776-025-9

Содержание

ВВЕДЕНИЕ

Основной задачей статистики является получение и соответствующая обработка статистической информации для принятия решений, направленных на достижение желаемого результата в хозяйственной, социально-экономической, научной, культурной и других видах творческой деятельности государства, общественных организаций, экономических структур общества[1].

Для гуманитарных наук, в том числе юридических, статистика является важнейшей наукой, изучающей особенности современного общества. Правовая статистика, как составляющая социальной статистики, имеет отношение не только к конкретным юридическим фактам, но и основана на юридически значимых массовых явлениях и процессах в обществе, и призвана способствовать выявлению наиболее острых проблем, возникающих в процессе развития правового государства.

Необходимым условием для деятельности юристов, особенно управленцев, является статистический анализ этих массовых явлений и процессов. Без достоверных статистических и демографических данных юридические науки могут оказаться в плену догматических представлений о конкретных правовых событиях, которые могут быть объяснимы для небольшого региона, но совершенно непригодны или мало пригодны для страны в целом.

Правоохранительные органы получат возможность наиболее эффективно вести борьбу с правонарушениями и преступлениями и их профилактикой, если примут во внимание данные о преступности, ее структуре и динамике, а также прогнозы о развитии преступных проявлений в будущем.

На основе статистического исследования состояния и уровня преступности, а также возможности и подготовленности оперативных и следственных подразделений полиции, правоохранительные органы смогут не только планировать и распределять свои силы и средства борьбы с преступными проявлениями, но и обосновать причины укрытия преступлений от учета, выявлять причины манипулирования цифрами, что стало модным в XXI веке.

На всех этапах статистического исследования в правовой статистике применяются всевозможные математические методы, на основании которых статистическая работа приобретает научный вид.

В данном пособии рассматриваются математические подходы для разрешения различного рода статистических задач. На первом этапе статистического исследования – статистическом наблюдении – применяются методы выявления ошибок ввода данных, рассчитываются объемы выборочного наблюдения для обеспечения репрезентативности выборки. На втором этапе исследования применяются математические приемы формирования вторичных группировок. Третий этап статистического исследования – статистический анализ –

[1] Лекции по статистике [Электронный ресурс]. Режим доступа: http://WWW.matstat.ksu.ru/Lectures.html.

целиком построен на теории вероятностей, математической статистике, теории корреляционных связей и индексных методах статистического исследования.

Данное пособие будет полезным для курсантов и слушателей образовательных организаций МВД России при изучении дисциплины «Правовая статистика», обучающихся по специальностям 030501.65 Юриспруденция и 030505.65 Правоохранительная деятельность, слушателям факультета заочного обучения и факультета повышения квалификации. Данное пособие выходит за рамки рабочей программы учебной дисциплины «Правовая статистика» в части использования математических методов анализа правовых явлений, поэтому может быть использована сотрудниками информационных и аналитических подразделений ОВД.

1. ЗНАЧЕНИЕ ТЕРМИНА «СТАТИСТИКА»

Слова «статистика», «статистические данные», «статистический анализ», «учет» в настоящее время стали привычны для всех и каждого, так как из средств массовой информации и специальной литературы проникли в нашу повседневную жизнь и наше подсознание.

Учетом и статистикой занимались еще в глубокой древности.

Впервые статистическая информация появилась в Книге чисел («Бемидбар» – «В пустыне») – четвертой книги Пятикнижия Ветхого Завета. В Книге чисел приводятся данные по исчислению народа и боеспособных мужчин, готовящихся к военной кампании. Их количество составляло более 600 тысяч израильтян, которые группировались по 12 коленам – племенам, образовавшим израильский народ[1].

Учет численности облагаемого данью населения в России осуществлялся с IX века. Первую перепись населения Киевской Руси провели татарские ханы в 1246 г. Правда, их целью был не сбор статистических данных, а учет домовладений, облагаемых данью. В 1646 году в целях осуществления налогообложения в России проводился учет подворий, параллельно проводилась перепись населения с учетом возраста, пола и рода занятий. Авторитетный российский историк статистики Авдей Ильич Гозулов считал, что «по программе и уровню организации подворные переписи XVII в. были для своего времени выдающейся формой изучения хозяйства, не имевшей на Западе равной себе системы учета»[2].

Люди, решая свои жизненные потребности, занимались статистическими операциями, не подозревая, что они фактически осуществляли статистическое наблюдение, чем заложили основы статистики.

Впервые термин «статистика» появился в пьесе Уильяма Шекспира под названием «Трагическая история о Гамлете, принце Датском» (1602 г., акт 5, сцена 2). Смысл этого слова у Шекспира – знать, придворные[3].

Статистика как наука стала развиваться с середины XVII века по двум направлениям: первое – английская школа политических арифметиков и второе – немецкая описательная школа[4].

Английская школа политических арифметиков ставила целью изучать общественные явления с помощью числовых характеристик. Эта школа имела два направления: демографическое, основателем которого стал Джон Граунт, и статистико-экономическое, разработанное Вильямом Петти. Первое направление было основано на анализе массовых демографических явлений, законе

---

[1] Электронная еврейская библиотека [Электронный ресурс]. Режим доступа: http://www.eleven.co.il/article/14688

[2] *Гозулов А.И.* История отечественной статистики. - М., 1957. С. 9.

[3] *Орлов А.И.* Нечисловая статистика. - М.: МЗ-Пресс, 2004.

[4] *Чалиев А.Д.* Статистика: курс лекций [Электронный ресурс]. Режим доступа: http://WWW.perepis2002.ru

больших чисел и методах случайных отклонений. Статистико-экономическое направление изучало возможность оценить то или иное явление в условиях нехватки числовых данных.

Немецкая школа, основателем которой стал Герман Конринг, занималась описанием в словесной форме государств, их устройства и политики, географического и климатического расположения, быта населения, религии и др. Такой материал не содержал цифр, что исключало возможность проведения анализа.

Появлением учебной дисциплины «статистика» мы обязаны последователю Г. Конринга, немецкому профессору истории и политической географии Мартину Шмейцелю, который в 1723 г. в Йенском университете начал читать курс политической статистики (Collegium politico-statisticum), в котором излагал политические и географические сведения о государстве.

С 1747 года ученик М. Шмейцеля, впоследствии профессор Мордургского университета Готфрид Ахенваль эту дисциплину назвал статистикой.

М. Шмейцель и Г. Ахенваль для названия дисциплины использовали латинские и итальянские слова о государстве: status – состояние, stato – государство, statista – знаток государства.

Определение научной статистики в XIX веке было дано бельгийским учёным Адольфом Кетле, который, написав 65 сочинений по статистике, положил начало третьему направлению статистической науки, объединившему односторонний математический и описательный – бесчисловой – подходы англичан и немцев.

А. Кетле добавил в науку статистику глубокое философское содержание, особенно в области социальной физики. В своем главном исследовании «О человеке и развитии его способностей» А. Кетле выдвинул идею о необходимости определения значений физического и духовного миров и применения теории вероятностей при обобщении этих значений. Ученый, изучив значения физиологических особенностей человека: веса, силы мускул, скорости дыхания, развития умственных способностей по возрастам и др., пришел к выводу о возможности приобретения человеком нравственных качеств и даже наклонности к совершению преступлений. Он допускал, что каждому человеку присуща известная наклонность или склонность к преступлению, способная, при известных условиях, превратить его в преступника. В каждом обществе существует средняя наклонность к преступлениям[1].

А. Кетле был проникнут убеждением, что общественные явления могут и должны быть изучаемы только на основании правильно устроенного единого систематического наблюдения, он организовал Первый международный статистический конгресс, который состоялся в Брюсселе в 1853 году. В нем участвовали руководители национальных статистических учреждений и видные учёные различных стран. С тех пор Международные статистические конгрессы

[1] *Елисеев С.А.* Зарубежная криминологическая мысль о причинах имущественной преступности: очерк истории // Вопросы уголовного права. Иркутск. Сибирский юридический вестник. 2012. № 1.

стали первой формой международного сотрудничества в области количественного изучения массовых общественных явлений. «Результаты этих конгрессов показали, что временные встречи и обмен мыслей между представителями отдельных стран, при многочисленности состава конгрессов, изменчивости и случайности состава их контингента, не представляли достаточных гарантий для намеченных конгрессом целей»[1].

Было проведено 9 конгрессов, после чего в 1885 году был создан Международный статистический институт с офисом в Гааге, который стал центром международной статистики, занимался сбором и обработкой статистических данных отдельных стран, готовил рекомендации по сопоставимости данных.

В 1970-е гг. на основе Международного статистического института сформировались 3 автономные секции: Международная ассоциация по применению статистики в физических науках, Международная ассоциация муниципальных статистиков, Международная ассоциация специалистов по выборочному методу[2].

Административно-командная система СССР снизила заинтересованность в развитии и совершенствовании новых статистических методов, особенно при организации деятельности коммерческого предприятия. В японских корпорациях все, от председателя совета директоров до рядового рабочего в цехе, знали хотя бы основы статистических методов. В СССР со словом статистика ассоциировали Госкомстат, занимавшийся учетом общественно значимых количественных показателей общества.

В 1989 году проведен анализ стандартов по прикладной статистике СССР, в результате чего 24 из 31 стандарта были отменены как содержащие грубые ошибки, излишне регламентирующие труд специалистов[3].

С таким отставанием вступать в рыночные отношения с предприятиями развитых стран было недопустимо. Необходимо было объединить усилия по организации статистики и ее методов как единой области научной и практической деятельности.

В октябре 1990 года была создана Всесоюзная статистическая ассоциация, президентом которой был назначен профессор, доктор экономических и технических наук, кандидат физико-математических наук А.И. Орлов.

С этого времени к понятию «математическая статистика», добавилось понятие «прикладная статистика». Математическая статистика – раздел математики, разрабатывающий методы регистрации, описания и анализа данных наблюдений и экспериментов с целью построения вероятностных моделей массовых случайных явлений[4]. Математическая статистика является фундаментом прикладной статистики. Прикладная статистика включает в себя методологию организации статистического исследования: как планировать исследование, как

---

[1] *Троицкий Н.А.* Отчет директора Центрального статистического комитета о первой сессии Международного статистического института в г. Риме. - СПб., 1987. С. 2.

[2] *Рябушкин Т.* Международная статистика. - М., 1965.

[3] *Орлов А.И.* Всесоюзная статистическая ассоциация. – М.: МЦНМО, Квант, 1991. № 7.

[4] Вероятностные разделы математики / под ред. Ю.Д. Максимова. - СПБ.: «Иван Фёдоров», 2001. С. 400. 592 с.

собирать данные, как подготавливать данные к обработке, как представлять результаты, а также организацию компьютерной обработки данных[1].

В результате появления прикладной статистики появилось целое направление статистики – статистика объектов нечисловой природы. К статистике нечисловых данных относят данные наблюдений объектов нечисловой природы, которые нецелесообразно описывать числами, в частности элементы различных нелинейных пространств. Примерами являются бинарные отношения (ранжировки, разбиения, толерантности и др.), результаты парных и множественных сравнений, множества, нечеткие множества, измерение в шкалах, отличных от абсолютных[2].

Таким образом, термин «статистика» используется для определения специфической отрасли науки, имеющей свой математический аппарат, в которой излагаются общие вопросы сбора, измерения, анализа массовых количественных и качественных, числовых и нечисловых данных, выявления закономерностей и взаимосвязей различных явлений и процессов.

Кроме науки термин «статистика» можно использовать как:

– совокупность цифровых сведений о массовых явлениях и процессах в природе и общественной жизни. В обиходе статистикой являются статистические данные, представляемые в отчетности предприятий, организаций, отраслей экономики, а также публикуемые средствами массовой информации. Эти данные являются результатом статистического исследования или наблюдения;

– деятельность людей по сбору, анализу и обнародованию массовой количественной информации о различных явлениях и процессах общественной жизни.

Следует отметить, что между разными значениями термина «статистика» существует неразрывная связь, такая же, как между любой наукой, ее практическим применением и отдельными научными фактами.

В последнее время стало модным выделять четвертое значение термина «статистика», как некий параметр ряда случайных величин, рассчитываемый по определенному алгоритму в результате наблюдения для проверки различных гипотез, относительно природы или значений отдельных показателей исследуемых данных. Это значение термина «статистика» в принципе можно отнести к любому направлению из трех выше перечисленных.

Определение статистики как науки дано в каждом учебнике или учебном пособии. Но в основном все определения сводятся к определению, данному в советской энциклопедии: статистика – это отрасль знаний, в которой излагаются общие вопросы сбора, измерения и анализа массовых статистических (количественных или качественных) данных, изучение количественной стороны массовых общественных явлений в числовой форме[3].

---

[1] *Орлов А.И.* Нечисловая статистика. - М.: МЗ-Пресс, 2004.
[2] Там же.
[3] Малая советская энциклопедия. - М.: Советская энциклопедия, 1960. Т. 8. С. 1090.

В учебнике по правовой статистике В.В. Лунеева дано похожее определение, которое наиболее точно отражает процесс изучения количественных показателей во времени и пространстве:

«Статистика есть универсальная наука, изучающая количественную сторону массовых явлений и процессов в целях раскрытия их качественного своеобразия и закономерности развития в конкретных условиях места и времени»[1].

В статистике как науке можно выделить три научных уровня:

1. Общая теория статистики – изучает наиболее общие правила и законы цифрового представления социально-экономических явлений.

2. Экономическая и социальная статистика. Первая изучает структуру, и пропорции общественного воспроизводства. Вторая разрабатывает систему показателей, характеризующих образ жизни населения.

3. Отраслевые статистики. К экономической статистике относятся: статистика промышленности, сельского хозяйства, строительства, транспорта, природных ресурсов и другие. К социальной статистике – демографическая статистика, политическая статистика, статистика здравоохранения, образования и науки, правовая статистика и другие.

Всякое статистическое исследование проходит три этапа: статистическое наблюдение, сводка и группировка наблюдения и анализ.

Рассматривая общие вопросы статистики, применить математический инструментарий общей теории статистики можно на любом из перечисленных уровней и этапов статистического исследования.

Вопросы для самоконтроля:

1. Время появления статистических данных.
2. Первоначальное значение термина «статистика».
3. Основные направления развития статистики как науки.
4. Появления учебной дисциплины «Статистика».
5. Значение деятельности Адольфа Кетле для становления статистики.
6. Понятие прикладной статистики.
7. Понятие статистики нечисловых данных.
8. Современное определение и использование термина «статистика».
9. Определение статистики как науки.
10. Научные уровни статистики.

[1] *Лунеев В.В.* Юридическая статистика: учебник. - М.: Юристъ, 2004. 392 с.

2. ПОНЯТИЕ СТАТИСТИЧЕСКИХ ДАННЫХ И СТАТИСТИЧЕСКИХ ПОКАЗАТЕЛЕЙ В СТАТИСТИЧЕСКОМ НАБЛЮДЕНИИ

Чтобы дать характеристику массовому явлению, необходимо на основе статистических правил собрать первичные данные и сведения о составляющих этого явления, которые впоследствии для достижения целей статистического исследования превратятся в информацию. Далеко не все данные общественного явления могут быть представлены в виде статистической информации.

Статистические данные могут быть собраны в информационные банки данных, и по мере надобности эти данные превратятся в статистическую информацию, необходимую для оценки и обработки. Для этого на этапе сводки и группировки статистические данные наблюдения формируются в группы по определенному критерию, что значительно упрощает процесс анализа.

Очень важно при статистическом наблюдении пройти следующие этапы:

А) Составление плана статистического наблюдения. В план проведения статистического наблюдения включают методологические и организационные вопросы.

При рассмотрении методических вопросов определяются цели, объекты, документы для сбора данных, место и время наблюдения, единицы совокупности и единицы наблюдения.

Единицы совокупности – это первичный элемент объекта, являющийся носителем признаков и являющийся единицей счета. Иначе говоря, единица совокупности – это то, что подвергается обследованию, а единица наблюдения – это источник получаемых сведений. Например, при проведении учета преступности на территории области единицей наблюдения являются территориальные отделы полиции, а единицей совокупности – преступления, лица, их совершившие.

К организационным вопросам относят: выбор способов наблюдения, круг лиц, организующих процесс сбора данных в документах первичного учета, разработку и заполнение форм отчетности и другие текущие задачи, связанные с проведением статистического наблюдения.

Б) Сбор статистических данных. Этот процесс обычно представляет собой заполнение статистических формуляров. Статистический формуляр – это документ единого образца, в котором содержатся вопросы по программе статистического наблюдения и перечень возможных ответов (статистический подсказ) на поставленные вопросы. Ответы на вопросы в документе могут фиксироваться в специально отведенных полях.

В) Контроль полученных данных на основе анализа взаимосвязей между количественными показателями и качественными признаками на всех этапах наблюдения.

Статистическое наблюдение называется точным, если вычисленный по материалам наблюдения показатель соответствует действительной величине. Расхождение между этим показателем и истинной величиной называется ошибкой наблюдения. Ошибки могут быть случайными и систематическими, т.е. имеющими одинаковую тенденцию к изменению значения показателя.

В зависимости от стадии возникновения различают ошибки регистрации, к которым относятся те неточности, которые возникают при записи данных в статистический формуляр. Различают ошибки измерения, репрезентативности, преднамеренные и непреднамеренные ошибки.

Для уменьшения ошибок проводят следующие мероприятия:

- тщательная проверка, отбор и обучение кадров;
- организация контрольных проверок правильности заполнения первичных документов учета;
- проведение арифметического контроля путем сравнения предварительно просчитанных контрольных сумм по строкам или столбцам с полученными итогами;
- проведение синтаксического контроля путем проверки правильности структуры формуляра, заполнению обязательно заполняемых реквизитов;
- логический контроль путем сопоставления кодов классификаторов и словарей с записанным в кодируемые области формуляра значением показателя.

В результате сводки и группировки статистических данных, выполняемых на втором этапе статистического исследования, можно получить показатели, которые могут быть как объемными, например, численность населения района, так и расчетными, например, средний возраст лиц, совершивших кражу или количество краж на одно убийство.

Статистические данные – это конкретные определенные численные значения статистических показателей, которые зависят от места и времени.

Объектом статистического исследования является статистическая совокупность. Каждый ее элемент называется единицей статистической совокупности. Признаком единицы статистической совокупности называют ее конкретное свойство, которое может быть количественно измерено. Эти измерения могут иметь различные количественные значения, т.е. иметь изменчивость – вариацию. Например, возраст преступника или время совершения краж.

Признаки единицы статистической совокупности делятся на качественные, если они не измеряются, а описываются, например, пол, место совершения преступления, предмет преступного посягательства и др. И количественные, если они выражены числом, например, количество погибших, материальный ущерб и др.

А также эти признаки могут быть первичными, выражающими абсолютную величину наблюдаемого явления, например, количество преступлений, и вторичными, которые определяются в процессе анализа данных, например, социальная опасность этих преступлений.

Кроме этого, признаки единицы статистической совокупности могут быть измерены за определенный период времени или в определенный момент времени, например, количество преступлений во время трансляции по телевидению отборочного матча на первенство мира по футболу с участием сборной

России или количество преступлений, зарегистрированных на последнее число месяца.

Статистические показатели делятся на учетно-оценочные, т.е. установленные в определенное время и определенном месте, и аналитические, т.е. основанные на расчетах средних величин, динамики, причинности и связи.

Важным понятием в статистике является понятие закономерности и статистической совокупности.

Закономерность, проявившаяся лишь в большой массе явлений через преодоление свойственной ее единичным элементам случайности, называется статистической закономерностью.

Множество элементов, обладающих массовостью, целостностью, взаимосвязанностью, а также наличием вариации признаков, характеризующих эти элементы, называется статистической совокупностью.

Статистические показатели выражаются с помощью абсолютных, средних и относительных величин.

2.1. Абсолютные статистические величины

Абсолютные показатели являются количественным выражением признаков статистических явлений.

Результатом статистического наблюдения являются показатели, которые характеризуют абсолютные размеры или свойства изучаемого явления у каждой единицы наблюдения. Они называются индивидуальными абсолютными показателями.

Если показатели характеризуют всю совокупность в целом, они называются обобщающими абсолютными показателями. Статистические показатели в форме абсолютных величин всегда имеют единицы измерения: натуральные или стоимостные.

Статистические показатели, которые учитывают физические единицы (метры, кг, штуки и т.д.) называются натуральными. Если показатель применяется при подсчете различных итогов по объектам одного качества, но широкого ассортимента, тогда такой показатель называется условно-натуральным.

Перевод в условное измерение осуществляется с помощью коэффициента пересчета:

К= ФПК / ЗЗК,

где К – коэффициент пересчета, ФКП – фактическое потребительское качество, ЗЗК – заранее заданное качество, эталон.

Натуральные единицы измерения бывают простыми, составными и условными.

Простые натуральные единицы измерения – это тонны, километры, штуки, литры, мили, дюймы и т.д. В простых натуральных единицах также измеряется объем статистической совокупности, т.е. число составляющих ее единиц, или объем отдельной ее части.

Составные натуральные единицы измерения имеют расчетные показатели, получаемые как произведение двух или нескольких показателей, имеющих простые единицы измерения. Например, учет затрат труда на предприятиях выражается в отработанных человеко-днях (число работников предприятия умножается на количество отработанных за период дней). На предприятиях грузового транспорта есть показатель тонна-километры (тонны перевезенного груза умножаются на пробег автомобиля в километрах).

Условно-натуральные единицы измерения широко используют в анализе производственной деятельности, когда требуется найти итоговое значение однотипных показателей, которые напрямую несопоставимы, но характеризуют одни и те же свойства объекта.

Натуральные единицы пересчитываются в условно-натуральные путем выражения разновидностей явления в единицах какого-либо эталона.

Например:

– для учета печатной продукции в полиграфии используют условные печатные листы (1 п.л. = 40000 печатных знаков, включая пробелы);

– для измерения алкогольной продукции используют декалитры абсолютного алкоголя, т.е. спирта, практически не содержащего воды;

– для подсчета общего объема работы транспорта складывают тонно-километры перевезенных грузов.

Перевод в условные единицы осуществляется с помощью специальных коэффициентов. Например, если имеется 100 кг вологодского масла жирностью 82,5% и крестьянского масла жирностью 72,5% – 200 кг, всего 300 кг.

В пересчете на масло жирностью 72,5% (крестьянское) получим общий объем 314 кг (82,5/72,5*100+72,5/72,5*200=114+200).

В пересчете на масло жирностью 82,5% (вологодское) получим общий объем 276 кг (82,5/82,5*100+72,5/82,5*200=100+176).

Пример: в колонии-поселении сельскохозяйственного типа отбывают наказание 120 осужденных к лишению свободы на срок:

1 год – 50 чел.
2 года – 30 чел.
3 года – 14 чел.
4 года – 16 чел.
5 лет – 10 чел.

Найти условно-натуральную величину «количество человек, отбывающих наказание сроком 2 года»:

Решение: задаем эталон – 2 года. Считаем коэффициент пересчета:

1/2 = *0,5*; 1/1 = *1*; 3/2 = *1,5*; 2/1 =*2*; 5/2= *2,5*.

Считаем условно-натуральную величину:

50**0,5* =25; 30**1*=30; 14**1,5*=21; 16**2*=32; 10**2,5*=25

25+30+21+32+25=133 осужденных на 2 года.

Данный показатель на первый взгляд кажется бессмысленным, но, в принципе, может быть использован при анализе различных взаимосвязей при изучении занятости и содержания заключенных.

В условиях рыночной экономики наибольшее значение и применение имеют стоимостные единицы измерения: рубли, доллары, евро, условные денежные единицы и др. Для оценки социально-экономических явлений и процессов используются показатели в текущих или фактически действующих ценах или в сопоставимых ценах.

Сама по себе абсолютная величина не дает полного представления об изучаемом явлении, не показывает его структуру, соотношение между отдельными частями, развитие во времени. В ней не выявлены соотношения с другими абсолютными величинами. Поэтому статистика, не ограничиваясь абсолютными величинами, широко использует общенаучные методы сравнения, обобщения.

Абсолютные величины имеют большое научное и практическое значение. Они характеризуют наличие тех или иных ресурсов и являются основой разнообразных относительных показателей[1].

2.2. Относительные и средние статистические величины

Множество элементов, обладающих массовостью, целостностью, взаимосвязанностью, а также наличием вариации признаков, характеризующих эти элементы, называется статистической совокупностью.

Относительные величины представляют собой обобщающие показатели, выражающие меру количественных соотношений, присущих конкретным явлениям или статистическим объектам.

Относительные показатели по сравнению с абсолютными позволяют точнее оценивать различные явления, например, в 2013 году количество убийств в Российской Федерации и Китае по абсолютным показателям было примерно равным – соответственно 14574 и 13410. В относительных величинах убийства на 100 тыс. жителей разница в 10 раз – в России 10,19, в Китае 1,0[2].

Относительные величины исчисляются как отношение двух чисел. При этом числитель называется сравниваемой величиной, а знаменатель – базой относительного сравнения. Относительные величины могут измеряться: в коэффициентах, в процентах, в промилле, продецимилле в именованных числах[3].

Например, в промилле – тысячная доля, 1/10 процента, обозначается (‰). Продецимилле – 1/10 промилле, или 1/100 процента, обозначается (‱).

[1] Энциклопедия экономиста. Статистика. Общая теория статистики [Электронный ресурс]. Режим доступа: http://www.grandars.ru.

[2] Портал правовой статистики. Генеральная прокуратура РФ [Электронный ресурс]. Режим доступа: http://crimestat.ru.

[3] *Щербина Л.Е.* Общая теория статистики / ЭСМО 2008 [Электронный ресурс]. Режим доступа: http://www.litres.ru/lidiya-scherbina/obschaya-teoriya-statistiki/?lfrom=15465955.

Промилле используется для обозначения количества тысячных долей чего-либо в целом, например, фраза «солёность воды составляет 11‰ (1,1%) означает, что если взять 1 кг воды, то в ней будет 11 г солей.

Данный показатель используется для оценивания состояния алкогольного опьянения человека или концентрации алкоголя в его крови. Считается, что до 0,5‰ алкоголь не оказывает никакого влияния на организм человека. 5‰ и выше может оказаться смертельной дозой[1].

По своему содержанию относительные величины подразделяются на следующие виды: выполнения плана; динамики; структуры; интенсивности; сравнения и др.

Относительная величина выполнения плана представляет собой отношение фактического выполнения обязательства по какому-либо показателю к уровню, предусмотренному планом. Эта величина отражает степень выполнения запланированного показателя и может быть выражена в виде числа или в процентах.

Относительными величинами динамики называются показатели, характеризующие изменение абсолютных величин общественных явлений во времени. Для расчета показателя «темп роста» значение абсолютной величины, соответствующей определенному периоду времени (в статистике иногда называют уровнем ряда), показателя делят на значение абсолютного показателя предыдущего времени. Используем официальные данные статистического наблюдения преступности по Белгородской области с 2010 по 2013 год[2] (табл. 2.1).

Показатель «темп роста» может рассчитываться не только по сравнению с предыдущим годом (цепные показатели), но и с базовым.

Базовым считается год, принятый за отсчет, по отношению к которому проводится сравнение параметров роста и делается соответствующий анализ исследуемых социальных процессов. При расчете показателей динамики указывается – «неподвижная база».

Обычно при анализе масштабных исторических периодов в качестве базового используется год, предшествующий году, в котором произошли значимые события, ставшие причиной крупных организационных изменений.

Например, с 2006 года была введена новая Единая система учета преступлений. Для анализа результатов ее введения целесообразно базовым считать 2005 год.

Темп прироста определяется делением значения абсолютного прироста на значение предыдущего или базового ряда. Темп прироста показывает, на сколько процентов изменился исследуемый показатель по сравнению с предыдущим (базовым) значением. При снижении показателя «темп прироста», его значение будет отрицательным, а темп роста в этом случае будет меньше 100%. Ряды динамики можно изобразить графически (См. рис. 2.1).

---

[1] Лечение и профилактика алкоголизма [Электронный ресурс]. Режим доступа: http://alko03.ru/voditelyam/0-0-promille.html

[2] Портал правовой статистики. Генеральная прокуратура РФ [Электронный ресурс]. Режим доступа: http://crimestat.ru.

Таблица 2.1.

Расчет показателей динамики

	2010	2011	2012	2013
Всего преступлений по Белгородской области[1]	16842	15895	14888	14930
Абсолютный прирост		15895-16842 -947	14888-15895 -1007	14930-14888 42
Темп роста (в коэффициентах) (в процентах)		15895/16842 0,944 94,4%	15895/15895 0,937 93,7%	15895/14888 1,003 100,3%
Темп роста (неподвижная база)		15895/16842 94,4%	15895/16842 88%	15895/16842 89%
Темп прироста (в коэффициентах) (в процентах)		-947/16842 -0,056 5,6%	-1007/15895 -0,063 6,3%	42/14888 0,003 0,3%

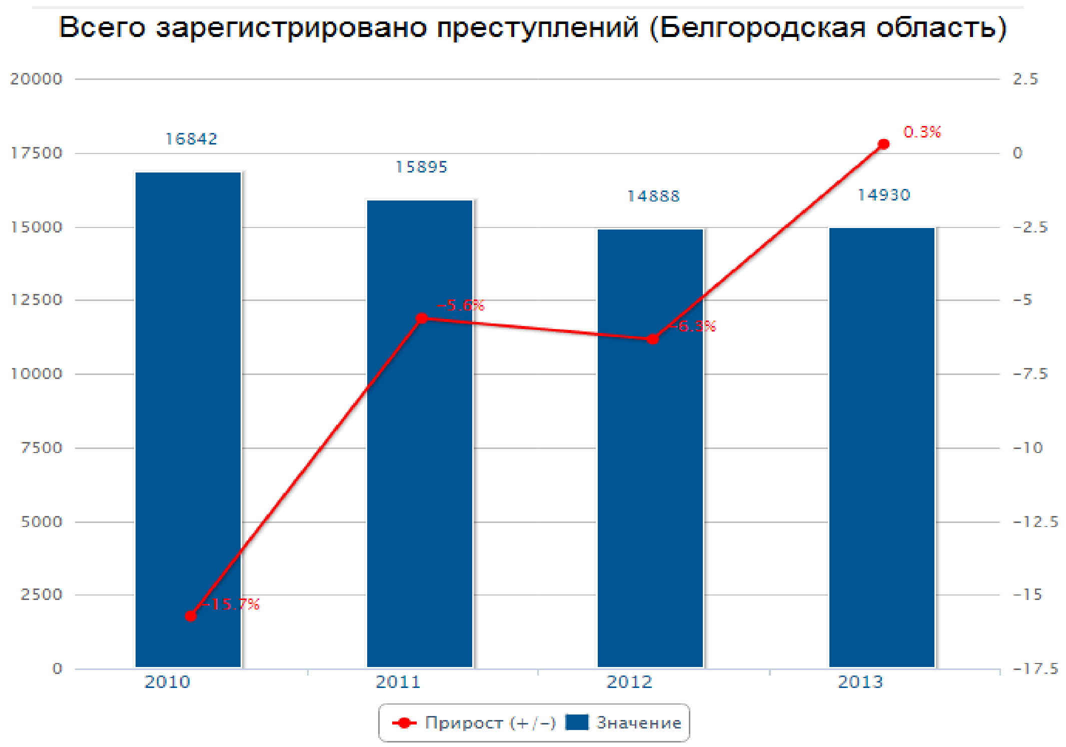

Рис. 2.1. Количество зарегистрированных преступлений в Белгородской области[2]

[1] Портал правовой статистики. Генеральная прокуратура РФ [Электронный ресурс]. Режим доступа: http://crimestat.ru.

[2] Там же.

Относительные величины структуры называются удельными весами составных элементов в общей совокупности и рассчитываются по формуле:

$$\text{Удельный вес части совокупности} = \frac{\text{Уровень части совокупности}}{\text{Вся совокупность}}.$$

Например, в Белгородской области в 2013 году было зарегистрировано 14930 преступлений, из них 483 особо тяжких и 2054 тяжких. Определить удельный вес преступлений небольшой и средней тяжести.

Определяем количество преступлений небольшой и средней тяжести:

14930-483-2054=12393.

Удельный вес этих преступлений в общей преступности составит:

12393/14930*100%=83%.

Относительными величинами интенсивности называются показатели, определяющие степень распространенности данного явления в какой-либо среде.

Для того чтобы выявить, к примеру, распространенность преступлений в различных областях страны, недостаточно назвать их общее количество, зарегистрированное в этих областях, поскольку размещенное в этих областях население по численности существенно отличаются.

Но если вычислить число преступлений, приходившихся на 10000 жителей от всего населения области, то можно ответить, в какой области криминогенная обстановка вызывает опасение.

Относительные величины интенсивности находят широкое применение в практике статистики и определяются путем сопоставления абсолютных величин одноименных уровней на различных территориях, взятых за определенный период времени. Если результаты вычисления слишком малы, их, для наилучшего восприятия, умножают на 10000 или 100000.

В правовой статистике таким образом считают коэффициент преступности КП по формуле:

$$\text{КП} = \frac{\text{Количество преступлений х 10000}}{\text{Численность населения}}.$$

Исходя из данных по Белгородской области, численность населения на 1 января 2014 года составляла1544,1 тысяч человек; Курской – 1118,9 тысяч человек. В Белгородской области зарегистрировано 14930 преступлений в Курской – 13205[1]. Расчет коэффициента преступности может выглядеть следующим образом:

$\text{КП}_{(\text{Белгородская обл.})}$ = (14930*10000)/1544100= 96,69

$\text{КП}_{(\text{Курская обл.})}$= (13205*10000)/1118900=118,01

В статистических отчетах часто встречаются показатели поражаемости преступностью различных слоев населения в виде коэффициента поражаемо-

[1] Портал правовой статистики. Генеральная прокуратура РФ [Электронный ресурс]. Режим доступа: http://crimestat.ru.

сти. Этот коэффициент представляет собой отношение двух удельных весов: удельного веса преступлений, совершенных определенной категорией граждан к общему количеству преступлений, и удельного веса данной группы населения в структуре населения. Данный коэффициент в процентах не измеряется.

Например, если население города составляет 380 тысяч человек, среди которых лиц пенсионного возраста 100 тысяч и ими за год было совершено 300 преступлений, при этом общее количество преступлений по данному городу составило 1500, то коэффициент поражаемости преступностью лиц пенсионного возраста составит:

$$\frac{300/1500}{100000/380000} = \frac{0{,}2}{0{,}263} = 0{,}76.$$

Большое значение в статистике отводится средним величинам, которые относятся к обобщающим показателям. В результате расчета средней величины случайные и нетипичные различия между признаками отдельных величин взаимопоглощаются и получается такое единственное значение, которое характерно для всей совокупности в целом.

Для группировки некоторого множества значений вокруг одного единственного числа в статистике применяют так называемые меры центральной тенденции. Точность этого числа зависит от уровня массовости качественно однородного изучаемого явления.

К основным мерам центральной тенденции относят степенные средние и структурные средние.

Степенные средние

К степенным средним относятся средние гармонические, арифметические, геометрические, квадратические, кубические.

При расчете выше перечисленных средних для одних и тех же исходных данных, значения будут разными. Согласно правилу мажорантности: чем выше показатель степени, тем выше средняя величина. Соответственно значения:

$$\overline{X}_{\text{гарм}} \le \overline{X}_{\text{геом.}} \le \overline{X}_{\text{арифм.}} \le \overline{X}_{\text{квадр.}} \le \overline{X}_{\text{куб.}}$$

Средние величины должны исчисляться так, чтобы при замене каждого индивидуального значения осредняемого показателя его средней величиной оставался без изменения определяющий показатель, связанный тем или другим образом с осредняемым. Значение определяющего показателя определяет формулу расчета средней величины.

Рассмотрим часто применяемые степенные средние:

А) **Среднее арифметическое** – представляет собой сумму всех индивидуальных значений признака, деленную на их количество в совокупности и выражается формулой:

$$\overline{X_a} = \frac{1}{n}\sum_{i=1}^{M}(m_i * x_i) = \frac{1}{n}(x_1 + \ldots + x_n)$$

Б) **Среднее арифметическое взвешенное** – среднее значение, учитывающее весовые коэффициенты для каждого значения. Например, в подразделении проходит службу 24 сотрудника рядового состава, 14 – младшего, 13 сотрудников среднего и 2 – старшего начальствующего состава (x). Каждому рядовому необходимо пройти дополнительную служебную подготовку в количестве 20 часов, младшему составу – 18 часов, среднему – 30 часов, старшему – 22 часа (f). Рассчитать средневзвешенную и среднюю арифметическую величину.

Средняя арифметическая рассчитывается делением «всего часов» на «количество сотрудников»: (20+18+30+22) / (24+14+13+2) =90/53=1,7 ч.

Средняя арифметическая взвешенная рассчитывается по формуле, в которую подставим соответствующие значения:

$$\overline{X} = \frac{\sum_{i=1}^{M} x_i f_i}{\sum_{i=1}^{M} f_i} x = \frac{24 * 20 + 14 * 18 + 13 * 30 + 2 * 22}{24 + 14 + 13 + 2} = 22 \text{ часа}$$

В среднем на каждого сотрудника необходимо затратить 22 часа.

В) **Среднее гармоническое** – число, обратное среднему арифметическому их обратных. Рассчитывается по формуле:

$$A_{-1}(x_1, \ldots, x_n) = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}$$

Из формулы видно, что конструкция средней гармонической выглядит сложнее, чем средняя арифметическая. Расчет среднего гармонического осуществляют тогда, когда статистические наблюдения используют обратные значения. Например, время на одно предварительное следствие или время на задержание одного преступника, время на обучение одного курсанта и т.д.

Для определения среднего времени проведения предварительного следствия сравним показатель среднего арифметического и среднего гармонического. Для этого возьмем продолжительность десяти предварительных следствий по делам особой сложности. Согласно ч.5 ст. 162 УПК РФ этот срок может составить 12 месяцев. Данные поместим в табл. 2.2 и рассчитаем средние арифметическое и гармоническое время на одно предварительное следствие дела особой сложности:

Таблица 2.2.

Определение суммы прямых и обратных величин

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	**Сумма**
Время x_i на одно предварительное следствие (месяц)	2	3,5	3,3	4	4,2	4,3	5,2	6	8,2	9,3	50
$1/x_i$ (обратная величина)	0,50	0,29	0,30	0,25	0,24	0,23	0,19	0,17	0,12	0,11	2,4

Среднее арифметическое =50/10=5 месяцев.

Среднее гармоническое =10/2,4=4,17 месяца.

Для данной задачи наиболее предпочтительным было бы использовать средний гармонический показатель – 4,17 месяца. Согласно правилу мажорантности среднее гармоническое меньше среднего геометрического и меньше среднего арифметического.

Г) **Среднее геометрическое** – это такое число из положительных вещественных чисел, которым можно заменить каждое из этих чисел так, чтобы их произведение не изменилось. Определяется по формуле:

$$G(x_1, x_2, \ldots, x_n) = \sqrt[n]{x_1 x_2 \cdots x_n} = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

Среднее геометрическое может применяться для расчета среднего темпа роста, средних валютных курсов или средних коэффициентов. Например, для расчета среднего темпа роста преступности по Белгородской области за последние три года (табл. 2.1) среднее геометрическое значение составит:

$$\sqrt[3]{94{,}4 * 93{,}7 * 100{,}3} = 96{,}08$$

Следует заметить, что полученное значение геометрической средней всегда будет меньше (равно) значения арифметической средней:

(94,4+93,7+100,3)/3=96,13.

При расчете геометрической средней коэффициентов преступности по Белгородской (96,69) и Курской областям (118,01) на 1 января 2014 года получится – 106,8. Средняя арифметическая – 107,35.

Таким образом, применение средней арифметической величины будет оправдано, если необходимо найти такое значение признака, которое качественно было бы равно удалено от минимального и максимального значений признака.

Д) **Средняя взвешенная для интервала** рассчитывается путем вычисления сначала средней величины для каждого интервала, как полусуммы верхней и нижней границ, а затем – среднюю всего ряда. Для открытых интервалов значение предельного интервала определяется по величине интервалов, примыкающих к ним.

Например, продолжим предыдущую задачу, рассчитав средний возраст сотрудников подразделения, используя среднюю взвешенную для интервального ряда возрастов. Расчет оформлен в табл. 2.3.

Таблица 2.3.

Расчет средней взвешенной интервального ряда

Возраст, лет	**Количество сотрудников**	**Среднее значение интервала**	**Произведение середины интервала на число сотрудников**
До 20	12	20 - (22-20)=18 граница нижнего предела (20+18)/2=19	19*12=228
20-22	6	(20 + 22) / 2 = 21	21*6=126
22-26	10	(22 + 26) / 2 = 24	24*10=240
26-30	12	(26 + 30) / 2 = 28	28*12=336
30 и более	13	граница верхнего предела: 30-26+30=24 (30 + 34) / 2 = 32	32*13=416
Итого:	53		1346

$$x = \frac{\sum x'f}{\sum f} = \frac{1346}{53} = 25{,}4 \text{ года.}$$

Таким образом, учитывая, что данные расчеты дают приблизительное значение, средний возраст сотрудников подразделения 25 лет.

Структурные средние

В качестве структурных средних рассматриваются:

А) **Мода** – наиболее часто встречающееся значение в выборке, например, модой для чисел 1,4,4,8,9,10 будет 4;

Например, кражи стали самым частым преступлением в России. Кроме этого, по информации Федеральной службы исполнения наказаний в 2008 году в 222 следственных изоляторах содержалось 4600 несовершеннолетних. Среди этих несовершеннолетних девять из десяти были осуждены за кражи мобильников.

В 2007 году органами внутренних дел страны зарегистрировано около 350 тысяч фактов похищений мобильных трубок. Но специалисты утверждают: если иметь в виду и незарегистрированные случаи, эта цифра может превысить и 1 миллион!

Таким образом, кражи, а особенно кражи телефонных трубок, в 2007 году были модальным видом преступлений[1].

Б) **Медиана** – значение, которое делит упорядоченные статистические наблюдения пополам. Например, медианой для чисел 1,4,4,8,9,10 будет 6.

[1] Самое распространенное в России преступление. Прокуратура республик Северная Осетия - Алания. Новости. 27.11.2008. [Электронный ресурс]. Режим доступа: http://www.procuror-osetia.ru/news/item/231.

Медиана (Me) – величина варьирующего признака, делящая совокупность на две равные части – со значениями признака меньше медианы и со значениями признака больше медианы.

Например, рассчитаем медиальное значение заработной платы в Белгородской области в 2013 году.

Медианное значение заработной платы делит совокупность пополам: одна половина работников имеет значение заработной платы ниже медианы, другая половина работников – выше медианы. Медиана рассчитывается по формуле:

$$Me = X_{Me} + h_{Me}\frac{\frac{\sum_{i=1}^{k} n_i}{2} - S_{(Me-1)}}{n_{Me}},$$

где: X_{Me} – нижняя граница медианного интервала (первого интервала, накопленная частота которого больше или равна половине суммы всех частот);

h_{Me} – величина медианного интервала;

$S_{(Me-1)}$ – сумма накопленных частот в интервале, предшествующем медианному;

n_i – частота (или частость) i-го интервала;

k – общее количество интервалов вариационного ряда;

n_{Me} – частота (или частость) медианного интервала[1].

Для иллюстрации применения медианного значения рассмотрим данные по заработным платам граждан Центрального федерального округа Российской Федерации. Из табл. 2.4 видно, что медианное значение заработной платы значительно меньше, чем среднее арифметическое значение. Оказывается, среднее арифметическое значение не может показать распределение зарплат. При этом говорить об оценке социального неравенства нельзя: средние показатели зарплаты 10 охранников и одного руководителя фирмы дают среднюю зарплату по предприятию. Учитывая, что зарплата руководителя в разы больше, то подавляющая часть работников среднюю зарплату не получат.

Для более точного анализа социального неравенства граждан по уровню их зарплат целесообразно применять расчет медианной или модальной зарплаты. Сближение отношений средней и медиальной (модальной) зарплат является показателем социального выравнивания.

---

[1] Среднее и медианное значение начисленной заработной платы работников организаций в целом по России и по субъектам Российской Федерации [Электронный ресурс]. Режим доступа: http://www.gks.ru/free_doc /new_site/population/bednost/tabl/3-1-5.htm.

Таблица 2.4.

Применение медианного значения при анализе заработной платы[1]

	2005 г.			**2009 г.**			**2013 г.**		
Территория	**МЕ ЗП, руб.**	**Ср ЗП, Руб.**	**МЕ/Ср**	**МЕ ЗП, Руб.**	**Ср ЗП, Руб.**	**МЕ/Ср**	**МЕ ЗП, Руб.**	**Ср ЗП, Руб.**	**МЕ/Ср**
Белгородская область	4655	6182	75,3	11357	13712	82,8	18776	22156	84,7
Курская область	3706	4971	74,6	9757	12126	80,5	16390	20239	81,0
г. Москва	9861	14421	68,4	26625	36672	72,6	39380	56914	69,2
Центральный федеральный округ	5736	8583	66,8	15437	22517	68,6	24272	35795	67,8

Из табл. 2.4 видно, что социальное обстановка выравниванием социального неравенства более благоприятна в Белгородской области. Причем за период с 2005 по 2013 годы можно отметить ее улучшение с 75,3% до 84,7%. При этом в Москве и в Центральном федеральном округе социальная обстановка по зарплатам практически не улучшилась и составила чуть меньше 70%.

Вопросы для самоконтроля:

1. Этапы статистического наблюдения.
2. Понятие единиц совокупности, единиц наблюдения и единиц измерения.
3. Понятие «статистический формуляр».
4. Понятие ошибки наблюдения и их виды.
5. Мероприятия для уменьшения ошибок наблюдения.
6. Понятие статистических данных.
7. Понятие статистических показателей. Основное отличие статистических данных от статистических показателей.
8. Понятие статистической закономерности и статистической совокупности.
9. Натуральные и условно-натуральные статистические показатели, коэффициент пересчета.
10. Виды единиц измерения. Пересчет натуральных единиц измерения в условные.
11. Понятие, виды и вычисление относительной величины.
12. Относительные величины динамики.
13. Понятие и вычисление удельного веса.
14. Расчет относительной величины интенсивности.
15. Относительные величины.
16. Понятие и требования к вычислению средней величины.
17. Вычисление степенных и структурных средних.

[1] Среднее и медианное значение начисленной заработной платы работников организаций в целом по России и по субъектам Российской Федерации [Электронный ресурс]. Режим доступа: http://www.gks.ru/free_doc/new_site/population/bednost/tabl/3-1-5.htm

3. ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ В СТАТИСТИЧЕСКОМ НАБЛЮДЕНИИ

3.1. Понятие и формы статистического наблюдения

В основе всех научных знаний лежит наблюдение. Для обнаружения обшей закономерности, которой подчиняется явление, необходимо многократно его наблюдать в одинаковых условиях.

Статистическое наблюдение – это научно организованный, планомерный учет фактов, характеризующих изучаемый процесс, и сбор полученных на основе этого учета массовых данных.

В правовой статистике учет состоит из пяти этапов, а именно: отслеживание самих событий преступлений; их регистрация; формирование отчетности путем обобщения информации о преступности; анализ отчетности и внедрение результатов этого анализа в правоохранительную практику[1].

Первые два этапа статистического наблюдения осуществляются сотрудниками правоохранительных органов, непосредственно касающимися расследования преступлений. Они заполняют статистические формуляры – документы первичного учета – статистические карточки и журналы.

Виды документов первичного учета и правила их заполнения оговорены соответствующей инструкцией, утвержденной совместным приказом силовых органов России, которые определили единый для всех органов порядок учета преступлений[2]. На основании этих статистических формуляров субъектами Российской Федерации в информационных центрах МВД РФ формируется отчетность о преступности, т.е. переход к третьему и последующим этапам статистического учета.

Официальный статистический учет в Российской Федерации осуществляется субъектами официального статистического учета. Субъекты официального статистического учета от имени Российской Федерации осуществляют правомочия обладателей официальной статистической информации, формируемой этими субъектами[3].

С 1 января 2012 г. вступила в силу новая редакция ст. 51 Федерального закона от 17 января 1992 г. № 2202-1 «О прокуратуре Российской Федерации».

---

[1] Правовая статистика: учебник для студентов вузов, обучающихся по специальности «Юриспруденция», для курсантов и слушателей образовательных учреждений МВД / В.С. Лялин, Е.А. Костыря, А.В. Симоненко, Е.И. Кузнецова, Е.Н. Барикаев. - 2-е изд., перераб. и доп. - М.: ЮНИТИ-ДАНА, 2010. С. 48.

[2] Приказ Генеральной прокуратуры РФ, МВД РФ, МЧС РФ, МИНЮСТ РФ, ФСБ, Минэкономразвития и торговли РФ, ФС РФ по контролю за оборотом наркотиков от 29.12.2005 г. №№ 39/1070/1021/253/780/353/399 «О Едином учете преступлений» // Российская газета. 2006. № 13.

[3] Федеральный закон от 29 ноября 2007 года № 282-ФЗ ст. 5 «Об официальном статистическом учете и системе государственной статистики в Российской Федерации» (с изм. 23.07.2013 № 251-ФЗ) [Электронный ресурс]. Режим доступа: http://base.consultant.ru/cons/cgi/online.-cgi?req=doc;base=LAW;n=166225;fld=134;dst=101782;rnd= 0.5816348623484373.

Таким образом, Генеральная прокуратура РФ стала «Субъектом официального статистического учета». При этом официальный статистический учет осуществляется в соответствии с федеральным планом статистических работ, который разрабатывается Росстатом совместно с субъектом официального статистического учета, т.е. с Генеральной прокуратурой РФ и утверждается Правительством РФ.

На основании ст. 17 и ст. 51 Федерального закона «О прокуратуре Российской Федерации» Генеральной прокуратурой с 2012 года введены 4 формы статистического федерального наблюдения в сфере правовой статистики:

– № 1-ЕГС «Единый отчет о преступности»;

– № 2-ЕГС «Сведения о лицах, совершивших преступления»;

– № 3-ЕГС «Сведения о зарегистрированных, раскрытых и нераскрытых преступлениях»;

– № 4-ЕГС «Сведения о состоянии преступности и результатах расследования преступлений»[1].

Государственная статистическая отчетность о работе правоохранительных органов является обобщением первичного учета показателей преступности и представлена в настоящий момент формами отчетности за год, квартал, месяц. Вся статистическая информация о преступности публикуется на сайте http://crimestat.ru Генеральной прокуратуры Российской Федерации.

Нарушение порядка представления статистической информации, а равно представление недостоверной статистической информации влечет ответственность, установленную статьей 13.19 Кодекса Российской Федерации об административных правонарушениях от 30.12.2001 № 195-ФЗ, а также статьей 3 Закона Российской Федерации от 13.05.92 № 2761-1 «Об ответственности за нарушение порядка представления государственной статистической отчетности»[2].

Кроме государственной отчетности имеет место внутриведомственная отчетность, которая проводится для своих оперативных нужд. Так, например, МВД РФ разработаны формы отчетности: форма «1-К» о составе кадров МВД РФ, форма «Дисциплина» о состоянии дисциплины, о ходе реализации мер по противодействию коррупции в системе МВД России – форма «ПКор»[3] и др.

Приказом МВД РФ от 03.03.2014 г. № 130 откорректирована форма федерального наблюдения №1-БДД – сведения о безопасности дорожного движения.

---

[1] Приказ Генеральной прокуратуры РФ от 2 июля 2012 г. № 250 «Об утверждении форм федерального статистического наблюдения № 1-ЕГС, № 2-ЕГС, № 3-ЕГС, № 4-ЕГС» / Режим доступа: ГАРАНТ.РУ: http://www.garant.ru/ products/ipo/prime/doc/70130276/#ixzz3GCjQKc8m.

[2] Приказ Федеральной службы государственной статистики от 26.02.2009 г. № 34 «Об утверждении статистического инструментария для организации статистического наблюдения за деятельностью следственных органов и органов дознания, рассмотрением заявлений и сообщений о преступлении» ГАРАНТ.РУ: http://www.garant.ru/products/ipo/prime/doc/-12065569/#ixzz3G7Cfg2tc.

[3] Приказ МВД РФ от 31.07.2012 № 747 «Об утверждении форм статистической отчетности» [Электронный ресурс]. Режим доступа: http://etkovd.ucoz.ru/index/prikazy_mvd_rf_-2012_god_s_601_po_800/0-103.

Разработана ведомственная форма отчета 4-Е «О результатах работы ОВД по обеспечению учетно-регистрационной дисциплины»[1].

Таким образом, можно выделить две основные формы статистического наблюдения: **государственная и ведомственная отчетность**. Кроме этих форм статистического наблюдения существует непрерывное регистровое наблюдение и специальное статистическое исследование.

Непрерывное регистровое наблюдение фиксирует во времени долговременные процессы и мониторинг – слежение за социальными индикаторами.

Существует Единый государственный регистр предприятий всех форм собственности (ЕГРПО), информационный фонд которого содержит: регистровый код, сведения о территориальной и отраслевой принадлежности, форме подчиненности, виде собственности и другие сведения о предприятии.

Существует регистр населения, который с помощью автоматизированной картотеки населения постоянно отслеживает демографические факты рождения, смерти, брака и другие демографические показатели.

Для осуществления мониторинга создаются различные службы мониторинга, например, финансового.

Федеральная служба по финансовому мониторингу (Росфинмониторинг) является федеральным органом исполнительной власти и под руководством Президента Российской Федерации разрабатывает государственную политику в этой сфере. Росфинмониторинг противодействует легализации доходов, полученных преступным путем и финансированию терроризма, является национальным центром по оценке угроз национальной безопасности, возникающих в результате легализации доходов, полученных преступным путем, финансирования терроризма и распространения оружия массового уничтожения[2].

Специальное статистическое исследование позволяет собрать сведения, отсутствующие в официальной отчетности, для чего разрабатываются специальные программы исследования.

Например, при изучении эффективности мер борьбы с преступностью может возникнуть необходимость изучить участие сотрудников правоохранительных органов в правовой пропаганде, работу по повышению их квалификации. В этих целях соответствующими подразделениями ОВД проводятся специальные проверки.

3.2. Выборка в несплошном статистическом наблюдении

Статистическое наблюдение может быть сплошным и несплошным.

Наблюдение преступности является сплошным, т.е. значения статистических показателей формируются на основе данных каждого уголовного дела.

---

[1] Приказ МВД РФ от 23.10.2002 № 1028 «Об утверждении формы статистической отчетности 4-Е» // СПС «КонсультантПлюс» [Электронный ресурс]. Режим доступа:http://base.consultant.ru/cons/cgi/online. cgi?req=doc;base=EXP;n=526887.

[2] Федеральная служба по финансовому мониторингу [Электронный ресурс]. Режим доступа: http://www.fedsfm.ru/about.

Несплошное наблюдение, предполагающее исследовать не всю, а только часть совокупности с помощью выборки определенного размера. По этой выборке можно будет судить о всей совокупности в целом.

Под выборочным наблюдением понимается способ организации наблюдения, при котором исследованию подлежит не вся совокупность, а лишь ее определенная и строго установленная часть[1].

Выборочный метод исследования является единственно возможным в случае бесконечной генеральной совокупности или в случае, когда исследование связано с уничтожением наблюдаемых объектов. Кроме того, он позволяет существенно экономить затраты ресурсов.

Недостатком его является появление ошибок исследования (их называют ошибками репрезентативности), которые связаны с тем, что изучается только часть объекта. Математическая статистика дает рекомендации, как организовать исследование, чтобы свести эти ошибки к минимуму, и дает методику оценки этих ошибок.

Чтобы по данным выборки иметь возможность судить о генеральной совокупности, выборка должна быть отобрана так, чтобы она давала правильное представление о генеральной совокупности.

Например, для проверки качества продукции отобрана партия втулок, изготовленная случайно выбранным рабочим. Но в цехе по производству втулок работают квалифицированные токари и начинающие. Ясно, что если эти втулки изготовлены квалифицированным токарем, то представление о качестве продукции, выпускаемой всем цехом, будет «завышенным», а если изучать втулки, изготовленные начинающим токарем, то «заниженным».

Для того чтобы выборка давала представление о генеральной совокупности, необходимо, чтобы соблюдался принцип равной возможности всем элементам генеральной совокупности быть отобранными в выборку. В приведенном примере в выборку попали втулки, изготовленные только одним рабочим, т.е. эти втулки при отборе имели преимущество и указанный принцип соблюден не был[2].

Выборка называется репрезентативной (представительной – от англ. representative), если она достаточно хорошо воспроизводит генеральную совокупность, т.е. это выборка, которая производится так, что все объекты генеральной совокупности имеют одинаковую вероятность попасть в выборку.

Обеспечить это условие можно различными средствами. Например, отбор можно производить просто на основе таблиц случайных чисел. Таких таблиц сейчас издано много: разработаны программы для ЭВМ – генераторы случайных чисел. Если изучается объект, состоящий из многих разнородных частей, например, мнение избирателей, надо позаботиться о том, чтобы в выборке

---

[1] Правовая статистика: словарь терминов / сост. Д.Н. Копонев, А.А. Дрога. - Белгород: Бел ЮИ МВД России, 2013.

[2] *Лукьянов Д.Б., Прокопенко А.Н.* Курс лекций по информатике и математике. - Белгород: Белгородский юридический институт МВД РФ, 2003.

в соответствующей пропорции были представлены все части системы. В ней должны быть представлены горожане и сельские жители, молодежь и пенсионеры, военные, рабочие, интеллигенция и т.д. из всех частей страны и в той же пропорции, что и во всей стране.

К чему может привести несоблюдение этого правила, показывают многочисленные случаи несбывшихся предвыборных прогнозов. Например, в 1936 году перед президентскими выборами в США журнал «Literary Digest» провел опрос 10 миллионов избирателей и предсказал, что Франклин Рузвельт проиграет выборы. Фамилии избирателей были взяты из телефонных книг. Но в 30-е годы во время депрессии люди, имевшие телефон, не представляли всех избирателей США, выборка оказалась не репрезентативной, и прогноз не оправдался. На телевидении вошло в моду проводить экспресс-опросы во время передачи – желающие сообщить свое мнение могут позвонить в студию и ответить на вопрос «да», «нет» или «не знаю». Такая форма опроса не дает репрезентативной выборки. Примером организации репрезентативного опроса может служить, в частности, метод отбора, который был применен в Англии при проведении обследования рациона питания среднего англичанина. Выборка извлекалась методом трехступенчатого отбора. На первом этапе было отобрано 50 избирательных округов. Затем из них было отобрано некоторое количество избирательных участков. На третьем – некоторое количество семей внутри этих участков. На каждом этапе отбор был строго случайным.

Существуют четыре вида отбора, обеспечивающие репрезентативность выборки: собственно-случайный, механический, типический, серийный и комбинированный отборы.

Собственно-случайный отбор – означает выборку конкретных единиц из генеральной совокупности без дробления последней на группы. При этом существует два варианта отбора – повторный и бесповторный. Повторный отбор предполагает участие уже отобранной единицы в последующей процедуре отбора. Бесповторный отбор такое участие не предполагает. Допустим, в библиотеке производится инвентаризация. Если книги отбираются и откладываются в сторону, то это бесповторная выборка, а если книги отбираются несколько раз, например, по алфавиту и по тематике, то суммарная выборка будет повторной.

Главный принцип организации собственно-случайной выборки состоит в том, что каждый элемент генеральной совокупности имеет равную возможность быть отобранным. При анализе собственно статистических процессов для формирования собственно-случайной выборки используются таблицы случайных чисел или специальные генераторы случайных чисел для ЭВМ.

При механической выборке генеральная совокупность делится на количество частей, равных объему выборки, потом из каждой части отбирается одна единица. Можно пронумеровать элементы генеральной совокупности, а затем в зависимости от объема выборки, установив шаг отбора, произвести их отбор.

Использование типической выборки предполагает деление генеральной совокупности по тому или иному признаку на типические группы. При этом стремятся максимально уменьшить в каждой из них колеблемость признака.

Количество единиц из каждой типической группы выбирается пропорционально среднему квадратическому отклонению признака в этой группе. Типическая выборка характеризуется более точными результатами, чем механическая или собственно-случайная.

Серийная (гнездовая) выборка подразумевает деление генеральной совокупности на ряд групп (серий). Затем случайным образом производят выборку из них. Серии, попавшие в выборку, обследуются сплошным образом. Она применяется в том случае, когда сложно организовать выборочное наблюдение тремя вышеуказанными способами в силу ограниченности ресурсов (людских, материальных и т.п.) участников выборочного наблюдения.

Комбинированная выборка характеризуется выборочным отбором серий на первом этапе и выборочным отбором единиц из этих серий на втором этапе.

3.3. Теоретические основы методов статистических испытаний

Выборочной совокупностью в статистике принято называть совокупность единиц изучаемого явления или процесса, отобранных для наблюдения, а генеральной – всю массу единиц изучаемого явления или процесса[1].

Ответ на вопрос, в какой зависимости находится выборочная средняя величина от аналогичной средней величины генеральной совокупности, дает закон больших чисел. Сущность данного закона заключается в том, что чем больше объем выборочной совокупности будет приближаться к объему генеральной, тем точнее средняя выборочная будет воспроизводить среднюю генеральную величину[2].

Основная задача выборочного наблюдения состоит в том, чтобы характеристики показателей средней и доли выборки позволяли достоверно судить о средней и доле генеральной совокупности. Выборочное наблюдение всегда предполагает сбор и обработку лишь части генеральной совокупности. Отобранная часть изучается, а результаты распространяются на всю исходную совокупность. Наблюдение организуется таким образом, что эта отобранная часть в уменьшенном масштабе представляет всю генеральную совокупность[3].

Рассмотрим оценки достоверности результатов статистических исследований.

При статистических исследованиях предполагается, что случайная выборка $x_1, x_2, x_3, \dots, x_n$, которая представляет конкретные значения случайной величины ξ, принадлежит некоторой, в общем, неизвестной (гипотетической)

---

[1] Правовая статистика: словарь терминов / сост. Д.Н. Копонев, А.А. Дрога. - Белгород: Бел ЮИ МВД России, 2013.

[2] Правовая статистика: учебник для студентов вузов, обучающихся по специальности «Юриспруденция», для курсантов и слушателей образовательных учреждений МВД / под ред. В.С. Лялина, А.В. Симоненко. - М.: ЮНИТИ-ДАНА, 2010.

[3] Выборочное наблюдение в правовой статистике: учебное пособие / Д.Н. Копонев, А.А. Дрога. - Белгород: Бел ЮИ МВД России, 2014. 44 с.

генеральной совокупности. Случайная величина из генеральной совокупности может принимать или конечное число N своих значений $X_1, X_2, X_3, \dots, X_N$, или же бесконечное. В случае конечного числа N предполагается, что это число намного больше объема выборки n, а вопрос – насколько больше – решается непосредственно для конкретного статистического исследования. Набор неизвестных значений $X_1, X_2, X_3, \dots, X_N$ означает те результаты, которые мы могли бы получить из воображаемой генеральной совокупности. Предполагается также, что полученные значения выборки $x_1, x_2, x_3, \dots, x_n$ находятся среди неизвестных значений случайной величины ξ из генеральной совокупности.

Основные числовые характеристики – математическое ожидание $M(\xi) = a$ и дисперсия $D(\xi)$ – случайной величины из генеральной совокупности будем называть, соответственно, генеральной средней и генеральной дисперсией. Важной характеристикой является корень квадратный из генеральной дисперсии, который будем называть генеральным средним квадратическим отклонением и обозначим как $\sqrt{D(\xi)} = \sigma$.

Для надежности и достоверности анализа и выводов из результатов статистических обследований на основе выборочных данных, в первую очередь, необходимо сделать наилучшую оценку генеральных характеристик. Очевидно, что значения $x_1, x_2, x_3, \dots, x_n$ от выборки к выборке будут различными и носить случайный характер, но, установив числовые характеристики генеральной совокупности по статистическим данным, возможно с определенной погрешностью и достоверностью судить о результатах и выводах статистического исследования.

Для оценки генеральных средней a и дисперсии σ^2 служат выборочное среднее арифметическое и выборочная дисперсия, которые обозначим, соответственно, как $\overline{x}$ и $\overline{S}^2$.

Выборочное среднее арифметическое для выборки $x_1, x_2, x_3, \dots, x_n$ объемом n определяется следующим образом:

$$\overline{x} = \frac{x_1 + x_2 + \dots + x_i + \dots + x_n}{n} = \sum_{i=1}^{i=n} x_i \qquad (3.1)$$

Если в выборке объемом n некоторые или все значения повторяются, например, значение x_1 появляется f_1 раз, значение x_2 – f_2 раз, … значение x_i – f_i раз, …, значение x_k – f_k раз, причем сумма $f_1 + f_2 + \cdots f_i + \cdots + f_k = n$, то выборочное среднее арифметическое называется еще взвешенным и вычисляется по формуле:

$$\overline{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots x_i f_i + \cdots + x_k f_k}{f_1 + f_2 + \cdots + f_k} = \frac{1}{n} \cdot \sum_{i=1}^{i=k} x_i f_i, \quad n = \sum_{i=1}^{i=k} f_i, \qquad (3.2)$$

а числа $f_1, f_2,\dots f_i,\dots, f_k$ часто называют частотами.

Выборочная дисперсия вычисляется по формуле

$$\bar{S}^2 = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+\dots+(x_n-\bar{x})^2}{n} = \frac{1}{n}\cdot\sum_{i=1}^{i=n}(x_i-\bar{x})^2 \qquad (3.3)$$

или по формуле

$$\bar{S}^2 = \frac{(x_1-\bar{x})^2 f_1+(x_2-\bar{x})^2 f_2+\dots+(x_k-\bar{x})^2 f_k}{f_1+f_2+\dots f_k} = \frac{1}{n}\cdot\sum_{i=1}^{i=k}(x_i-\bar{x})^2 f_i,$$

$$n = \sum_{i=1}^{i=k} f_i. \qquad (3.4)$$

Для упрощения численных расчетов используется также следующая формула

$$\bar{S}^2 = \frac{1}{n}\cdot\sum_{i=1}^{i=k}(x_i f_i)^2-(\bar{x})^2. \qquad (3.5)$$

Как известно, выборочный метод позволяет сильно сократить трудоемкость на получение статистических данных, а в некоторых случаях он является единственно возможным, например, при статистическом контроле годности снарядов или других изделий, или признаков, которые разрушаются в результате их испытания. В выборочном методе выборки, которые будут нами рассмотрены, могут быть без возвращения (бесповторные) и с возвращением (повторные).

В общей теории вероятностей и математической статистики доказывается, что статистические оценки характеристик генеральной совокупности должны удовлетворять определенным условиям. Статистическая оценка должна быть: состоятельной, несмещенной и эффективной. Выборочное среднее $\bar{x}$, которое находится по формулам (3.1), (3.2) этим условиям удовлетворяет. Что касается генеральной дисперсии $\bar{S}^2$ (формулы (3.3) – (3.5)), то она является смещенной, поэтому вводят исправленную или несмещенную дисперсию, которую обозначим как S^2 и которая связана с выборочной дисперсией $\bar{S}^2$ соотношением:

$$S^2 = \frac{n}{n-1}\bar{S}^2. \qquad (3.6)$$

Из формулы (3.6) видно, что дисперсия S^2 дает лучшую оценку характеристикам генеральной совокупности при малых объемах выборки n, так как она является несмещенной. Однако, если число n достаточно большое, то различие между обеими дисперсиями практически исчезает.

Обратимся теперь к вопросу об ошибках, которые возникают при вычислении генеральных характеристик (среднего арифметического и дисперсии) на основе выборочных данных. Здесь следует иметь формулы, которые дают возможность определить ошибку, например, выборочного среднего при заданном объеме выборки по отношению к его генеральному значению и, наоборот, заранее вычислить объем выборки при условии, чтобы значения ошибки находились в определенном интервале с необходимой точностью. Для этих целей существуют разные методы различной сложности.

Ниже приведем некоторые относительно простые формулы, позволяющие найти ошибки репрезентативности статистических данных и оценить объем выборки, гарантирующий определенную погрешность с заданным уровнем доверительной вероятности. Репрезентативность непосредственно из перевода этого слова на русский язык означает представление или обнаружение всех существенных свойств, которые присущи генеральной совокупности, на основе статистических исследований небольшой ее части.

Разность между выборочными и генеральными характеристиками измеряется средней квадратичной ошибкой выборки, которая вычисляется в случае бесповторной выборки по формуле

$$\varepsilon = \sqrt{\frac{\sigma^2}{n}\left(1 - \frac{n}{N}\right)}, \qquad (3.7)$$

где σ^2 – генеральная дисперсия, n – объем выборки, N – объем генеральной совокупности. Для случая повторной выборки ее средняя квадратичная ошибка находится по формуле

$$\varepsilon = \sqrt{\frac{S^2}{n}}, \qquad (3.8)$$

где S^2 – несмещенная (или исправленная) выборочная дисперсия. Из формул (3.7) и (3.8) видно, что первая формула содержит неизвестную генеральную дисперсию σ, а вторая – известную выборочную несмещенную дисперсию S^2. Но если объем выборки достаточно большой (его объем приближенно можно оценить заранее перед проведением статистических исследований), то для оценки погрешности вместо генеральной дисперсии σ^2 можно использовать исправленную выборочную дисперсию S^2.

Ошибку при определении выборочной доли по какому-то конкретному признаку в общей выборке можно оценить следующим образом

$$\varepsilon = \sqrt{\frac{P(1-P)}{n}}, \qquad (3.9)$$

здесь P – доля этого конкретного признака в выборке, а $(1-P)$ – доля его отсутствия. С точки зрения теории вероятностей, статистическая величина P есть ничто иное, как оценка вероятности того, что в выборочных испытаниях появится событие с исследуемым признаком.

Для пояснения формулы (3.9) обратимся к формуле (3.2), в которой, к примеру, буквой f обозначено число (частота, как иногда называют) появления случайной величины с определенным признаком и с ее конкретным значением, равным x_i во всей выборке с объемом n. Тогда доля этого признака, который обозначен P, очевидно, определяется отношением $P = f / n$. Приведенные выше формулы (3.7) – (3.9) позволяют провести расчет средней квадратичной ошибки выборки по отношению к генеральной совокупности.

Однако можно провести более информативные расчеты ошибок при статистических обследованиях в следующей постановке. Для заданной величины ошибки ε найти интервал, в котором будут находиться характеристики генеральной совокупности, например, генеральное среднее, с заданной доверительной вероятностью γ. Этот интервал называется доверительным интервалом, а величина γ – доверительной вероятностью.

Введем предельно допустимую ошибку выборки

$$\Delta = t \cdot \varepsilon, \tag{3.10}$$

где величина t – так называемый коэффициент доверия, который определяется заданным уровнем доверительной вероятности γ.

Из теории вероятностей известно, что вероятность, с которой случайная величина ξ (то ли выборочная средняя, то ли выборочная дисперсия), распределенная по нормальному закону, принимает свои конкретные значения от числа x_1 до числа x_2, вычисляется по формуле $P(x_1 \le \xi \le x_2) = \Phi(x_2) - \Phi(x_1)$, где функция $\Phi(x)$ есть функция нормального распределения. Далее аргумент функции нормального распределения будем обозначать буквой t. Значения функции $\Phi(t)$, так как она есть несобственный интеграл и не выражается через элементарные функции, численно вычислены при разных значениях аргумента t, табулированы и приводятся во многих книгах и учебниках по теории вероятностей и математической статистике. Однако надо предостеречь, что буквой $\Phi(t)$ разные авторы обозначают близкую, но все-таки другую функцию. Здесь же функция $\Phi(x)$ есть функция нормального распределения, как интеграл от $-\infty$ до t, и которая, таким образом, определена однозначно.

Для вычисления разности между выборочным средним и генеральным средним с условием (требованием) выполнения неравенства $(-\Delta \le a - \overline{x} \le +\Delta)$ для предельно допустимой ошибки при достаточно большом объеме выборки имеем следующее приближенное уравнение

$$P\left(-t \le \frac{a - \overline{x}}{\varepsilon} \le +t \right) \approx 2\Phi(t) - 1. \tag{3.11}$$

Правая часть выражения (3.11) равна назначаемой доверительной вероятности:

$$2\Phi(t) - 1 = \gamma. \tag{3.12}$$

Условие (3.11) означает, что выборочное среднее будет равно генеральному среднему с погрешностью, равной $\Delta = \varepsilon \cdot t$, то есть генеральное среднее будет равно любому числу из интервала $\overline{x} - t \cdot \varepsilon \le a \le \overline{x} + t \cdot \varepsilon$. Поэтому такая оценка генеральной характеристики называется интервальной в отличие от точечной оценки, которая вычисляется по формулам (3.7) – (3.9).

Отметим, что в случае несмещенной оценки (симметрично расположенной относительно генеральной средней a) формула для вычисления вероятности может быть записана в виде $P(-t \le \xi \le +t) = 2\Phi(t) - 1$, что принято во внимание в уравнениях (3.11) и (3.12).

Отметим, что в случае несмещенной оценки (симметрично расположенной относительно генеральной средней a) формула для вычисления вероятности может быть записана в виде $P(-t \le \xi \le +t) = 2\Phi(t) - 1$, что принято во внимание в уравнениях (3.11) и (3.12).

При статистических исследованиях величину доверительной вероятности принимают равной 0,95, 0,954, 0,997 и более, даже 0,9999[1]. К примеру, доверительный уровень вероятности $\gamma = 0{,}954$ означает, что в 1000 случаях только в 46 случаях ошибка может не находиться в доверительном интервале, при назначенном значении вероятности $\gamma = 0{,}997$ – в трех случаях, при величине $\gamma = 0{,}999$ – только в одном случае. Если, к примеру, в планируемых статистических исследованиях выбрана доверительная вероятность $\gamma = 0{,}9545$, то соответствующее значение коэффициента доверия t из уравнения $\Phi(t) = (1 + 0{,}9545) / 2 = 0{,}97725$ можно найти из соответствующей таблицы, которое равно $t = 2{,}0$. Этот результат будет использован нами ниже в одном из рассмотренных примеров.

Если величина γ задана, то из уравнения (3.12) при помощи таблицы, которая в нашем пособии не приводится, для функции $\Phi(t)$ определяется значение аргумента t. Зная теперь конкретное число для параметра t, вычисляем предельную ошибку $\Delta = t \cdot \varepsilon$ по формуле (3.10) и строим интервал, в котором находятся значения генеральной характеристики с ошибкой Δ и с гарантированной вероятностью γ. В частности, средняя величина генеральной совокупности будет находиться в интервале

$$\bar{x} - \Delta \le a \le \bar{x} + \Delta. \qquad (3.13)$$

Аналогично, таким же образом можно вычислить доверительный интервал для генеральной дисперсии.

В правовой статистике, в первую очередь, требуется информация об определенных потенциальных объектах, например, лиц с наклонностями совершения правонарушений или уже с реальными правонарушителями. Следовательно, вопрос статистического обследования таких объектов, т.е. вопрос исследования по какому-то определенному признаку является вполне насущной областью статистического исследования. Поэтому оценка дисперсия для доли

---

[1] В большинстве книг и учебников по правовой и юридической статистике используют таблицы, вычисленные известным статистиком А.Я. Боярским, который всегда полагал доверительную вероятность равной $\gamma = 0{,}954$.

определенного признака в статистических исследованиях выборочным методом действительно важна и определяется по формуле:

$$S^2 = \omega\,(1-\omega), \tag{3.14}$$

где величина $\omega = f\ /\ n$ вычисляется по статистическим данным из выборки объемом n, в которой определенный для исследования признак появляется f раз.

Важной задачей при планировании статистических исследований выборочным методом является предварительное определение оптимального объема выборки для заданной средней ошибки репрезентативности и заданной доверительной вероятностью. Если планируется проводить методом повторной выборки, то объем выборки можно оценить по формуле

$$n = \frac{t^2 S^2}{\Delta^2}, \tag{3.15}$$

а в случае бесповторной выборки этот объем определяется формулой

$$n = \frac{N t^2 S^2}{N\Delta^2 + t^2 S^2}. \tag{3.16}$$

Все входящие в формулы (3.15) и (3.16) величины определены ранее.

Если планируется статистическое обследование по конкретному признаку, например, по возрасту из общего числа осужденных за преступления, то объем выборки, исходя из неравенства Чебышева, можно оценить по формуле:

$$n = \frac{1}{4(1-\gamma)\varepsilon^2} \tag{3.17}$$

Величина выборки, вычисленная по формуле (3.17), сильно превышает ее необходимое значение для достоверных предсказаний, но в силу универсального закона больших чисел увеличение числа объема выборки только увеличивает ее достоверность.

Однако, в реально проводимых статистических исследованиях выборочным методом, в общем, проще и надежно объем выборки и желаемую ее погрешность без каких-то вычислений использовать результаты в виде готовых таблиц, которые в свое время получил наш известный статистик А.Я. Боярский.

Таблицы А.Я. Боярского дают, в общем, реальные ответы на два основных вопроса:

1) если погрешность статистических исследований известна (хотя бы теоретически), то какой при этом объем выборки, который достоверно гарантирует статистические результаты с вероятностью, равной $\gamma = 0{,}954$;

2) если объем выборки заранее задан, то какова вероятность (надежность статистических данных предсказания на весь исследуемый генеральный кол-

лектив) того, что выводы статистических данных справедливы для всего исследуемого коллектива.

Предельная выборочная ошибка Δ % находится на пересечении первого столбца с величиной известного процента ω % данного качественного признака и строки с данным числом n выборки.

Таблица 3.1.

*Определение величины предельной выборочной ошибки Δ %
при статистическом исследовании выборочным методом по какому-либо качественному признаку из всей совокупности, в которой процент ω % наличия данного признака и объем выборки n известны*[1]

Процент данного признака, ω, %	**Объем выборки (число наблюдений), n**								
	100	**200**	**300**	**400**	**500**	**600**	**700**	**800**	**900**
10	6,0	4,3	3,5	3,0	2,7	2,5	2,3	2,1	2,0
15	7,2	5,1	4,1	3,6	3,2	2,9	2,7	2,5	2,4
20	8,0	5,7	4,6	4,0	3,6	3,3	3,0	2,8	2,7
30	9,2	6,5	5,3	4,6	4,1	3,7	3,5	3,2	3,1
35	9,6	6,8	5,5	4,8	4,3	3,9	3,6	3,4	3,2
40	9,9	7,0	5,6	4,9	4,4	4,0	3,7	3,5	3,3
45	10,0	7,1	5,7	5,0	4,5	4,1	3,8	3,5	3,3
55	10,0	7,1	5,7	5,0	4,5	4,1	3,8	3,5	3,3
65	9,6	6,8	5,5	4,8	4,3	3,9	3,6	3,4	3,2
70	9,2	6,5	5,3	4,6	4,1	3,7	3,5	3,2	3,1
75	8,7	6,2	5,0	4,3	3,9	3,5	3,3	3,1	2,9
80	8,0	5,7	4,6	4,0	3,6	3,3	3,0	2,8	2,7

Табл. 3.1 с определенной и достаточной надежностью дает ответ на вопрос: какова достоверность результатов, полученных в результате выборочных исследований (опросов) небольшого, в общем, числа, т.е. некоторой части (доли) всей общей совокупности, чтобы сделать вывод, точнее распространить выборочные результаты на всю совокупность.

К примеру, мы провели среди населения некоторой ее части исследование (анонимный опрос, как его организовать – это особая задача, впрочем, решаемая) об их отношении к правоохранительным органам. Насколько будет достоверным истинное отношение всего населения к ответам всего некоторой его части?

Конечно, естественно возникает первый и основной вопрос: на каком реальном фундаменте основан данный выше ответ в виде какой-то таблицы?

---

[1] Таблица взята из учебного пособия (стр.106): *Горемыкина Т.К.* Общая и правовая статистика. - М.: МГИУ, 1999. 164 с. В этом учебном пособии указано, что эти таблицы содержат выдержки из работы А.Я. Боярского «Таблицы для определения достоверности статистических показаний и числа наблюдений в статистическом исследовании». - М., 1947.

Как обнаружили гениальные умы человечества при исследовании случайных явлений, что и здесь имеются законы, которые диктуются самой природой независимо от воли и прихоти людей. Основным законом в случайных явлениях является закон, названный законом больших (массовых) чисел. Не вдаваясь в теорию этих универсальных законов, но в их подтверждение приведем не подлежащее сомнению утверждение: число новорожденных мальчиков и девочек за месяц, за год и более примерно одинаково.

При выборочном методе исследования возникает следующий вопрос. Какое минимальное небольшое число, например, человек, необходимо исследовать (опросить), чтобы результаты такого статистического исследования с определенной погрешностью представляли весь намного больший по числу коллектив? Это минимальное необходимое число из всей совокупности, называемое объемом выборки, можно определить по табл. 3.2.

Таблица 3.2.

Объем выборки n, который необходим для того, чтобы выборочная ошибка ε % не превысила заданной ей предельной величины Δ % при данном проценте ω % исследуемого качественного признака[1]

Процент данного признака, ω, %	**Предельное значение выборочной ошибки, Δ %**					
	1	**2**	**3**	**4**	**5**	**10**
10	3 600	900	400	230	150	37
15	5 100	1 300	570	320	210	52
20	6 400	1 600	710	400	260	65
25	7 500	1 900	830	470	300	76
30	8 400	2 100	930	530	340	85
35	9 100	2 300	1 010	570	370	92
40	9 600	2 400	1 070	600	390	97
45	9 900	2 500	1 100	620	400	100
50	10 000	2 500	1 110	630	400	100

Планируемый необходимый объем выборки n находится на пересечении первого столбца, в котором приведена величина процента ω % данного качественного признака и строки с заданной предельной ошибкой Δ %.

Таким образом, таблицы А.Я. Боярского с достоверностью $\gamma = 0{,}954$ позволяют без вычислений дать оценку предельной ошибки Δ % при заданной величине погрешности ε % и известном объеме выборки n при выборочном исследовании по какому-то качественному признаку, а при планировании тако-

[1] Таблица взята из учебника (стр. 129): *Лунев В.В.* Юридическая статистика. - М.: Юристъ, 2007. 394 с. В этом учебнике указано, что эта таблица приводится в сокращенном виде из работы А.Я. Боярского «Таблицы для определения достоверности статистических показаний и числа наблюдений в статистическом исследовании». - М., 1947.

го выборочного исследования – оценить объем выборки при заранее заданном проценте данного признака и предельной ошибки. К тому же следует добавить, что в правовой статистике исследование по качественному признаку представляет наибольший практический интерес.

Для более точных предсказаний, конечно, следует использовать приведенные выше в этом разделе формулы, которые доказаны в общей теории математической статистики.

Ниже приведем примеры, в которых будут проведены, в первую очередь, расчеты по теоретическим формулам.

Рассмотрим относительно простой пример, в котором генеральная совокупность известна и ее объем небольшой, что позволяет провести вычисления как генеральных характеристик, так и соответствующих выборочных характеристик и выполнить сравнительный анализ полученных результатов.

Пример. Численность курсантов, обучающихся на потоке, составляет 100 человек, и их оценки приведены в таблице. Определить успеваемость (средний балл) курсантов в двух случаях:

1) с учетом всех курсантов, рассматривая их как генеральную совокупность;

2) на основе 20% выборки. Сравнить полученные результаты для этого примера, а также с теоретическими предсказаниями по формулам, приведенным выше в этом разделе.

Таблица 3.3.

Распределение оценок (баллов)

Оценки (баллы)	**Количество курсантов**
Неудовлетворительно, «2»	10
Удовлетворительно, «3»	45
Хорошо, «4»	30
Отлично, «5»	15
Всего курсантов	100

Полагаем, что весь поток курсантов составляет генеральную совокупность объемом $N = 100$. Оценка (балл) является случайной величиной, которая может принимать четыре значения x_i,.($i = 1, 2, 3, 4$), то есть $x_1 = "2"$, $x_2 = "3"$ $x_3 = "4"$, $x_4 = "5"$.

Из приведенной исходной табл. 3.3 находим, что оценка «2» ($x_1 = 2$) появляется с частотой $f_1 = 10$, оценка «3» ($x_2 = 3$) – с частотой $f_2 = 45$, оценка «4» ($x_3 = 4$) – с частотой $f_3 = 30$ и оценка «5» ($x_4 = 5$) – с частотой $f_4 = 15$. Сумма всех частот равна объему $N = 100$ генеральной совокупности

$$\sum_{i=1}^{i=4} f_i = f_1 + f_2 + f_3 + f_4 = 10 + 45 + 30 + 15 = 100.$$

По известным формулам находим генеральные взвешенные среднее a и дисперсию σ^2:

$$a = \frac{1}{N} \cdot \sum_{i=1}^{i=4} x_i f_i, \qquad a = \frac{2 \cdot 10 + 3 \cdot 45 + 4 \cdot 30 + 5 \cdot 15}{10 + 45 + 30 + 15} = \frac{350}{100} = 3,50,$$

$$\sigma^2 = \frac{1}{N} \cdot \sum_{i=1}^{i=4} (x_i - a)^2 f_i,$$

$$\sigma^2 = \frac{(2 - 3,50)^2 \cdot 10 + (3 - 3,50)^2 \cdot 45 + (4 - 3,50)^2 \cdot 30 + (5 - 3,50)^2 \cdot 15}{10 + 45 + 30 + 15} =$$

$$= \frac{75}{100} = 0,75,$$

а также генеральное среднее квадратичное отклонение от среднего значения

$$\sqrt{\sigma^2} = \sigma \approx 0,866.$$

Итак, для генеральной совокупности имеем следующие характеристики

$$a = 3,50, \quad \sigma^2 = 0,75, \quad \sigma \approx 0,866.$$

Для сравнения с генеральными характеристиками из нее возьмем случайно 20%-ю выборку объемом $n = 20$ и вычислим аналогичными выборочные характеристики.

Как организовать случайный 20%-й выбор оценок «2», «3», «4», «5» из всей генеральной совокупности объемом $N = 100$?

Один из приемов состоит в использовании таблиц случайных чисел, которым и воспользуемся. Пусть оценки расположены в следующем порядке:

1 – 10 находится оценка «2»,
11 – 55 находится оценка «3»,
56 – 85 находится оценка «4»,
86 – 100 находится оценка «5»,

которые условно пронумерованы от 1 до 100 (конечно, порядок нумерации может быть произвольным).

При помощи таблицы случайных чисел[1] получаем следующую выборку объемом $n = 20$:

Оценка «2» ($x_1 = 2$) появляется с выборочной частотой $\overline{f}_1 = 1$, оценка «3» ($x_2 = 3$) – с частотой $\overline{f}_2 = 10$, оценка «4» ($x_3 = 4$) – с частотой $\overline{f}_3 = 6$ и оценка «5» ($x_4 = 5$) – с частотой $\overline{f}_4 = 3$.

Способ получения случайной выборки из таблицы случайных чисел опишем позже в конце этого примера.

Для проверки общего числа случайного выбора из таблицы случайных чисел находим, что объем выборки $\overline{f}_1 + \overline{f}_2 + \overline{f}_3 + \overline{f}_4 = 1 + 10 + 6 + 3 = 20$ действительно равен $n = 20$.

Вычисляем выборочное среднее $\overline{x}$ и выборочную несмещенную дисперсию S^2:

$$\overline{x} = \frac{1}{n} \cdot \sum_{i=1}^{i=4} x_i \overline{f}_i, \quad s^2 = \frac{2 \cdot 1 + 3 \cdot 10 + 4 \cdot 6 + 5 \cdot 3}{1 + 10 + 6 + 3} = \frac{71}{20} = 3{,}55,$$

$$S^2 = \frac{1}{n} \cdot \sum_{i=1}^{i=4} (x_i - \overline{x})^2 \overline{f}_i,$$

$$S^2 = \frac{(2 - 3{,}55)^2 \cdot 1 + (3 - 3{,}55)^2 \cdot 10 + (4 - 3{,}55)^2 \cdot 6 + (5 - 3{,}55)^2 \cdot 3}{1 + 10 + 6 + 3} =$$

$$= \frac{12{,}950}{20} \approx 0{,}6475$$

и выборочное среднее квадратичное отклонение от среднего значения $\sqrt{S^2} = S \approx 0{,}8047$.

Таким образом, для выборки получены следующие ее характеристики

$$\overline{x} = 3{,}55, \quad S^2 \approx 0{,}6475, \quad S \approx 0{,}8047.$$

Вычислим абсолютную и относительную ошибки для выборочного среднего:

$$abs(a - \overline{x}) = 0{,}05, \quad abs\left(\frac{a - \overline{x}}{a}\right) = \frac{0{,}05}{3{,}50} \approx 0{,}01429,$$

относительная ошибка составляет примерно 1,4 %.

Вычислим абсолютную и относительную ошибки для выборочной дисперсии:

---

[1] Источник: The RAND Corporation. A Million Random Digits with 100 000 Normal Deviates. – N.Y.: Free Press, 1966. [Электронный ресурс]. Режим доступа: http://www.dpva.info/Guide/-GuideMathematics/GuideMathematicsFiguresTables/RandomDigitsTable/

$$\sigma^2 - S^2 = 0{,}1025, \qquad \frac{\sigma^2 - S^2}{\sigma^2} \approx 0{,}137,$$

а также для выборочного среднего квадратичного отклонения:

$$\sigma - S \approx 0{,}06135, \qquad \frac{\sigma - S}{\sigma} \approx 0{,}071.$$

Как видно, для выборочного среднего квадратичного отклонения относительная ошибка составляет примерно 7,1 %.

Теперь вычислим теоретическую ошибку по формуле (3.7), пометив ее нижним индексом «теор», в которой положим $\sigma = 0{,}866$, $N = 100$ и $n = 20$:

$$\varepsilon_{теор} = \sqrt{\frac{0{,}866 \cdot 0{,}866}{20} \cdot \left(1 - \frac{20}{100}\right)} \approx 0{,}17320,$$

В процентном выражении равная 17,3 %. Реальная же ошибка, вычисленная выше, равна 7,1 %. Это не удивительно по той причине, что теоретические оценки дают завышенные числа. Но эти завышенные планируемые числа только на пользу, в том смысле, что они еще более усиливают достоверность статистических данных и их выводов.

Вычислим по теоретической формуле (3.16) необходимое число объема выборки, обозначив его также нижним индексом «теор», в которой для нашего примера положим $\varepsilon = 0{,}1732$ и получим, что это число равно 18, $n_{теор} = 18$. В нашем примере объем выборки равен $n = 20$, что вполне достаточно для достоверности результатов и выводов таких статистических исследований. К этому добавим что, как написано выше, это находится в согласии с теоремой П.Л. Чебышева о законе больших чисел, которая утверждает, что с увеличением объема выборки надежность и достоверность данных статистических обследований предсказания только увеличивается.

Теперь коротко опишем способ, как в нашем примере случайно отбирается выборка $n = 20$ из генеральной совокупности объемом $N = 10$. Из таблицы случайных чисел[1] выписываются подряд, в нашем примере с первого числа, один за другим числа из этой таблицы. Затем этот ряд из чисел разбивается на группы из трех чисел (так как наша генеральная совокупность имеет объем $N = 100$), из которых отбираем числа, значения которых являются двузначными от 01 до 100. Реально этот процесс занимает длительное время, которое намного можно уменьшить при использовании современных компьютеров.

Пример. Используя данные нижеприведенной таблицы[2], вычислить средний возраст осужденных мужчин, а также найти возрастной интервал, на который приходится наибольшее число осужденных.

---

[1] Источник: Источник: The RAND Corporation. A Million Random Digits with 100 000 Normal Deviates. – N.Y.: Free Press, 1966. [Электронный ресурс]. Режим доступа: http://www.dpva.info/Guide/GuideMathematics/GuideMathematicsFiguresTables/RandomDigitsTable/

[2] Из портала правовой статистики генеральной прокуратуры РФ, режим доступа: crimestat.ru

Таблица 3.4.

Распределение лиц (мужчины), совершивших преступление, по возрасту (данные по РФ за 2013 год)

	Номера возрастных интервалов					
	1	**2**	**3**	**4**	**5**	**6**
Возраст осужденных, годы	14 – 15	16 – 17	18 – 24	25 – 29	30 – 49	50 и старше
Количество осужденных (частота), f_i	17 155	36 907	196 310	163 770	363 630	78 523

Здесь f_i обозначает частоту (или вес) i-й возрастной группы, то есть число осужденных из этой группы.

Для нахождения среднего возраста осужденных воспользуемся вначале следующей формулой:

$$\overline{x} = \frac{\sum_{i=1}^{i=6} x_i f_i}{\sum_{i=1}^{i=6} f_i},$$

в которой переменная x_i есть середина возрастного интервала, равная полусумме крайних чисел этого i-го интервала. Но заметим, что для шестого интервала, для которого правая граница не задана, его величину условно выбираем равной длине предыдущего пятого интервала, то есть равной 19 годам. По данным исходной табл. 3.4 вычисляем сумму $\sum_{i=1}^{i=6} f_i$:

$$\sum_{i=1}^{i=6} f_i = f_1 + f_2 + f_3 + f_4 + f_5 + f_6 =$$

$$= 17\,155 + 36\,907 + 196\,310 + 163\,770 + 363\,630 + 78\,523 = 856\,295,$$

т.е. общее число осужденных равно $N = 856\,295$

и сумму $\sum_{i=1}^{i=6} x_i \cdot f_i$:

$$\sum_{i=1}^{i=6} x_i \cdot f_i = x_1 \cdot f_1 + x_2 \cdot f_2 + x_3 \cdot f_3 + x_4 \cdot f_4 + x_5 \cdot f_5 + x_6 \cdot f_6 =$$

$$=17155*14{,}5 + 36907*16{,}5 + 196310*21{,}0 +$$

$$+ 163770*27{,}0 + 363630*39{,}5 + 78523*59{,}5 = 28437516{,}5.$$

Затем по приведенной выше формуле находим среднюю величину возраста осужденных:

$$\overline{x} = \frac{28\,437\,516{,}5}{856\,295} \approx 33{,}2\ (\textit{года}).$$

Исходные данные и результаты вычислений записываем в следующую итоговую табл. 3.5.

Таблица 3.5.

Результаты вычислений

№№ п/п, i	**Возраст осужденных, годы**	**Количество осужденных (частота),** f_i	**Середина интервала,** x_i**, годы**	**Произведение,** $x_i \cdot f_i$
1.	от 14 до 15	17 155	14,5	248 747,5
2.	от 16 до 17	36 907	16,5	608 965,5
3.	от 18 до 24	196 310	21,0	4 122 510,0
4.	от 25 до 29	163 770	27,0	4 421 790,0
5.	от 30 до 49	363 630	39,5	14 363 385,0
6.	от 50и старше (69)	78 523	59,5	4 672 118,5
Всего		$N=\sum_{i=1}^{i=6} f_i = 856$	–	28 437 516,5
Средний возраст осужденных $\overline{x} \approx 33,2 \ (\text{года})$				

Средний возраст осужденных и каждой возрастной группы долю в общем числе $N = 856\,295$ всех осужденных вычислим по-другому.

Как известно, доля или удельный вес ω_i ($i = 1,2,3,4,5,6$) возрастных групп в общем количестве осужденных вычисляется по формуле:

$$\omega_i = \frac{f_i}{N}, \qquad N = \sum_{i=1}^{i=6} f_i,$$

а средний возраст осужденных находится по формуле:

$$\overline{x} = \sum_{i=1}^{i=6} x_i \cdot \omega_i .$$

Используя данные табл. 3.4, последовательно рассчитываем:

$$\omega_1 = \frac{f_1}{N} = \frac{17\,155}{856\,295} \approx 0,020 . \qquad \omega_2 = \frac{f_2}{N} = \frac{36\,907}{856\,295} \approx 0,043 ,$$

$$\omega_3 = \frac{f_3}{N} = \frac{196\,310}{856\,295} \approx 0,229 , \qquad \omega_4 = \frac{f_4}{N} = \frac{163\,770}{856\,295} \approx 0,191 ,$$

$$\omega_5 = \frac{f_5}{N} = \frac{363\,630}{856\,295} \approx 0,425 , \qquad \omega_6 = \frac{f_6}{N} = \frac{78\,523}{856\,295} \approx 0,092 ,$$

$$\overline{x} = \sum_{i=1}^{i=6} x_i \cdot \omega_i = x_1 \cdot \omega_1 + x_2 \cdot \omega_2 + x_3 \cdot \omega_3 + x_4 \cdot \omega_4 + x_5 \cdot \omega_5 + x_6 \cdot \omega_6 =$$

$$= 0{,}29 + 0{,}71 + 4{,}81 + 5{,}16 + 16{,}77 + 5{,}46 = 33{,}2 \text{ (года)}.$$

Исходные данные и результаты вычислений записываем во вторую итоговую табл. 3.6.

Таблица 3.6.

Результаты вычислений

№№ п/п, i	Возраст осужденных, годы	Количество осужденных (частота), f_i	Удельный вес i-й возрастной группы, $\omega_i = \frac{f_i}{N}$	Удельный вес i-й возрастной группы, $\omega_i \cdot$ 100 %	Произведение, $x_i \cdot \omega_i$
1.	от 14 до 15	17 155	0,020	2,0	0,29
2.	от 16 до 17	36 907	0,043	4,3	0,71
3.	от 18 до 24	196 310	0,229	22,9	4,81
4.	от 25 до 29	163 770	0,191	19,1	5,16
5.	от 30 до 49	363 630	0,425	42,5	16,77
6.	от 50 и старше (69)	78 523	0,092	9,2	5,46
Всего		$N = \sum_{i=1}^{i=6} f_i = $ $= 856\ 295$	$\sum_{i=1}^{i=6} \omega_i = \frac{1}{N} \cdot \sum_{i=1}^{i=6} f_i = $ $= 1$	Сумма=100%	$\sum_{i=1}^{i=6} x_i \cdot \omega_i \approx$ $\approx 33,2$ (лет)
Средний возраст осужденных $\overline{x} \approx 33,2$ (*года*)					

Как и должно быть, результаты для величины среднего возраста осужденных, полученные разными способами, совпадают. Кроме того, из табл. 3.6 видно, что наибольшая доля осужденных приходится на пятую $i = 5$ возрастную группу (30 - 49 лет); эта доля составляет 0,425 от единицы или 42,5 %.

Как видно из выше приведенной таблицы, число осужденных в возрасте от 30 до 49 лет составляет 363 630 человек, что составляет 42,5% от всей части (доли) осужденных. Чтобы определить объем выборки для установления этой доли в результате выборочных исследований, следует провести статистическую выборку.

Если задать предельную ошибку равную 1% при доле этой части осужденных равной 42,5 %, то согласно таблице А.Я. Боярского, достаточно провести опрос (объем выборки) среди от 9 600 до 9 900 всех осужденных, т.е. примерно объем выборки равен 9750. Однако, как правило, точный расчет этой величины немного ниже, что в силу естественного закона больших чисел собственно только увеличивает надежность (достоверность) выборочных исследований при намного меньших затратах по сравнению со сплошным исследованием.

Вопросы для самоконтроля:

1. Понятие статистического наблюдения.
2. Формы статистического наблюдения.
3. Организация официального статистического учета в России.
4. Формы статистического федерального наблюдения.
5. Непрерывное статистическое наблюдение. Регистры.
6. Сплошное и несплошное статистическое наблюдение.
7. Понятие и недостатки выборочного наблюдения.
8. Понятие «репрезентативная выборка».
9. Виды отбора, обеспечивающие репрезентативность выборки.
10. Понятие выборочной совокупности.
11. Основная задача выборочного наблюдения.
12. Оценка достоверности результатов статистического исследования.
13. Оценка предельной ошибки с помощью таблицы А.Я. Боярского.
14. Формулы, позволяющие найти ошибки репрезентативности статистических данных и оценить объем выборки.

4. ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ ДЛЯ ПРОВЕДЕНИЯ СТАТИСТИЧЕСКОЙ СВОДКИ И ГРУППИРОВКИ

Статистическое наблюдение – первый этап статистического исследования, дает огромное количество сведений, которые отражены в статистических формулярах. В первоначальном виде результат статистического наблюдения представляет, в лучшем случае, реляционную базу данных, реализованную на компьютере, в худшем – неупорядоченное множество заполненных бумажных носителей – статистических формуляров.

На втором этапе выполнения статистического исследования этот неупорядоченный массив информации необходимо соответствующим образом группировать, т.е. преобразовать так, чтобы первичные данные были бы отобраны по одному из реквизитов заполненного формуляра.

Выглядеть такая группировка на компьютере будет в виде таблиц, сформированных по значению выбранного из формуляра реквизита, например, наименованию отдела ОВД. Бумажные носители могут выглядеть в виде альбомов или ящиков, наполненных формулярами, отобранными из общей массы формуляров по значению выбранного реквизита.

Дальнейшая группировка сведется к формированию отдельных папок, состоящих из множества формуляров в пределах каждого журнала (для компьютерной базы каждой таблицы) по выбранным значениям реквизитов, например, «категория преступлений». Это позволит осуществить подсчет общих итогов путем подсчета количества строк в таблицах или формуляров в папках.

Для бумажной группировки при дальнейшей детализации может возникнуть проблема, когда один и тот же формуляр необходимо будет отнести к разным папкам. Компьютерная база данных предусматривает выполнение группировки – запроса каждый раз при вводе нового условия поиска формуляров по значению искомого реквизита. При этом количество поисковых реквизитов может быть любым. Следует отметить, что чем больше реквизитов будет отражено в запросе, тем меньше формуляров будет найдено в базе данных. В терминологии баз данных формуляр может называться записью.

Собранные в компьютерной базе данные могут быть преобразованы в формы отчетности с выполнением соответствующих подсчетов итоговых значений. Такой способ группировки называется машинным.

Ручной способ группировки предполагает выполнение задаваемых пользователем условий поиска записей (формуляров) в компьютерной базе данных.

Группировки позволяют выявить качественно однородные совокупности, раскрывать структуру совокупностей, наблюдать структурные сдвиги в зависимости от различных показателей и исследовать взаимосвязи между юридически значимыми показателями и социальными явлениями. Группировки могут быть представлены тремя видами:

Типологическая – расчленение совокупности явлений на отдельные качественно однородные группы.

Вариационная – это группировка, в результате которой типически однородные группы распределяются по количественным признакам, которые могут изменяться во времени.

Аналитическая – проводится для изучения взаимосвязи и зависимости изучаемых показателей.

При возникновении необходимости преобразовать группировку, построенную на основе количественных признаков в качественно однородные, а также при проведении двух и более группировок с различными интервалами к одной сопоставимой или при образовании укрупненных групп, в которых яснее проявляются реальные тенденции, используют вторичные группировки.

Вторичные группировки осуществляются путем сглаживания, укрупнения и смыкания ряда дробных показателей.

– **Сглаживание** рядов динамики выполняется путем вычисления средних показателей из данных первичной группировки. После чего ряд приобретает плавный, сглаженный вид. В качестве примера рассмотрим группировку, полученную в результате сглаживания динамического ряда преступности по среднеарифметическим данным, что устраняет случайные колебания в отдельные годы и выявляет главную тенденцию сокращения или роста преступных проявлений в регионе (рис. 4.1). По ежегодным данным состояния преступности вычисляем среднее значение за пятилетку. По данным средних значений строим график и наблюдаем сглаженную динамику роста преступности[1].

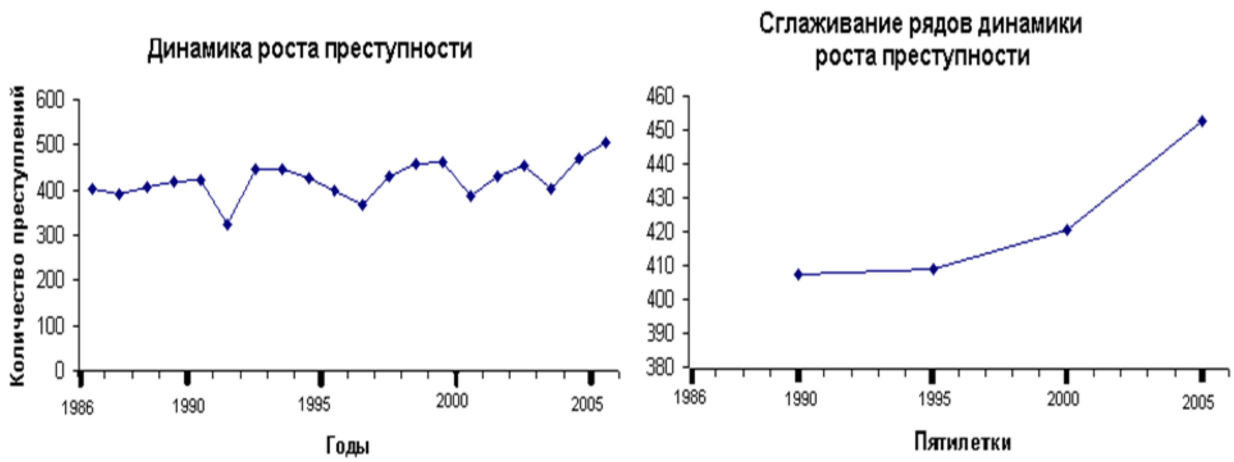

Рис. 4.1. Вторичная группировка путем сглаживания ряда

Сглаживание динамического ряда можно выполнить путем вычисления скользящей средней. Для вычисления скользящей средней сначала определяются укрупненные периоды, затем вычисляются средние значения членов укрупненных рядов, начиная с первого, затем со второго, третьего и т.д.

Полученные средние значения как бы скользят по динамическому ряду, шаг за шагом передвигаясь на минимальный временной интервал от начала до конца ряда.

---

[1] *Лунев В.В.* Юридическая статистика: учебник. - 2-е изд. перераб. и доп. - М.: Юрист, 2005.

В качестве примера рассмотрим тот же динамический ряд роста преступности с 1990 года (табл. 4.1). График этой вторичной группировки изображен на рис. 4.2.

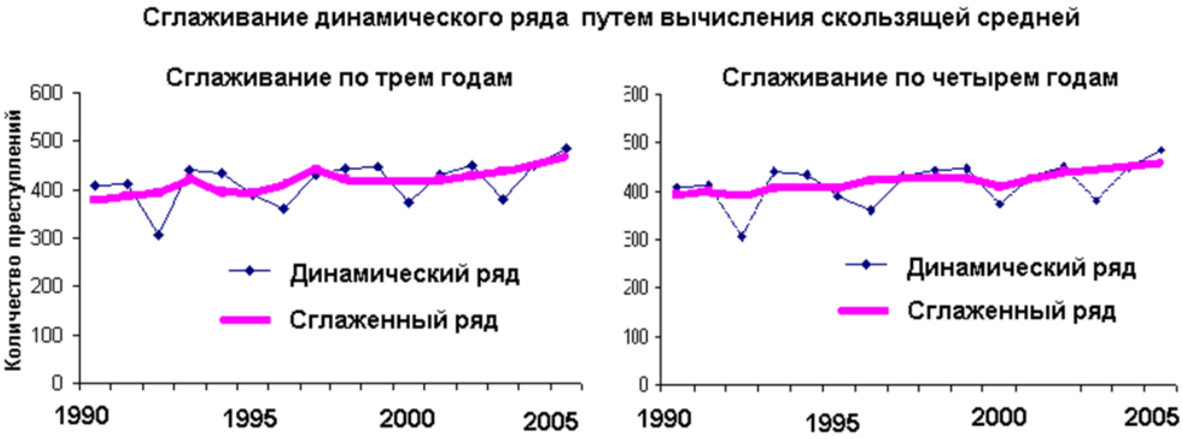

Рис. 4.2. Сглаживание ряда путем вычисления скользящей средней

Для расчета сглаживания путем вычисления скользящей средней можно использовать таблицы:

Таблица 4.1.

Вычисление скользящей средней в таблице

Годы	Количество преступлений	Сглаживание по трем годам	Сглаживание по четырем годам
1990	410		
1991	412		
1992	305	375,7	
1993	440	385,7	391,8
1994	435	393,3	398
1995	390	421,7	392,5
1996	360	395	406,3
1997	430	393,3	403,8
1998	444	411,3	406
1999	448	440,7	420,5
2000	374	422	424
2001	432	418	424,5
2002	450	418,7	426
2003	380	420,7	409
2004	450	426,7	428
2005	485	438,3	441,3

– **Укрупнение** предполагает суммирование данных за продолжительные отрезки времени.

Например, годовые сведения о преступности суммируются по годам, что при последовательном укрупнении показателей преступности на каждом этапе может дать реальный совокупный прирост (рис. 4.3).

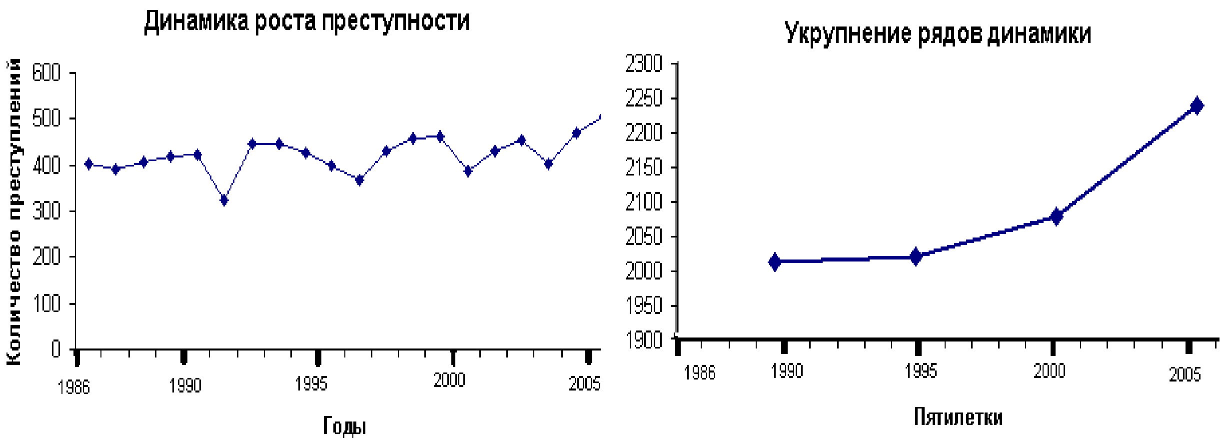

Рис. 4.3. Вторичная группировка, осуществляемая путем укрупнения

– **Смыкание** рядов динамики применяется при наличии несопоставимости анализируемых показателей.

Например, раньше преступность учитывалась в уголовных делах, а сейчас учитывается в преступлениях. Тогда «за базу» берут год, за который получены данные в прежнем измерении, а затем в новом. Каждый объем представляют за 100% и от него вперед и назад строится динамический.

Таблица 4.2

Смыкание рядов динамики

Периоды времени	**1**	**2**	**3**	**4**	**5**	**6**
Измерение преступности количеством преступлений	320	328	336	340		
Измерение преступности количеством осужденных				432	358	302
СМЫКАНИЕ (Количество осужденных)	406	416	426	432	358	302

Коэффициент перевода 432/340=1,27
П1=320*1,27=406;
П2=328*1,27=416;
П3=336*1,27=426

Полученные данные будут не точными, но они более или менее правильно отражают статистические закономерности.

Статистика располагает и более сложными приемами преобразования, которые используют достаточно сложные математические методы, но их рассмотрение требует специальной подготовки[1].

[1] *Лунев В.В.* Юридическая статистика: учебник. – 2-е изд. перераб. и доп. - М.: Юрист, 2005.

Вопросы для самоконтроля:

1. Содержание второго этапа статистического исследования.
2. Отличие машинной обработки данных от ручной при осуществлении статистической сводки.
3. Возможности статистической группировки.
4. Виды базовых статистических группировок.
5. Сглаживание рядов динамики при вторичной группировке.
6. Смыкание рядов динамики при вторичной группировке.
7. Укрупнение рядов динамики при вторичной группировке.

5. ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ ПРИ ПРОВЕДЕНИИ СТАТИСТИЧЕСКОГО АНАЛИЗА

Статистический анализ – расчленение явления или процесса на составные части, их количественное измерение, установление между ними взаимосвязей. Основными функциями статистического анализа являются:

- организационно-практическая – позволяет влиять на законодательную, исполнительную и судебную власти;
- описательная – позволяет оперировать качественно-количественными характеристиками исследуемого явления, описать его составные части и установить их соотношение;
- объяснительная – позволяет проникнуть вглубь изучаемого процесса или явления, выявить его внутренние и внешние взаимосвязи, установить причину происходящих событий;
- прогностическая – позволяет сделать многовариантный, вероятностный прогноз развития изучаемого явления.

Самый простой метод осуществления прогноза – графическая экстраполяция. Этот способ предполагает продолжение линии графика до следующих значений ряда. В данном примере значение показателя «количество преступлений в России» определено в 2302 тыс. Реальное значение показателя оказалось равным 2206,2 тыс., при этом ошибка прогноза всего 4,3%.

Рис. 5.1. Прогноз с помощью графической экстраполяции на 2013 г.

В основу теории статистического анализа положена теория вероятностей.

5.1. Теория вероятностей – основа математического прогнозирования

Теория вероятностей – это математическая наука, изучающая закономерности в случайных явлениях.

Если в обыденных представлениях, в житейской практике считается, что случайные события представляют собой нечто крайне редкое, идущее вразрез установившемуся, закономерному развитию событий, то в теории вероятностей мы откажемся от этих представлений. Случайные события, как они понимаются в теории вероятностей, обладают рядом характерных особенностей, в частности, все они происходят в массовых явлениях.

Во всех случаях, когда применяются вероятностные методы исследования, их цель состоит в том, чтобы, минуя слишком сложное (и часто практически невозможное) изучение отдельного явления, обусловленного очень большим количеством факторов, обратиться непосредственно к законам, управляющим массами случайных явлений. Изучение этих законов позволяет не только осуществить научный прогноз в своеобразной области случайных явлений, но в ряде случаев помогает целенаправленно влиять на ход случайных явлений, контролировать их, ограничивать сферу действии случайности.

Теория вероятностей – это наука, изучающая количественные закономерности однородных случайных событий массового характера и разрабатывающая методы количественной оценки влияния случайных факторов на различные события.

Классическое определение вероятности основано на равновозможности или равновероятности элементарных исходов в некотором опыте. Например, при бросании игральной кости, которая имеет точную форму куба и изготовлена из однородного материала, равновероятными элементарными исходами будут выпадение какого-либо определенного числа очков (от 1 до 6), обозначенного на гранях этого куба, поскольку в силу наличия симметрии ни одна из граней не имеет объективного преимущества перед другими.

В общем случае рассмотрим полную группу, состоящую из конечного числа элементарных, равновозможных несовместных событий (исходов) некоторого опыта. Такую группу называют группой возможных результатов опыта или испытания. Те из возможных результатов опыта, на которые подразделяется событие А, называют результатами опыта, благоприятствующими А.

Классическое определение вероятности события формулируют следующим образом:

вероятность P(A) события A равна отношению числа возможных результатов опыта (M), благоприятствующих событию A, к числу всех возможных результатов опыта (N):

$$P(A)=M/N.$$

Вычисление вероятности по формуле $P(A)=M/N$ вызывает в некоторых случаях затруднения при определении значений M и N. Эти затруднения связаны, в частности, с тем, что при решении ряда задач требуется применение формул из комбинаторики. Поэтому полезными являются нижеследующие комбинаторные формулы.

Число всевозможных перестановок из n различных элементов равно n! (обозначает произведение всех целых чисел от 1 до n, т.е. $n!=l*2*3*\ldots*n$)

Статистический способ определения вероятности основан на предварительном проведении большого числа испытаний. При этом вероятность рассчитывают по формуле:

$$h(A) = k/L,$$

где k – число появления события А, называемое частостью А;

L – общее число событий, наступивших в некоторой серии испытаний при определенном неизменном комплексе условий;

h(A) – статистическая вероятность (частота) события.

При большом числе испытаний значение частоты h(A) стабилизируется и приближается к величине вероятности P(A). Однако статистическую вероятность события можно достоверно определить только при достаточно большом числе испытаний.

Понятие условной вероятности является основным инструментом теории вероятностей.

Вероятность события А в предположении, что уже произошло событие В, называют **условной вероятностью события А при условии В** и обозначают P(A|B).

<u>Например</u>, в урне два белых шара и один черный; два человека вынимают последовательно из урны по одному шару. Рассмотрим два события: В – появление белого шара у первого человека; А – появление белого шара у второго человека.

Тогда P(B)=2/3. Теперь вычислим P(A|B), $P(A|\overline{B})$.

P(A|B) – это вероятность того, что второй человек вытащит белый шар, при условии, что первый человек уже вытащил один белый шар. Так как в урне остался один белый шар, то P(A|B)=1/2.

$P(A|\overline{B})$ – это вероятность того, что второй человек вытащит белый шар при условии, что первый человек вытащил черный. Так как в урне осталось два белых шара и ни одного черного, то P(A|B) = 1. Сравнивая полученные вероятности, можно сделать вывод, что событие А зависит от события В и вероятность появления события А носит условный характер.

Теорема умножения:

вероятность произведения двух событий равна произведению вероятности одного из событий на условную вероятность другого при условии, что первое произошло.

$$P(A*B)=P(A)*P(B/A)=P(B)*P(A|B).$$

Из теоремы умножения следует формула для вычисления условной вероятности события $P(A|B)=P(A*B)/P(B)$.

Этой формулой можно пользоваться, если $P(B) \neq 0$. Если же $P(B)=0$ (т.е. событие B – невозможно), то по теореме умножения,

$P(A|B)=0$ и $P(A*B)=0$.

Таким образом,

$$P(A|B)=\begin{cases} \dfrac{P(A*B)}{P(B)}, \text{ если } P(B)\neq 0 \\ 0, \text{ если } P(B)=0 \end{cases}.$$

Таким образом, рассмотрение условных вероятностей при одном и том же данном событии B равносильно выбору B в качестве нового пространства элементарных исходов с вероятностями, пропорциональными первоначальным. Коэффициент пропорциональности P(B) необходим для того, чтобы сделать вероятность нового пространства равной 1. Все основные теоремы о вероятностях остаются справедливыми для условных вероятностей, взятых относительно некоторого фиксированного события B.

Например, при медицинском освидетельствовании лиц, претендующих на получение разрешения на вождение автомобиля, оказалось, что из N (общего числа претендентов) N_a страдают дальтонизмом, N_b женщины.

Решение. Пусть A и B означают события, состоящие соответственно в том, что случайно выбранное лицо страдает дальтонизмом или является женщиной. В частности, может оказаться необходимым найти вероятность того, что случайно выбранная женщина страдает дальтонизмом, т.е. найти:

$$P(A/B) = P(AB)/P(B),$$

где: P(A/B) – вероятность события A (дальтонизм) при условии, что произошло событие B (выбрана женщина);

P(AB) – вероятность того, что женщина страдает дальтонизмом;

P(B) – вероятность того, что случайно выбранное лицо из N – женщина.

Так, при $N = 240$ и $N_b = 30$;

$$P(B) = N_b/N= 0{,}125.$$

По данным генетических исследований известно, что для женщин $P(AB) = 0{,}0001$ (для мужчин вероятность быть дальтоником равна 0, 01).

Таким образом,

$$P(A/B) = 0{,}0001/0{,}125 = 0{,}0008.$$

Следует отметить, что если

$$P(B/A) = P(B) \text{ и } P(A/B) = P(A),$$

то события B и A взаимонезависимы.

Следствием теорем сложения и умножения вероятностей является формула полной вероятности.

Пусть требуется определить вероятность некоторого события A, которое может произойти вместе с одним из событий $H_1, H_2,\dots,H_n$, образующих полную группу несовместных событий. Будем эти события называть гипотезами.

В этом случае вероятность события А вычисляется по формуле полной вероятности как сумма произведений вероятности каждой гипотезы на условную вероятность события при этой гипотезе:

$$P(A)=\sum_{i=1}^{n} P(H_i)*P(A|H_i).$$

5.2. Индексный метод в статистическом анализе

Индексы являются одними из самых распространенных статистических показателей. С их помощью изучаются развитие правовой системы государства в целом, ее отраслей, функционирование системы учреждений, осуществляющих правоохранительную деятельность. Индексы характеризуют изменение важнейших статистических показателей и исследуют роль и влияние отдельных факторов, определяющих эти изменения.

Слово «индекс» (index) в переводе с латинского языка означает показатель. Индексы прежде всего – относительные показатели, отражающие соотношение во времени (динамический индекс) или в пространстве (территориальный индекс) социально-правовых и экономических явлений или процессов[1].

Индекс – это любой обобщающий показатель двух и более совокупностей, состоящих из элементов, которые нельзя соизмерять, характеризующих изучаемое явление[2]. Индекс может показать, во сколько раз уровни исследуемых одинаковых явлений отличаются друг от друга в зависимости от определенных условий их проявления.

По степени охвата явления индексы бывают индивидуальные и общие.

Индивидуальные индексы – это результат сравнения двух единичных показателей, относящихся к одному объекту или изменение одного элемента совокупности, например, изменение количества обращений граждан в правоохранительный орган за год.

Для сравнения выбирается базисный показатель. Для данного примера сравнивается количество обращений в отчетном периоде по сравнению с базовым или изменение уровня полезности обращений, что важно для возбуждения уголовных дел, также с базовым периодом. В первом случае индекс будет количественным показателем, а во втором качественным.

Индивидуальные индексы представляют собой относительные величины динамики, выполнения плана, сравнения, и их расчет не требует знания специальных правил.

---

[1] Правовая статистика: учебник для студентов вузов, обучающихся по специальности «Юриспруденция», для курсантов и слушателей образовательных учреждений МВД / В.С. Лялин, Е.А. Костыря, А.В. Симоненко, Е.И. Кузнецова, Е.Н. Барикаев. - 2-е изд., перераб. и доп. - М.: ЮНИТИ-ДАНА, 2010. С. 216.

[2] *Лукьянов Д.Б.* Правовая статистика. Курс лекций: учебное пособие. - Белгород: ОНиРИД БелЮИ МВД России, 2008.

Таблица 5.1.

Количество обращений граждан в правоохранительные органы

Наименование	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Число обращений, млн.	12,4	13,56	14,2	15,12	15,98	16,38	17,39	18,64
Число зарегистрированных преступлений, тыс.	2596,9	2474,3	2208,7	2064,6	1810	1646,4	1554,8	1489,6
Количество заявлений на 1 преступление	4,8	5,5	6,4	7,3	8,8	9,9	11,2	12,5

Если выбрать 2006 год базисным, то индивидуальный индекс будет рассчитываться отношением текущего периода к показателю 2006 г. Рассчитать можно как качественные, так и количественные индексы (табл. 5.2).

Расчет количественного индивидуального индекса:13,56/12,4=1,09, т.е. число обращений граждан увеличилось на 9%.

Расчет качественного индивидуального индекса: 5,5/4,8=1,14, т.е. количество заявлений на одно преступление увеличилось на 14%.

Таблица 5.2.

Расчет индивидуальных индексов

	2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Количественный	база	1,09	1,15	1,22	1,29	1,32	1,4	1,5
Качественный	база	1,14	1,33	1,52	1,83	2,06	2,33	2,6

Общий индекс характеризует изменение всех показателей сложного явления. Общие индексы могут быть агрегатными индексами или средними взвешенными.

При расчете агрегатного индекса (aggregatus означает «складываемый, суммируемый»), применяются коэффициенты-измерители с помощью которых можно выразить итоговые показатели несоизмеримых совокупностей в разные временные промежутки или на разных территориях.

Примером агрегатного индексирования может быть оценивание всей совокупности каких-либо изделий, количество которых будет индексированной величиной, а их цена – коэффициентом-измерителем.

Если в совокупности несколько наименований продукции и у каждой есть свой объем, то применяют средние взвешенные арифметические и гармонические индексы. Средневзвешенный индекс характеризует в среднем изменение несоизмеримых показателей.

Примером среднего взвешенного индекса может служить валовый внутренний продукт (ВВП), состоящий из суммы стоимости всех товаров и услуг.

В правовой статистике примером среднего взвешенного может быть социальная опасность преступлений, которая состоит из среднего значения всей совокупности преступлений с соответствующими коэффициентами-измерителями.

Статья 15 УК РФ определяет 4 категории преступлений: преступления небольшой, средней тяжести, тяжкие и особо тяжкие.

Было бы неправильным оценивать и сравнивать преступность по количеству совершенных преступлений на определенной территории или за определенный период времени. У преступлений разная степень общественной опасности.

Для выбора коэффициента-измерителя используют срок лишения свободы, предусмотренный уголовным кодексом. Критериями оценивания тяжести преступлений можно взять санкции закона и назначенный судом срок лишения свободы.

Наиболее справедливым подходом для выбора коэффициента-измерителя следует выбирать то, что реально соответствует общественной опасности преступлений. Поэтому оценивать общественную опасность преступлений по срокам лишения свободы будет не совсем правильно. Ведь суд в рамках закона может варьировать сроками наказания. Кроме этого, лицо может проходить по нескольким статьям Уголовного кодекса.

Для оценки социальной опасности преступлений важен сам факт преступления, поэтому рекомендовано в качестве коэффициента-измерителя использовать максимальный срок по категории преступления.

Таким образом, коэффициент-измеритель для преступлений небольшой тяжести составит 3, средней – 5, тяжких преступлений – 10 и преступлений особой тяжести – 20.

Далее для расчета индекса социальной опасности преступлений (СИТП) из всей совокупности преступлений следует выбрать количественные суммарные показатели, соответствующие категориям тяжести преступлений и умножить на соответствующий коэффициент-измеритель. Для получения итогового значения результаты умножения сложить. Формула для расчета этого индекса выглядит следующим образом:

$$\text{СИТП} = \frac{\sum \text{П}_{\text{ТП}} \cdot \text{Б}}{\sum \text{П}_{\text{БП}} \cdot \text{Б}}$$

где $\text{П}_{\text{ТП}}$ – преступления текущего периода по всем 4 видам, Б – коэффициенты-измерители, $\text{П}_{\text{БП}}$ – преступления базового периода по всем 4 видам.

Пример. Предположим в городе N. в базовом 2010 году совершено 20 особо тяжких преступлений, 15 тяжких, 100 средней тяжести, 150 небольшой тяжести, – всего 285 преступлений. В 2011 году соответственно 10, 10, 120, 160 – всего 300 преступлений. Подставим значения в формулу:

СИТП = (10*20+10*10+120*5+160*3)/(20*20+15*10+100*5+150*3) =1380/1440=0.95

Из данного примера видно, что, несмотря на увеличение количества преступлений в 2011 году по сравнению с 2010 годом, совокупный индекс тяжести составил 95%. Это свидетельствует о снижении степени общественной опасности преступлений.

В сводных статистических отчетах по России не всегда давались показатели по всем категориям преступлений. Можно было встретить показатель «удельный вес тяжких и особо тяжких преступлений» В этом случае, располагая данными о «преступлений всего», можно установить числовые значения для двух комбинированных категорий преступлений. Социальная опасность преступлений, совершенных на территории России в период с 2006 по 2012 годы, составила (табл. 5.3):

Таблица 5.3.

Расчет индекса социальной опасности преступлений

Наименование	2006	2007	2008	2009	2010	2011	2012	2013	2006
Социальный индекс тяжести преступлений	17,6	16,5	14,5	13,7	11,9	10,7	10,0	9,5	17,6

В настоящее время индексный метод успешно развивается в трудах многих отечественных статистиков и получил широкое применение в практике статистических работ, в основном применительно к экономической статистике, в известной мере и в правовой статистике. Индексные показатели вычисляются на высшей ступени статистического обобщения и опираются на результаты сводки и обработки данных статистического наблюдения[1].

5.3. Статистические взаимосвязи и причинности

Некоторые явления в природе и обществе, называемые причинными признаками-факторами, могут обусловливать появление других явлений – признаков-следствий, т.е. обладать наличием причинно-следственной связи. В этом случае говорят о корреляционной связи, которая определяется с определенной долей вероятности.

Раздел статистики, в котором изучаются взаимосвязи социальных явлений, называется теорией корреляции и регрессии. Основоположниками этой теории являются английские ученые Ф. Гальтон, выдвинувший идею о статистических измерениях в психологии, и К. Пирсон, создавший математический анализ[2].

Корреляционный анализ обеспечивает измерение тесноты и направления связи. Регрессионный анализ оценивает функциональную связь и зависимость определенного значения результативного следственного признака от факторных признаков. Измерению подвергаются и те, и другие.

Корреляционный анализ начинается с того, что необходимо определить факторы, которые могут влиять на результативный показатель. Эти факторы должны выполнять следующие условия: быть измеримы, их количество долж-

---

[1] Классификация индексов и особенности их применения в уголовно-правовой статистике: лекции [Электронный ресурс]. Режим доступа: http://studentu-vuza.ru/pravovaya-statistika.

[2] Правовая статистика: учебник для студентов вузов, обучающихся по специальности «Юриспруденция», для курсантов и слушателей образовательных учреждений МВД / В.С. Лялин, Е.А. Костыря, А.В. Симоненко, Е.И. Кузнецова, Е.Н. Барикаев. - 2-е изд., перераб. и доп. - М.: ЮНИТИ-ДАНА, 2010. С. 174.

но быть большим, и они находятся в причинно-следственной связи с результативным показателем.

При сборе информации нужно быть уверенным в ее достоверности и однородности. После чего выбирают математическое уравнение, которым отражают сущность исследуемой зависимости.

Если связь многофакторная, то составляется матрица, в ячейки которой заносятся результаты наблюдений. Обработка матрицы осуществляется на компьютере с использованием довольно сложного математического аппарата.

Наиболее простым способом расчета корреляционной зависимости является парная корреляция – связь одного признака-фактора с одним признаком-следствием.

Для определения силы или плотности парной корреляционной связи устанавливают коэффициент корреляции, который изменяется от 0 до 1 или (-1) при обратной связи.

Если значение коэффициента близко к 0, то это означает, что линейная связь между признаками отсутствует. При значениях меньших 0,3 – связь слабая, от 0,3 до 0,5 – умеренная. При значениях коэффициента корреляции от 0,5 до 0,7 связь значительная, при значениях от 0,7 до 0,9 – сильная, а при коэффициенте близком к 1 связь очень сильная.

Для изучения корреляционных связей разработаны разные методы, каждый из которых решает свои конкретные задачи. Одни коэффициенты пригодны для измерения взаимосвязи качественных признаков, другие – для количественных. Это коэффициенты А.А.Чупрова, К.Пирсона, Фехнера, Кендалла и Спирмена[1].

Наиболее простыми и наглядными являются способы определения парного коэффициента корреляции по Фехнеру и ранговой корреляции по Спирмену. Выполним расчеты указанными способами и сравним полученный результат.

Поставим вопрос: верно ли утверждение, что рост количества обращений граждан в правоохранительные органы (признак-фактор) приводит к увеличению количества уголовных дел (признак-следствие).

Используем метод Фехнера, основанный на сравнении параллельных рядов, представляющих количество обращений о преступлениях и правонарушениях в выбранные условные подразделения ОВД и количества возбужденных в этих подразделениях уголовных дел. Данные поместим соответственно в столбцы 3 и 4 табл. 5.4.

[1] *Лунев В.В.* Юридическая статистика: учебник. – 2-е изд. перераб. и доп. - М.: Юрист, 2005.

Таблица 5.4.

Определение коэффициентов корреляции по Фехнеру и Спирмену

N	Наименование отдела	К-во обращений в ОВД с заявлением о преступлении *(x)*	К-во возбужденных уголовных дел по обращениям *(y)*	по Фехнеру		по Спирмену			
				Знаки отклонения от средней		**Ранги по признакам**		**Разность рангов**	
				Х	***У***	***Х***	***У***	***d***	***d^2***
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	ОВД1	120	12	-	-	7	9	-2	4
2	ОВД2	312	21	+	+	1	2	-1	1
3	ОВД3	104	15	-	-	9	6	3	9
4	ОВД4	111	13	-	-	8	8	0	0
5	ОВД5	98	10	-	-	10	10	0	0
6	ОВД6	142	14	-	-	6	7	-1	1
7	ОВД7	153	20	-	+	4	3	1	1
8	ОВД8	147	16	-	-	5	5	0	0
9	ОВД9	225	19	+	+	3	4	-1	1
10	ОВД10	231	22	+	+	2	1	1	1
Среднее арифметическое		**164,3**	**16,2**	Число совпадений знаков С=9 Число несовпадений знаков Н=1				Сумма квадратов разности рангов $\sum d^2=18$	

После подсчета среднего арифметического количества правонарушений и количества преступлений осуществляется сравнение среднего арифметического с соответствующим реальным значением в столбцах 3 и 4.

Если реальное значение больше среднего, то в столбцах 5 и 6 против реального значения ставится знак «плюс», меньше – «знак минус». Коэффициент Фехнера рассчитаем по формуле:

К$_{(\text{по Фехнеру})}$=(С-Н)/(С+Н)
К$_{(\text{по Фехнеру})}$=(9-1)/(9+1)=0,8

Если вопрос исследования поставить иначе: верно ли утверждение, что в том ОВД, где будет больше всего зарегистрировано сообщений о преступлениях и правонарушениях, будет больше всего возбуждено уголовных дел.

Коэффициент парной корреляции по Спирмену рассчитывается по тем же рядам 3, 4 (табл. 5.4.) по формуле:

$К_{(\text{по Спирмену})} = 1 - 6\sum d^2/(n(n^2-1))$

Для расчета необходимо определить номера мест от меньшего к большему значений количества обращений граждан и количество возбужденных уголовных дел, затем соответствующие номера мест записать в столбцы 7 и 8. Т.е. первое место в столбце 7 присваивается самому большему значению количества обращений, т.е. ОВД, а в столбце 8 - самому большему значению количества уголовных дел. Ранжировать можно и в обратном порядке.

К$_{(\text{по Спирмену}))}$=1- 6*18/(10(100-1))=0,89

Таблица 5.5.

Определение линейного коэффициента корреляции по Пирсону

N	Отделы	Обращения (x)	Уголовные дела (y)	Отклонение от средней		Суммарные значения для формулы		
				X	**У**	**X*Y**	**X^2**	**Y^2**
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	ОВД1	120	12	-44,3	-4,2	186,06	1962	18
2	ОВД2	312	21	147,7	4,8	708,96	21815	23
3	ОВД3ОВ	104	15	-60,3	-1,2	72,36	3636	1
4	Д4	111	13	-53,3	-3,2	170,56	2841	10
5	ОВД5	98	10	-66,3	-6,2	411,06	4396	38
6	ОВД6	142	14	-22,3	-2,2	49,06	497	5
7	ОВД7ОВ	153	20	-11,3	3,8	-42,94	128	14
8	Д8	147	16	-17,3	-0,2	3,46	299	0
9	ОВД9	225	19	60,7	2,8	169,96	3684	8
10	ОВД10	231	22	66,7	5,8	386,86	4449	34
		среднее 164,3	среднее 16,2			сумма 2115,4	сумма 43708	сумма 152

Несколько сложнее выглядит линейный корреляционный анализ по методу Пирсона. Формула расчета коэффициента корреляции построена таким образом, что если связь между признаками имеет линейный характер, коэффициент Пирсона точно устанавливает тесноту этой связи[1].

$$К_{(\text{по Пирсону})} = \frac{\sum(x_i - \bar{X}) * (y_i - \bar{Y})}{\sqrt{\sum(x_i - \bar{X})^2 * \sum(y_i - \bar{Y})^2}}$$

Используем данные столбцов 7, 8, 9 (табл. 5.5), подставим в формулу для расчета линейного коэффициента Пирсона.

Подставив полученные значения в формулу, получим:

$$К_{(\text{по Пирсону})} = \frac{2115{,}4}{\sqrt{43708 * 152}} = \frac{2115{,}4}{2574} = 0{,}82$$

Оценивая полученные коэффициенты, попадающие в условие от 0,7 до 0,9, можем установить, что корреляционная связь сильная, причем при расчете коэффициента корреляции по Спирмену она приближается к очень сильной.

---

[1] *Нижегородцева Н.В., Мишина Т.В.* Методические рекомендации по написанию и оформлению курсовой и выпускной квалификационной работы по психологии и конфликтологии / Ярославский государственный педагогический университет им. К.Д. Ушинского 2006 г. [Электронный ресурс]. Режим доступа: http://cito-web.yspu.org/link1/metod/met125/ node35.html.

Сущность статистических методов исследования зависимости состоит в том, чтобы по имеющимся числовым данным извлечь определенные закономерности развития системы. Но даже при наличии сильной статистической связи между двумя показателями нельзя быть уверенным в их причинно-следственной обусловленности, так как может существовать третья причина, которая и определяет их статистическую взаимосвязь.

Кроме того, в отдельных случаях отсутствие статистической связи не говорит об отсутствии причинной, а заставляет искать другие пути и средства их выявления, если содержательная концепция и практический опыт указывают на ее возможное существование.

Итак, корреляция - это вероятная, или статистическая зависимость, которая возникает тогда, когда зависимость одного признака от другого осложняется наличием ряда случайных факторов.

При корреляционной зависимости результативный признак под действием факторного признака принимает ряд значений, варьирующих около их средней величины. Это происходит потому, что на значение результативного признака влияет не только изменение факторного признака, но и действия каких-то других, неизвестных нам факторов. В связи с этим корреляционная зависимость проявляется не в единичном случае, а в массе[1].

Вопросы для самоконтроля:

1. Понятие «статистический анализ».
2. Функции статистического анализа.
3. Суть графической экстраполяции.
4. Классическое определение вероятности события.
5. Статистический способ определения вероятности.
6. Понятие условной вероятности.
7. Теорема умножения.
8. Полная вероятность наступления события.
9. Понятие и виды индексов.
10. Расчет агрегатного индекса.
11. Расчет социального индекса тяжести преступлений, выбор коэффициента-измерителя.
12. Понятие корреляционной связи. Обратная связь.
13. Понятие парной корреляционной связи.
14. Этапы корреляционного анализа.
15. Соответствие значений коэффициента корреляции тесноте связи.
16. Порядок и цели расчета коэффициента корреляции по Фехнеру.
17. Порядок и цели расчета коэффициента корреляции по Спирмену.
18. Порядок и цели расчета коэффициента корреляции по Пирсону.

[1] *Лунев В.В.* Юридическая статистика: учебник. - 2-е изд. перераб. и доп. - М.: Юрист, 2005.

СПИСОК ЛИТЕРАТУРЫ

1. Федеральный закон от 29 ноября 2007 года № 282-ФЗ ст. 5 «Об официальном статистическом учете и системе государственной статистики в Российской Федерации» (с изм. 23.07.2013 № 251-ФЗ) [Электронный ресурс]. Режим доступа: http://base.consultant.ru/cons/cgi/online.cgi?req=doc;base =LAW;-n=166225; fld=134;dst=101782;rnd= 0.5816348623484373.
2. Приказ Федеральной службы государственной статистики от 26.02.2009 г. № 34 «Об утверждении статистического инструментария для организации статистического наблюдения за деятельностью следственных органов и органов дознания, рассмотрением заявлений и сообщений о преступлении» ГАРАНТ.РУ: http://www.-garant.ru/products/ipo/prime/doc/12065569/#ixzz3G7Cfg2tc.
3. Приказ Генеральной прокуратуры РФ от 02.07.2012 г. № 250 «Об утверждении форм федерального статистического наблюдения № 1-ЕГС, № 2-ЕГС, № 3-ЕГС, № 4-ЕГС» / Режим доступа: ГАРАНТ.РУ: http://www.garant.ru/ products/ipo/prime/doc/70130276/#ixzz3GCjQKc8m.
4. Приказ Генеральной прокуратуры РФ, МВД РФ, МЧС РФ, МИНЮСТ РФ, ФСБ, Минэкономразвития и торговли РФ, ФС РФ по контролю за оборотом наркотиков от 29.12.2005 г. №№ 39/1070/1021/253/780/353/399 «О Едином учете преступлений» // Российская газета. 2006. № 13.
5. Вероятностные разделы математики / под ред. Ю.Д. Максимова. – СПб.: «Иван Фёдоров», 2001. С. 400. 592 с.
6. *Дрога А.А., Копонев Д.Н.* Выборочное наблюдение в правовой статистике: учебное пособие. – Белгород: Бел ЮИ МВД России, 2014. 44 с.
7. *Дрога А.А., Лукьянов Д.Б., Прокопенко А.Н.* Информатика и математика: учебное пособие. - Белгород: Бел ЮИ МВД России, 2008.
8. *Елисеев С.А.* Зарубежная криминологическая мысль о причинах имущественной преступности: очерк истории // Вопросы уголовного права. Иркутск. Сибирский юридический вестник. 2012. № 1.
9. *Лукьянов Д.Б.* Причины снижения статистических показателей преступности в России // Сборник научных трудов Белгородского юридического института МВД России. Вып. 2. – Белгород: Бел ЮИ МВД России, 2013. 169 с.
10. *Лунев В.В.* Юридическая статистика: учебник. – 2-е изд. перераб. и доп. – М.: Юрист, 2005.
11. *Орлов А.И.* Нечисловая статистика. – М.: МЗ-Пресс, 2004.
12. Правовая статистика: словарь терминов / сост. Д.Н. Копонев, А.А. Дрога. – Белгород: Бел ЮИ МВД России, 2013.
13. Правовая статистика: учебник для студентов вузов, обучающихся по специальности «Юриспруденция», для курсантов и слушателей образовательных учреждений МВД / В.С. Лялин, Е.А. Костыря, А.В. Симоненко, Е.И. Кузнецова, Е.Н. Барикаев. -2-е изд., перераб. и доп. – М.: ЮНИТИ-ДАНА, 2010. 255 с.

14. *Савюк Л.К.* Правовая статистика: учебник. – М.: Юристъ, 2006.

15. *Щербина Л.Е.* Общая теория статистики [Электронный ресурс]. Режим доступа: http://www.e-reading.me/bookreader.php/103879/Shcherbina_ Obshchaya_teoriya_ statistiki.html.

16. Энциклопедия экономиста. Статистика. Общая теория статистики [Электронный ресурс]. Режим доступа: http://www.grandars.ru.

УЧЕБНОЕ ИЗДАНИЕ

Копонев Дмитрий Николаевич;
Лукьянов Дмитрий Борисович;
Прокопенко Алексей Николаевич,
кандидат технических наук;
Чеканов Николай Александрович,
доктор физико-математических наук

Общие вопросы статистики

Учебное пособие

Редактор ***О.Н. Пендюрина***
Техн. редактор ***Т.Л. Ковалева***

Подписано в печать 28.11.2014 г., формат бумаги 60х90/16, уч.изд.л. 3,2,
бумага офсетная, печать трафаретная
Тираж экз., заказ №

Отпечатано в отделении полиграфической и оперативной печати
Белгородского юридического института МВД РФ
г. Белгород, ул. Горького, 71

9 785917 760259