

НАПРАВЛЕНИЕ I

МАТЕМАТИКА И УПРАВЛЕНИЕ

Итоги выборов. Анализ электоральных предпочтений

В. Е. Чуров, В. Л. Арлазаров, А. В. Соловьев

В статье изложен методологический подход к определению основных тенденций электоральных предпочтений на основе статистических данных. Приведены результаты анализа данных итоговых протоколов участковых избирательных комиссий на основании фактора территориальной принадлежности избирателей. Приведены выводы о дальнейших направлениях исследований электоральных предпочтений, а также краткое описание построения математической модели электоральных предпочтений.

Обозначения и сокращения

- ЧФ – человеческий фактор
- ИС – информационная система
- ЧС – чрезвычайная ситуация
- БД – база данных
- УИК – участковая избирательная комиссия
- ТИК – территориальная избирательная комиссия
- ОДП – оперативно-диспетчерский персонал
- ПО – программное обеспечение

Введение

Последнее время в публикации появились статьи, в которых делаются попытки обработать электоральную статистику с помощью математических методов и выявить очевидные и неочевидные тенденции, влияющие на проведение избирательного процесса [1–3].

Поскольку практически все эти исследования носят критический характер, попробуем и мы разобраться в том массиве статистических данных, которые накопились после очередных федеральных выборов.

Для начала определим, что мы собственно хотим исследовать и для чего. Обычно в статьях оппонентов дается анализ статистических данных по двум независимым параметрам, например, распределение количества участковых избирательных комиссий (УИК) в зависимости от явки и распределение голосов за кандидатов (партии) также в зависимости от явки. На основании анализа данных протоколов УИК, находящихся в открытом доступе [4], делается обычно один и тот же вывод: поскольку распределение (например, числа УИК в зависимости от явки) не соответствует нормальному закону распределения (закону больших чисел) относительно средней явки, то результаты выборов являются недостоверными.

Такой вывод можно было бы сделать, если бы мы имели дело с однородным массивом, как избирателей, так и УИК, которые, подчиняясь некоему случайному процессу, участвовали в выборах. Однако это не совсем так. Во-первых, сам процесс выборов отнюдь не случаен, так как слишком много факторов на него влияет, а во-вторых, сам массив, как избирателей, так и УИК отнюдь не однороден.

К тому же, исследуя различные социологические статистические зависимости (например, статистику автомобильных аварий, статистику доходов и т. д.), мы практически никогда не получим красивое нормальное распределение. Однако вывод о том, что результаты автомобильных аварий «подтасованы» явно некорректен.

Поэтому, более грамотным и глубоким подходом к анализу электоральной статистики был бы, на наш взгляд, следующий.

1. Построение модели электоральных предпочтений

Итак, что же мы хотим исследовать.

На наш взгляд, говорить об однородности избирателей в стране целом несколько некорректно, следовательно, мы должны построить модель и по ней исследовать «идеальные» электоральные предпочтения (политические предпочтения избирателей), сравнить их с реальными, полученными на выборах, и на основании этого сравнения делать вывод

о том, насколько они расходятся с «идеальным» образом среднего постсоветского избирателя.

Итак, мы хотим разбить все множество избирателей на некоторые однородные подмножества, для каждого подмножества выделить основные тенденции электоральных предпочтений и построить идеальные функции факторов, влияющих на эти предпочтения. Далее, на основании данных протоколов УИК можно построить реальную картину для каждого подмножества избирателей и сравнить ее с идеальной.

Затем, просуммируем все тенденции и предпочтения различных подмножеств избирателей и сравним полученную интегральную картину с общей реальной картиной на основании данных протоколов УИК на выборах.

Построение «идеальной» модели электоральных предпочтений начнем с выделения наиболее значимых параметров, на основании которых эти предпочтения могут формироваться. По нашему мнению, таких параметров несколько. Они могут быть отнесены условно к «основным» и «дополнительным».

К основным параметрам модели электоральных предпочтений отнесем следующие:

- 1) место проживания (принадлежность к определенной электоральной географии):
 - город/село, регион;
 - экономическое состояние региона;
 - социокультурные традиции (особенности электоральной культуры);
- 2) динамика экономического благосостояния:
 - динамика по группам населения (молодежь, работающие, пенсионеры и т. д.);
 - протестные настроения;
- 3) категории населения:
 - возраст/пол/трудовой статус и др.

На основании данных многочисленных исследований ВЦИОМ, ФОМ, РОМИР, ЦПТ, ЦЕССИ и др. можно построить картину электоральных предпочтений с точностью до города, района, отдельной категории населения и уж точно для региона в целом.

Конечно, протоколы УИК не содержат никаких сведений о возрасте, экономическом благосостоянии и т. д., но содержат однозначную привязку к региону, который во многом и характеризуется проживающим там

населением (вспомним хотя бы «красный пояс» регионов, традиционно поддерживающих КПРФ).

К дополнительным параметрам модели электоральных предпочтений можно отнести:

- 1) влияние СМИ в регионе (те же социологические опросы показывают, что около 97 % населения РФ получает информацию именно с помощью телевидения, при этом 17–20 % считают, что делают выбор при голосовании под влиянием СМИ [8]);
- 2) динамика политический предпочтений (статистика ВЦИОМ, ФОМ и др.);
- 3) работа партий в регионах (статистика деятельности региональных отделений партий);
- 4) влияние ошибок УИК/ТИК на достоверность информации (здесь можно использовать известные математические модели влияния усталостных ошибок, ошибок некомпетентности на основании данных об опыте работы сотрудников УИК/ТИК);
- 5) оценка влияния административного ресурса;
- 6) влияние вида выборов (регион/местные/федеральные и партии/персоналии), известно, например, что за персоналии голосуют охотнее, чем за партии (статистика ВЦИОМ).

Можно выделить еще много других параметров, однако получившаяся модель может оказаться сильно перегруженной, и расчеты по ней методами факторного дисперсионного или регрессионного анализа могут оказаться весьма громоздкими (см. [5]).

На наш взгляд вышеприведенных параметров модели вполне достаточно, чтобы построенная модель была бы достаточно адекватной реальной картине, а расчеты по ней достаточно точными, даже, несмотря на искажения и погрешности статистических данных всевозможных социологических опросов.

2. Проверка верности основных положений модели

Для проверки адекватности предположения о влиянии различных параметров на модели, проведем небольшой анализ статистических данных протоколов УИК.

Были взяты данные протоколов УИК 25 регионов европейской части РФ по выборам Президента РФ 2 марта 2008 года. На территории этих регионов численность электората составила 22 223 670, что составляет

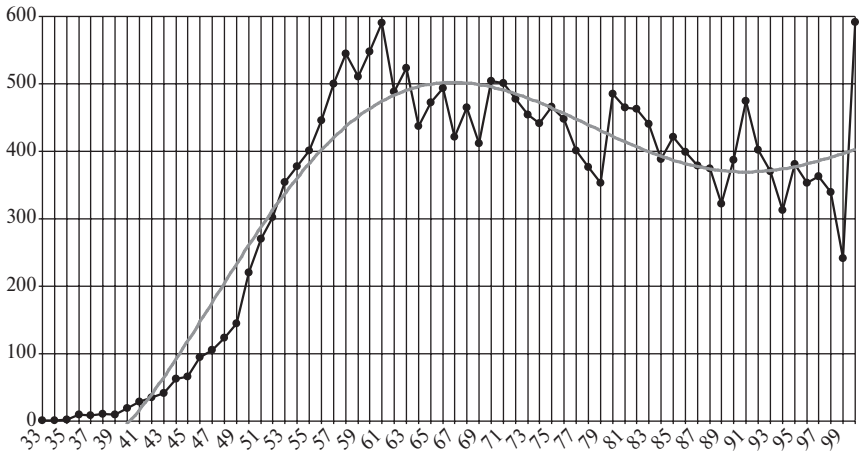


Рис. 1. Зависимость количества УИК от явки избирателей

около 21 % избирателей страны, количество УИК — 22 318 (примерно 23 % УИК).

Поскольку данными социологических опросов мы не располагали, проверим верность положения модели о неоднородности множества избирателей на основе имеющихся данных.

Так же как и оппоненты проведем исследование зависимости количества УИК от явки избирателей. На первый взгляд картина полностью соответствует картине, полученной оппонентами. Ни о каком нормальном распределении речи не идет, однако кривая регрессии (красный цвет) дает предположение о наличии неоднородного массива данных, видимо с разными тенденциями (рис. 1, см. [6, 7]).

Разделим все множество УИК на принадлежность к определенному месту проживания. Даже грубое разделение на «город—село» дает следующую картину (рис. 2).

Под «остальными» будем понимать УИК, которые формально не относятся ни к городу, ни к селу: т. е. судовые, военные, закрытые и иные УИК, их количество достаточно большое (890 штук или 3,99 %), однако количество избирателей на них около 1 % (в среднем около 290 избирателей по списку на один УИК), следовательно, вклад в общую копилку голосов также невелик.

Характерно, что даже такое грубое разделение всего множества УИК уже выделило 3 совершенно разные и в то же время устойчивые тенденции. Что, правда, совсем не удивительно. В селах, где взаимосвязи между людьми гораздо прочнее и глубже, чем в городе, гораздо выше средняя явка

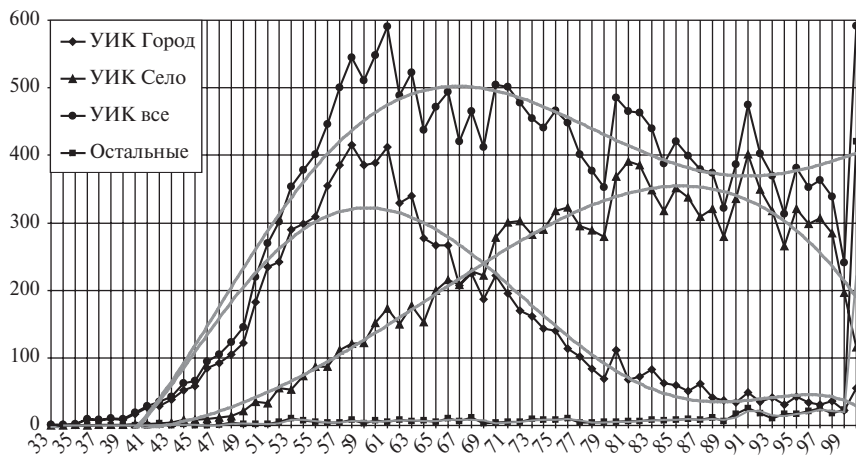


Рис. 2. Зависимость количества УИК, разделенных по принципу "город–село" от явки избирателей

и однороднее предпочтения. В городах, где эти связи слабые, явка ниже и предпочтения изменяются от региона к региону, что говорит о большей неоднородности самого массива городов, по отношению к селам.

Что касается судовых, закрытых и иных УИК, то тут, хотя вклад их по голосам незначителен, однородность достаточно большая (около половины имеют явку 100%). Данное явление тоже объяснимо, особенно для закрытых «военных» участков. Дисциплина, как известно, должна быть как на подводной лодке, железная, иначе что это за военная часть? Так что в 9:00 построение, в 9:15 голосование, направо, шагом марш. Но, подчеркнем, вклад их в общую копилку голосов достаточно мал, несмотря на то что данные взяты по Калининградской, Архангельской, Мурманской областям, а также по Ненецкому автономному округу.

Если данные закрытых участков исключить из общей картины голосов, качественная картина распределения голосов в зависимости от явки избирателей не изменится. И, кстати, не факт, что на этих закрытых участках было нарушение «прав человек» (см. например, отрицательный наклон графика в точке 100% голосов за Дмитрия Медведева).

Если дополнительно разложить УИК городского и сельского типа по количеству людей в списке (до 1000, от 1000 до 2000, более 2000), то получается следующая картина.

Из графиков видно, что в городах основной вклад, причем практически одинаковый, в общую тенденцию внесли УИК со списком более 1000 избирателей (более 95% избирателей), остальные 5% пришлось на мелкие

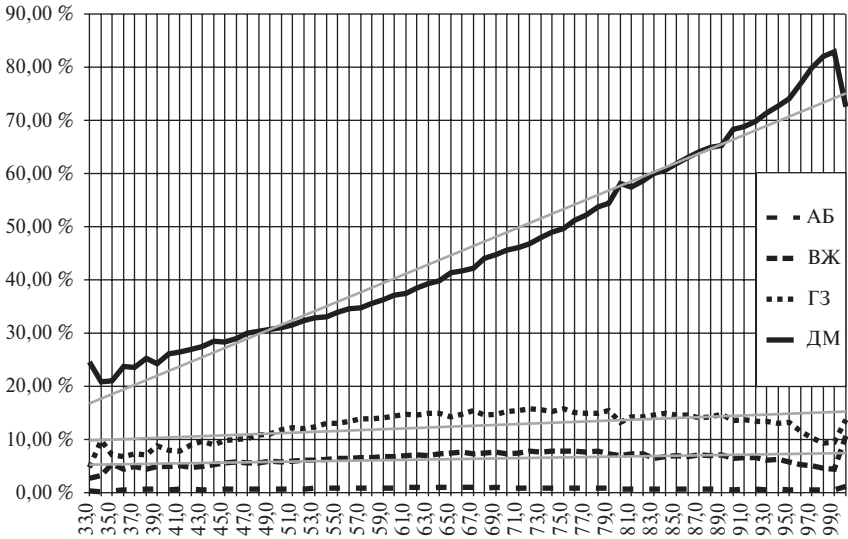


Рис. 3. Зависимость распределения голосов по кандидатам в процентах от списочного состава УИК в зависимости от явки

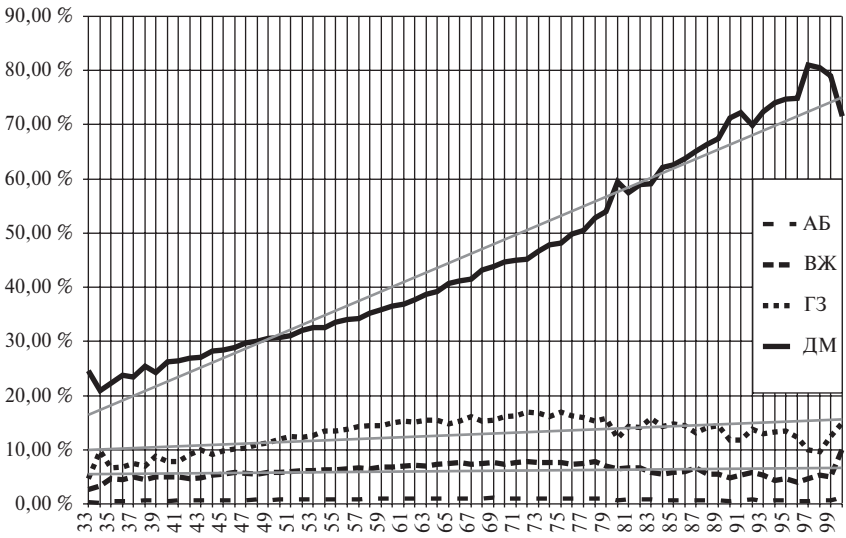


Рис. 4. Зависимость распределения голосов по кандидатам в процентах от списочного состава УИК в зависимости от явки (город)

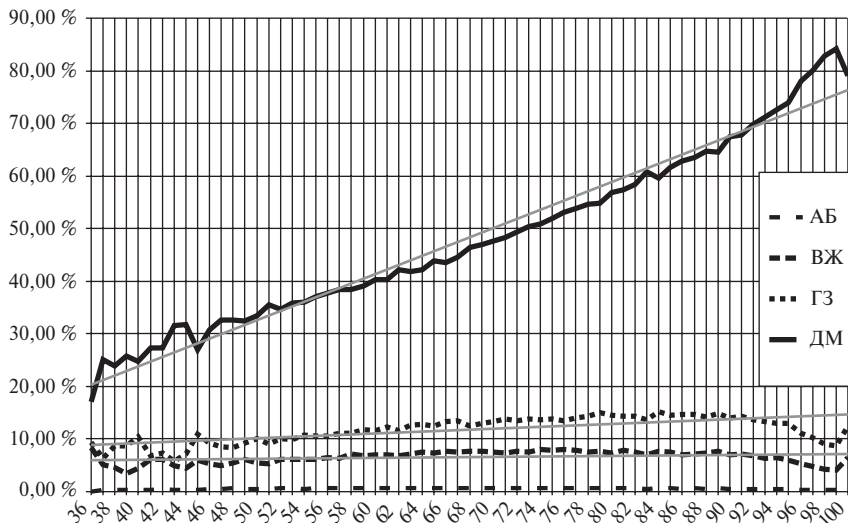


Рис. 5. Зависимость распределения голосов по кандидатам в процентах от списочного состава УИК в зависимости от явки (село)

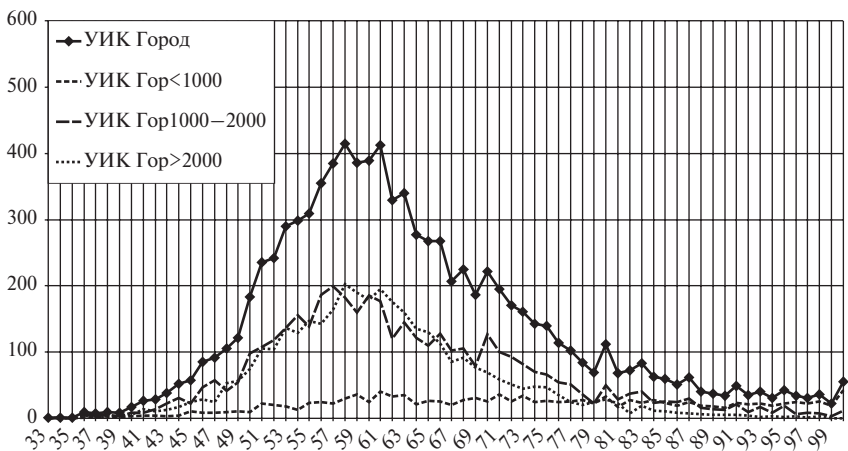


Рис. 6. Зависимость распределения количества УИК в зависимости от явки и размеров УИК (только город)

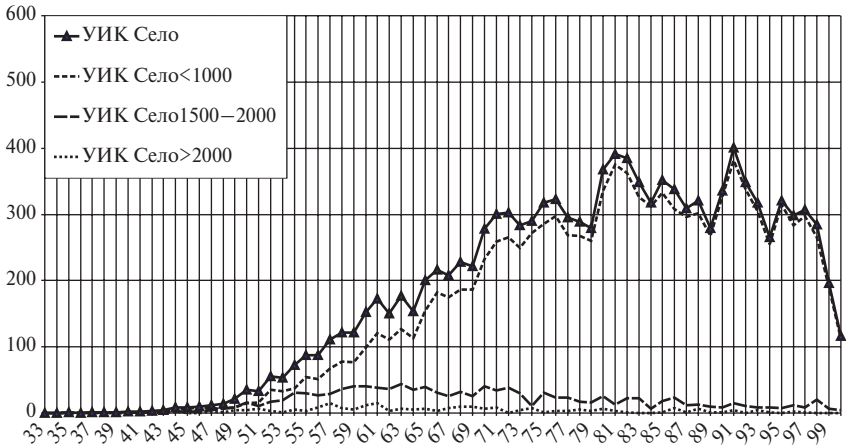


Рис. 7. Зависимость распределения количества УИК в зависимости от явки и размеров УИК (только село)

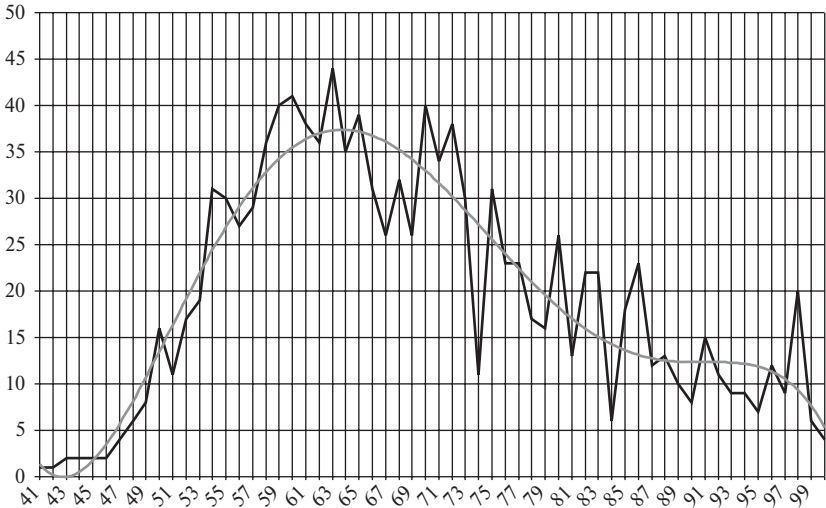


Рис. 8. Зависимость распределения количества УИК в зависимости от явки и размеров УИК (только село, УИК от 1000 до 2000 избирателей)

УИК. Тенденция мелких городских УИК отличается от общей тенденции крупных и претендует на проведение дополнительных исследований, например, по регионам, однако в виду малого вклада в общую копилку голосов, проводить их нет большого смысла.

В селах же наоборот, основной вклад у сельских УИК (более 65 % избирателей и самая высокая средняя явка — более 80 %). Они и определили общую тенденцию для всего множества сельских УИК. Чем крупнее сельские УИК, тем меньше их вклад (вернее множеств таких УИК) в общую копилку голосов и ниже явка.

Если детально рассмотреть тенденцию «средних» сельских УИК (от 1000 до 2000 избирателей, более 26 % избирателей), то получим следующую картину.

Крупные сельские УИК имеют похожую зависимость, однако, их вклад (около 8 % избирателей) на общую картину распределения голосов практически не влияет.

Если сложить все тенденции, определяемые каждым конкретным подмножеством УИК, то, во-первых, можно говорить о неоднородности всего множества УИК, и, во-вторых, что в пределах каждого подмножества закон распределения УИК оказывается близок к нормальному, за исключением «закрытых» УИК, которые по оговоренным выше причинам отнесем к подмножеству «особой электоральной культуры». Но опять же повторимся, что вклад этого подмножества в общую картину невелик, а утверждение о нарушениях законодательства спорно.

Кстати, аналогичные приближения можно получать не только аппроксимацией функциями, но, например, по методу наименьших квадратов, или путем кластерного анализа исходной статистики. Качественная картина при этом не изменяется, а приближение получается достаточно точное (см. графики ниже). Средняя явка (математическое ожидание) для городов равна 62,28 %.

Данные графиков подтверждают еще раз, что в пределах условно однородного подмножества УИК, выделенного по параметру «место проживания», фактическое распределение УИК оказывается близким к нормальному.

Если каждому УИК поставить в соответствие вес, равный количеству проголосовавших на нем избирателей, то картина будет сходной с распределением количества УИК в зависимости от явки, однако менее выражен будет подъем на конечном участке графика из-за того, что УИК со 100 % явкой очень маленькие («закрытые», судовые и иные в среднем имеют по 140 избирателей при 100 % явке), и, следовательно, их вклад в общую копилку голосов невелик.

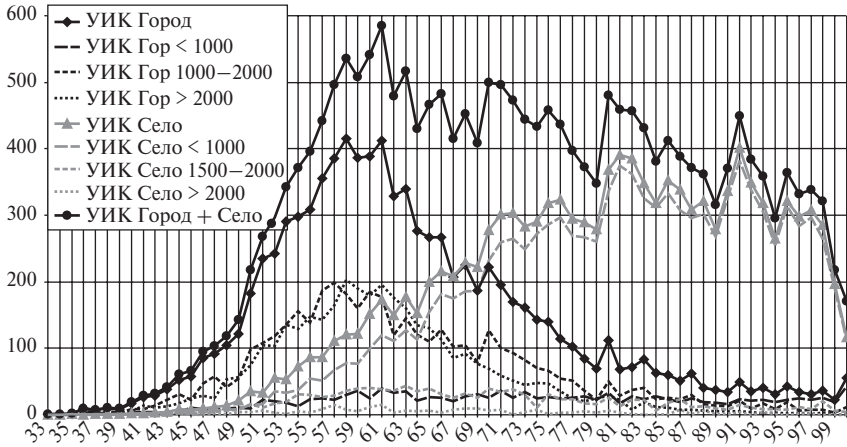


Рис. 9. Зависимость распределения количества УИК в зависимости от явки по различным подмножествам, определяемым местом проживания (за исключением подмножества «особой электоральной культуры»)

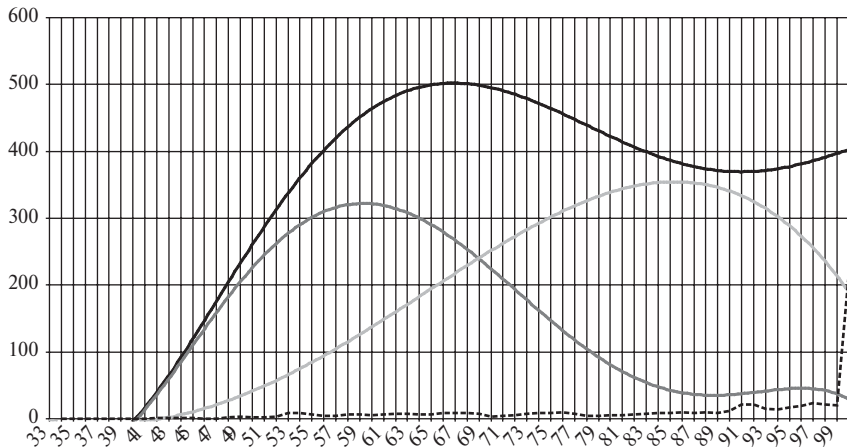


Рис. 10. Тенденции зависимости распределения количества УИК от явки по различным подмножествам (красный — суммарная картина, зеленый — город, голубой — село, малиновый — остальные), определяемым местом проживания (данные аппроксимации)

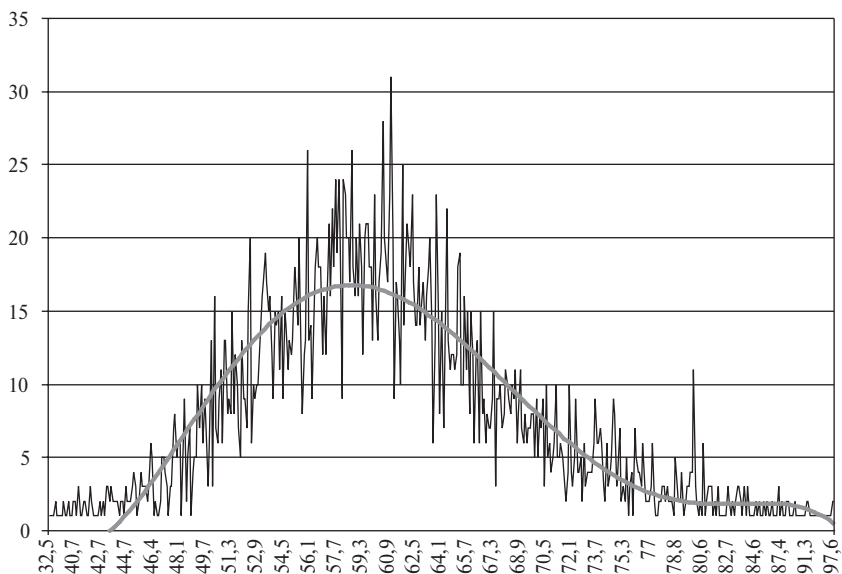


Рис. 11. Зависимость распределения количества УИК в зависимости от явки (только город) без округления до целых процентов

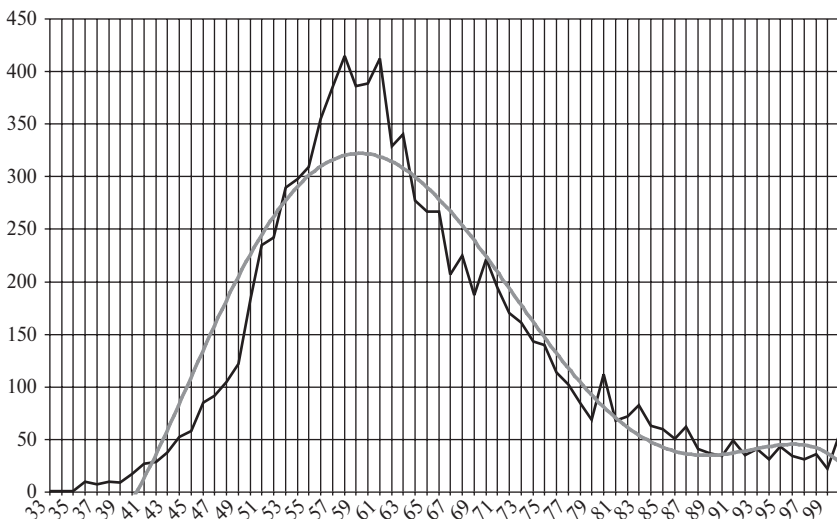


Рис. 12. Зависимость распределения количества УИК в зависимости от явки (только город) с точностью до единиц процентов

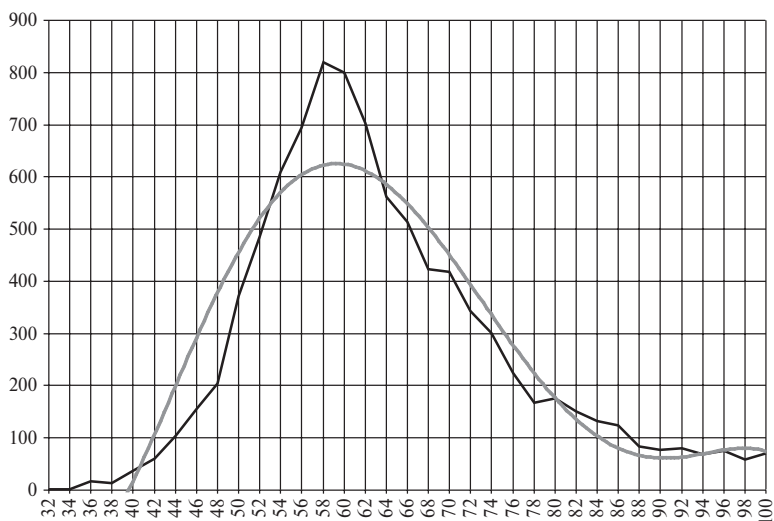


Рис. 13. Зависимость распределения количества УИК в зависимости от явки (только город) с точностью до двух процентов

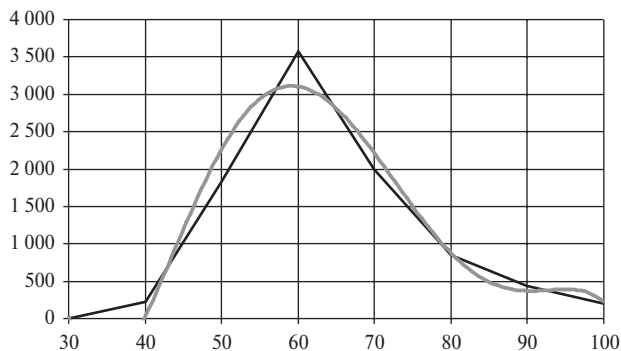


Рис. 14. Зависимость распределения количества УИК в зависимости от явки (только город) с точностью до 10 процентов

На следующем рисунке представлено сравнение зависимостей количества УИК и взвешенного количества УИК от явки избирателей.

Если проанализировать зависимость взвешенного количества УИК от места проживания избирателей (город—село—остальные), получим следующие графики.

По данному графику видно, что вклад «остальных» УИК совсем незначителен и дает небольшую прибавку только в точке 100 % явки, кото-

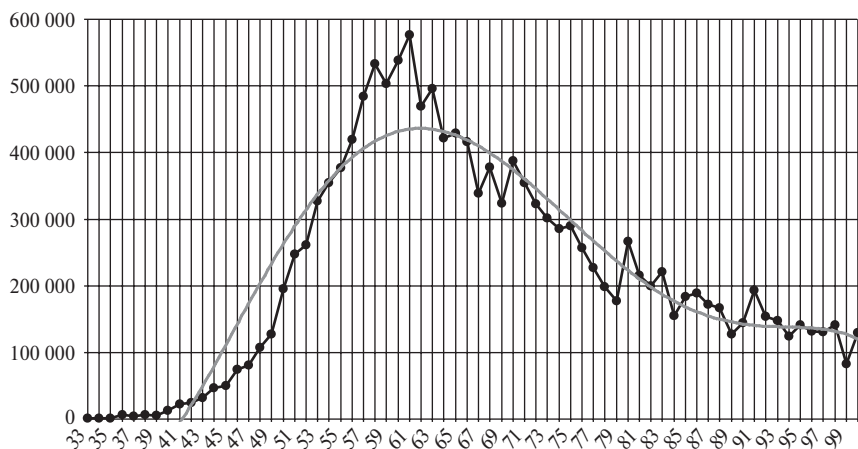


Рис. 15. Зависимость взвешенного количества УИК (вес равен количеству проголосовавших) от явки избирателей

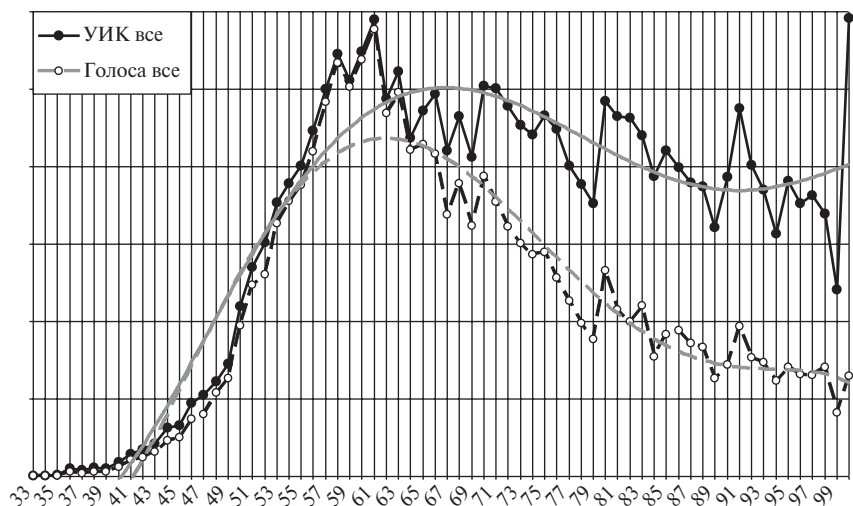


Рис. 16. Сравнение зависимостей количества УИК и взвешенного количества УИК (пунктирная линия) от явки избирателей

рая, однако, не влияет, на общую картину, так как эта прибавка составляет менее 0,4% общего количества поданных голосов. При этом заметно, что «городская» тенденция практически «нормальная» (слегка положительно — слева — скошена) с небольшими вкладами на участках 67–75%

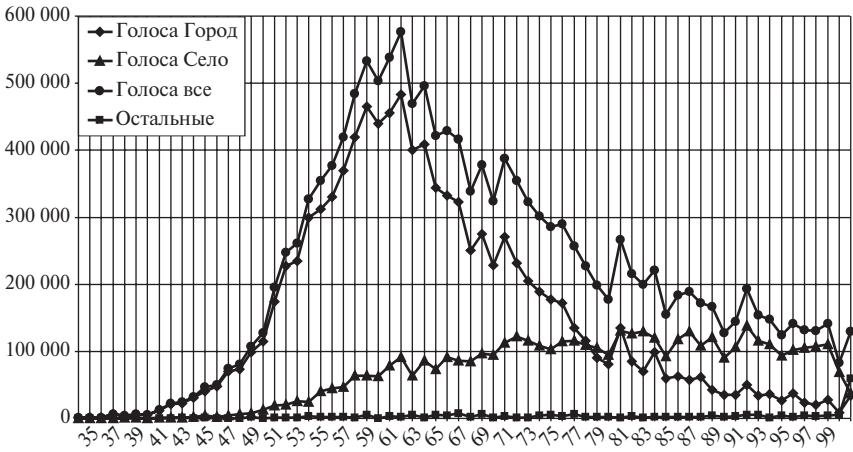


Рис. 17. Зависимость взвешенного количества УИК, разделенных по принципу "город—село" от явки избирателей

и 79–85 %, требующих более детального изучения. «Сельская» же становится более «размытой», что говорит о значительной неоднородности, скорее всего связанной с региональными различиями.

В целом сохраняются и пропорции поданных голосов за кандидатов.

На конечном участке графика (см. следующий рисунок) пропорция несколько изменена, что соответствует доминированию «сельской» тен-

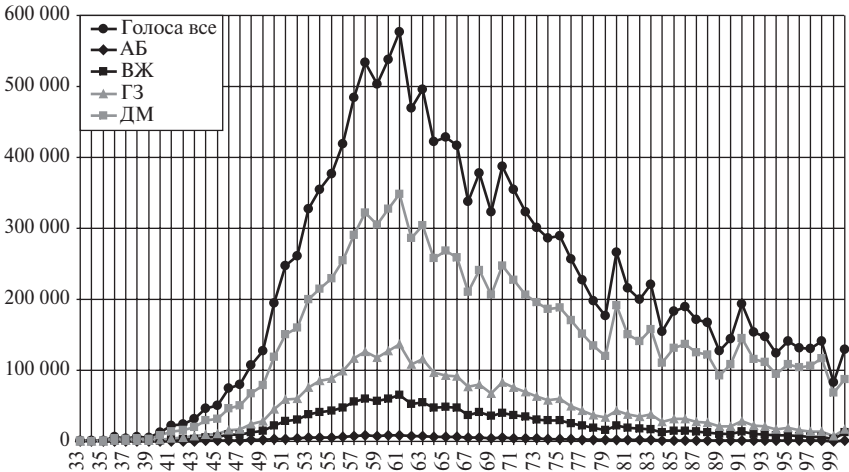


Рис. 18. Зависимость голосов за кандидатов от явки избирателей

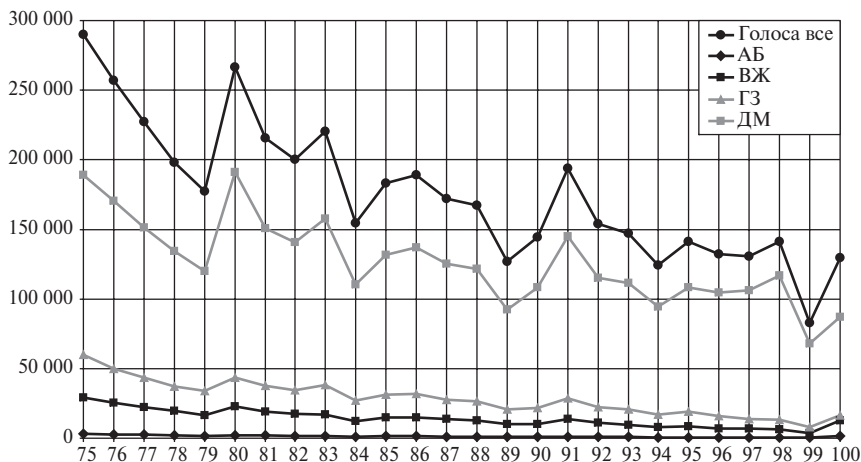


Рис. 19. Зависимость голосов за кандидатов от явки избирателей (увеличенный конечный участок графика)

денции (см. также графики 3–5), где пропорции несколько иные, чем в городах. Вообще же города дали 67,3 % от общего процента поданных голосов, села — 31,3 %, «остальные» — 1,4 %.

Выводы

На небольшом примере мы убедились в верности гипотезы о неоднородности множества избирателей.

На однородных подмножествах избирателей четко прослеживается независимость распределений по независимым параметрам (например, зависимости количества УИК от явки избирателей).

Возможно проведение дальнейших исследований по детализации подмножеств избирателей на основании подхода, описанного в разделе 1, с целью построения полной модели электорального поведения избирателей в РФ.

Необходимо сравнение расчетных данных по модели с реальными данными электоральной статистики.

Литература

1. Бузин А. Курс аномальной электоральной вероятности. [Электронный ресурс].
2. Шпилькин С. Обработка данных результатов федеральных выборов 2007–2008 гг. [Электронный ресурс].

3. *Любарев А. Е.* Статистические аномалии на федеральных выборах 2007–2008 гг. [Электронный ресурс].
4. Интернет-сайт Центральной избирательной комиссии.
5. *Орлов А. И.* Прикладная статистика. М.: Экзамен, 2006. 671 с.
6. *Романов В. Н.* Теория измерений. Методы обработки результатов измерений. СПб.: СЗТУ, 2006. 127 с.
7. *Елисеева И. И., Юзбашев М. М.* Общая теория статистики. М.: Финансы и статистика, 1996. 368 с.
8. ВЦИОМ: Телевидение — главное оружие избирательных кампаний / [Электронный ресурс] <http://www.iamik.ru/?op=full&what=content&ident=33605>