

# Статистика исследовала выборы

Статистический анализ результатов думских выборов-2011 показывает «искусственность» результатов «Единой России». Коллега главы ЦИК физик Сергей Шпилькин рассказывает, как на основании общедоступных данных Центризбиркома получить истинные результаты выборов.

Каждые выборы – это огромный эксперимент по определению мнений и предпочтений избирателей. Федеральные выборы дают 96 с небольшим тысяч «результатов измерений» – протоколов избирательных комиссий, по несколько десятков чисел в каждом. Даже если мы подозреваем, что эти данные как-то искажены, анализируя эти «измерения», можно многое узнать и об объекте измерения – предпочтениях избирателей, и об «измерительном приборе», то есть самой выборной системе.

Прежде чем приступать к статистическому анализу такого объема данных, полезно привести его к обозримому виду, и проще всего это сделать в графической форме. Мы рассмотрим два графических подхода, которые удобно применять для анализа выборных данных.

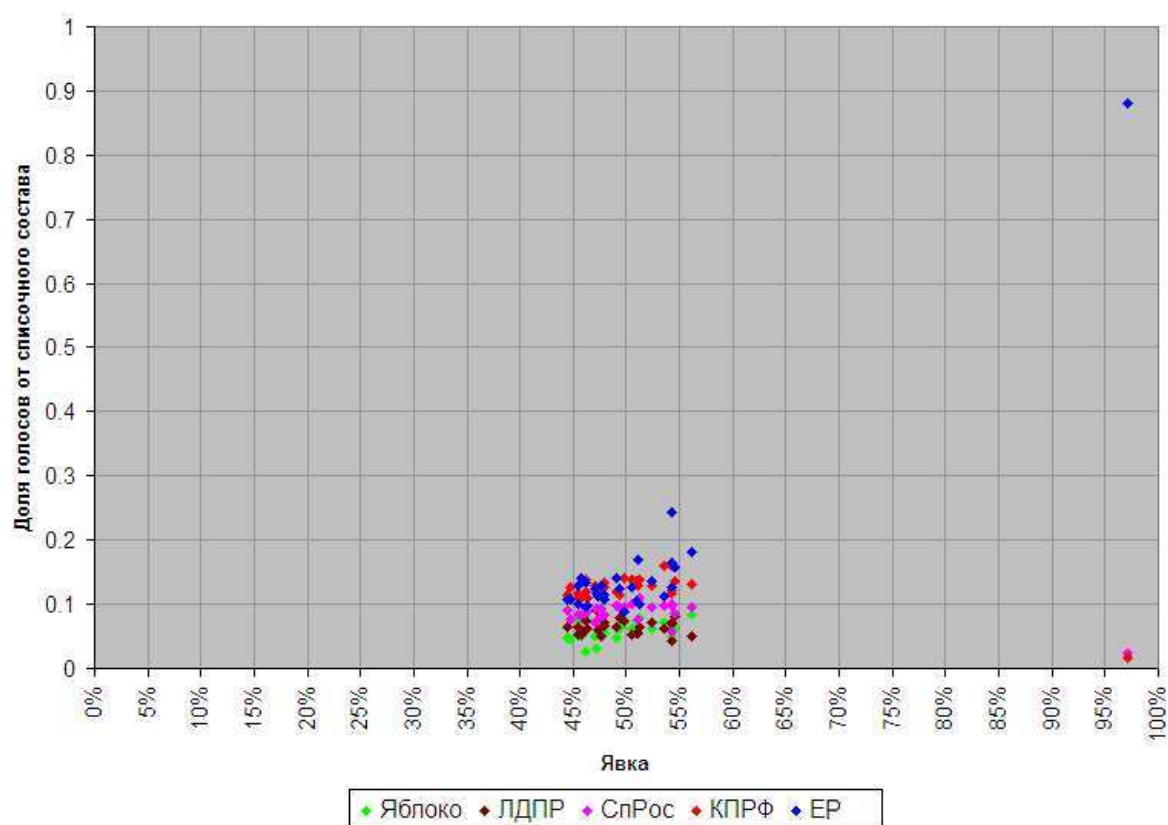
## Диаграммы рассеяния

Диаграмма рассеяния – это график, позволяющий выявить взаимосвязь между двумя показателями. Результаты измерений наносятся в виде точек на координатную плоскость с координатами, соответствующими интересующим нас показателям.

Преимущество диаграммы рассеяния состоит в том, что она позволяет исследовать даже совсем небольшие наборы данных.

В качестве примера ниже представлены в виде диаграмм рассеяния результаты думских выборов 2011 года в трех районах Москвы – Южном Тушино, Строгино и Гольяново. Каждая точка соответствует результатам одной из партий на одном избирательном участке. По горизонтальной оси отложена явка на участке, по вертикальной – доля зарегистрированных избирателей, проголосовавших за соответствующую партию на этом участке. Каждому избирательному участку соответствует пять точек – по одной для результата каждой из показанных партий (данные для других партий не приведены).

## Москва, Южное Тушино

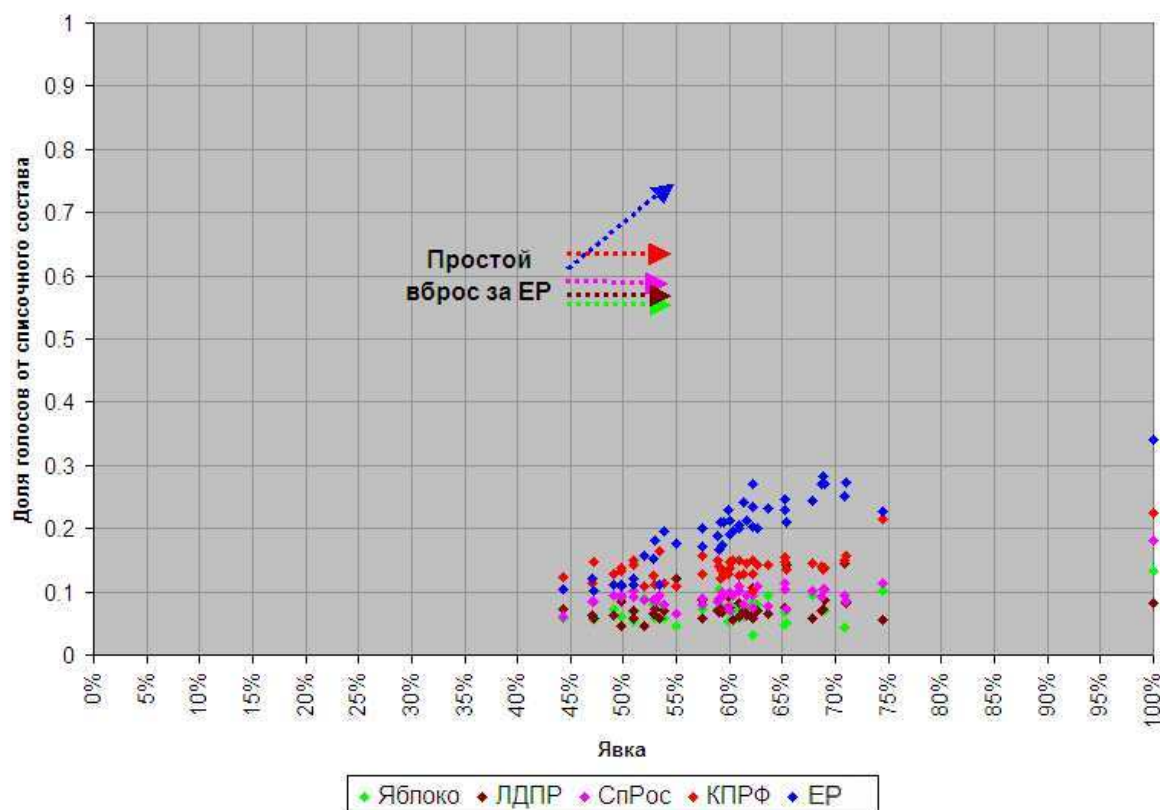


Все три территории – спальные районы Москвы с обычным составом населения, однако наблюдаемые картины различны. Картина на диаграмме Южного Тушино (рисунок выше) выглядит совершенно естественно.

Диапазон явки невелик, что ожидаемо для небольшой территории, соотношение голосов, поданных за разные партии, на всех участках более или менее одинаково.

Единственная выбивающаяся из общей картины точка с явкой 96% – небольшой (меньше трехсот человек) «закрытый» участок.

## Москва, Строгино

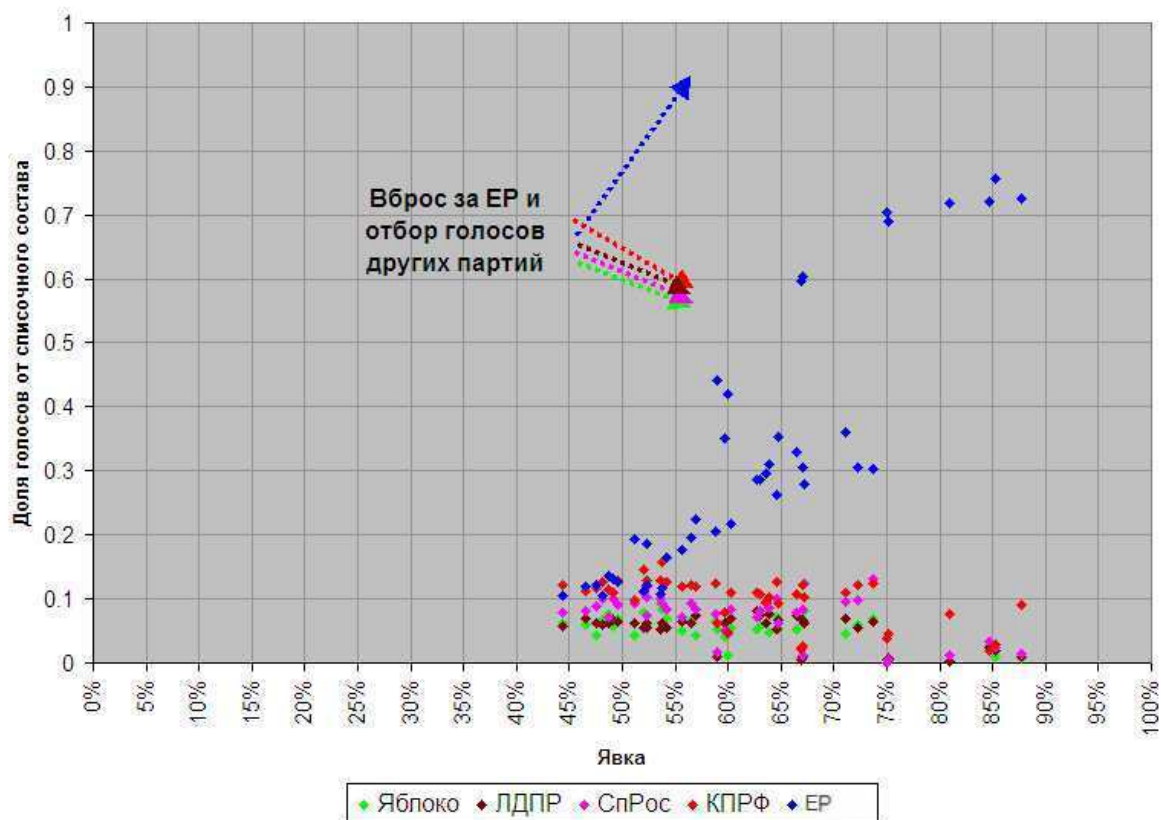


На диаграмме для Строгино ситуация другая. Диапазон явки больше, но на участках с большей величиной явки почему-то увеличивается только доля избирателей, проголосовавших за ЕР, а доли остальных партий остаются постоянными.

Такое поведение легко объяснить, если понять, как меняется положение избирательного участка на диаграмме при искусственном добавлении голосов за одну из партий (показано стрелками).

Во-первых, точки всех партий для этого участка сдвигаются в сторону большей явки на столько процентов, сколько составляют добавленные голоса от численности избирателей на участке. Во-вторых, точка партии, за которую добавляют голоса, поднимается вверх на столько же процентов. В итоге точка партии, за которую происходит «вброс» (для простоты мы будем называть добавление голосов вбросом, хотя оно вполне может быть просто результатом изменения протокола), сдвигается вправо по диагонали, а точки других партий – по горизонтали. Именно это мы и видим: точки голосования за ЕР распределены по диагонали, а точки голосования за остальные партии – по горизонтали.

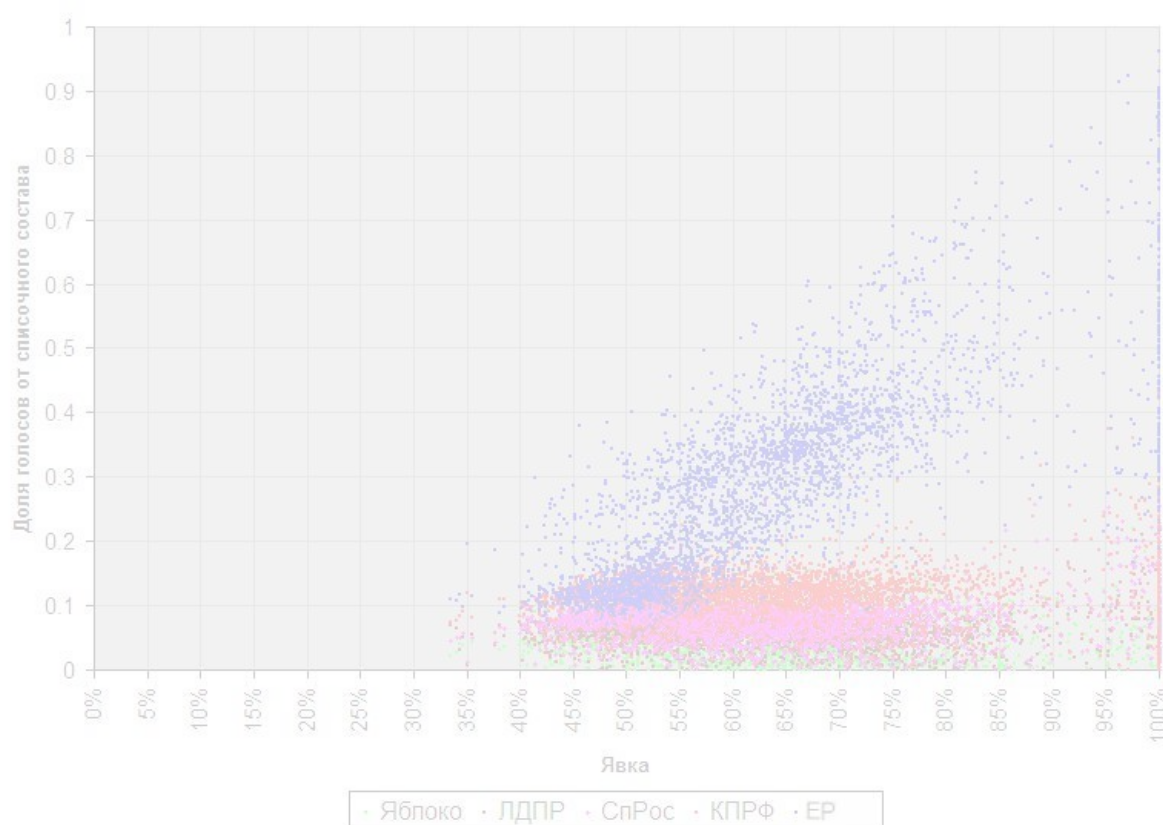
## Москва, Гольяново



На диаграмме района Гольяново проявляется еще одно явление. Кроме групп точек, укладывающихся в картину добавления голосов за ЕР, как на предыдущей диаграмме, появляются группы, в которых доли голосов за остальные партии резко понижены по сравнению с соседними участками, а доля голосов за ЕР увеличена еще больше.

Это соответствует ситуации, когда ЕР добавляют голоса, а у других партий отнимают.

## Вся Москва



На диаграмме для всех московских участков внизу слева выделяется плотная группа синих точек (голосование за ЕР), расположенных примерно так, как в Южном Тушино. Естественное предположить, что это участки с честным подсчетом. Остальные точки голосования за ЕР размазаны в диагональное облако, соответствующее искусственному завышению голосов за ЕР и, возможно, занижению голосов за другие партии. Если внимательно присмотреться, в этом облаке можно заметить плотную наклонную полосу немного выше линии, на которой доля голосов за ЕР равна половине явки.

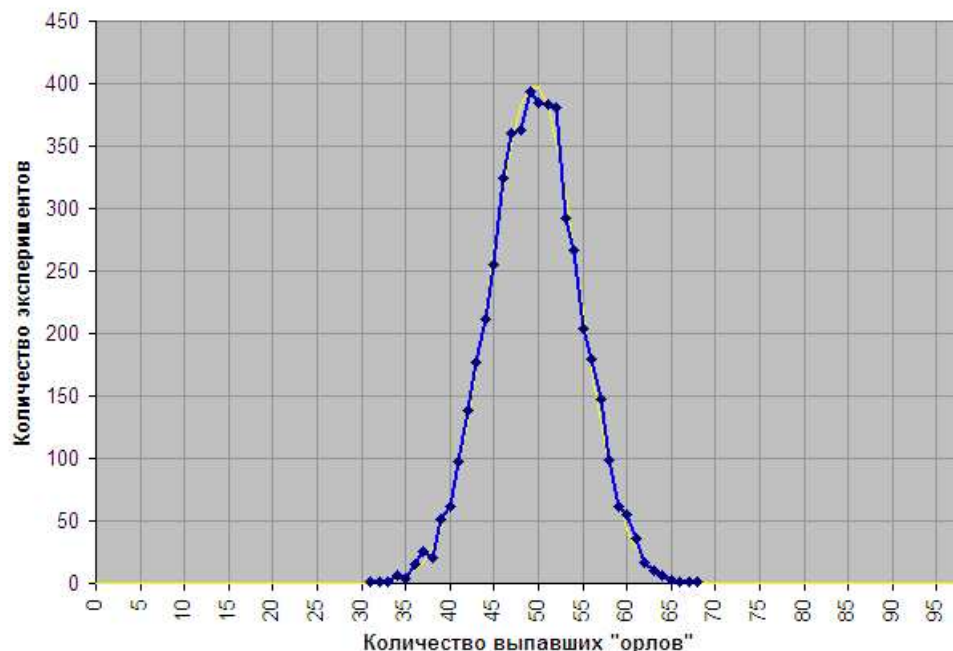
В пересчете на число проголосовавших избирателей это означает, что на этих участках ЕР получила немного больше 50% голосов.

Судя по тому, что такие участки выделяются в виде отдельной группы, показатель в 50% был для части участков плановым заданием.

### Гистограммы

Гистограммы распределения голосов – еще более удобный и показательный способ визуализации результатов выборов. Поясним, что они собой представляют и как работают на классическом примере игры в орлянку – «орел» или «решка».

Вероятность выпадения «орла» при бросании монетки составляет 50%. Но это не означает, что, если бросить монетку 100 раз, «орел» выпадет ровно 50 раз. Наоборот, если проделать множество (например, 10000) экспериментов, бросая монетку по сто раз в каждом (естественно, бросание монетки можно смоделировать на компьютере), и построить гистограмму распределения исходов по количеству выпавших орлов, мы получим что-то подобное изображенному на графике.



Количество выпадений конкретного числа орлов является случайной величиной, поэтому полученное распределение не будет гладким (и будет каждый раз немного разным), но в целом его форма с хорошей точностью описывается известным из теории вероятностей т. н. нормальным, или Гауссовым, распределением (тонкая желтая линия). График Гауссова распределения имеет вид симметричной колоколообразной кривой.

Гауссово распределение типично для многих величин, которые зависят от множества случайных независимых факторов (как в нашем случае количество выпавших орлов зависит от результатов бросания монетки в ходе эксперимента).

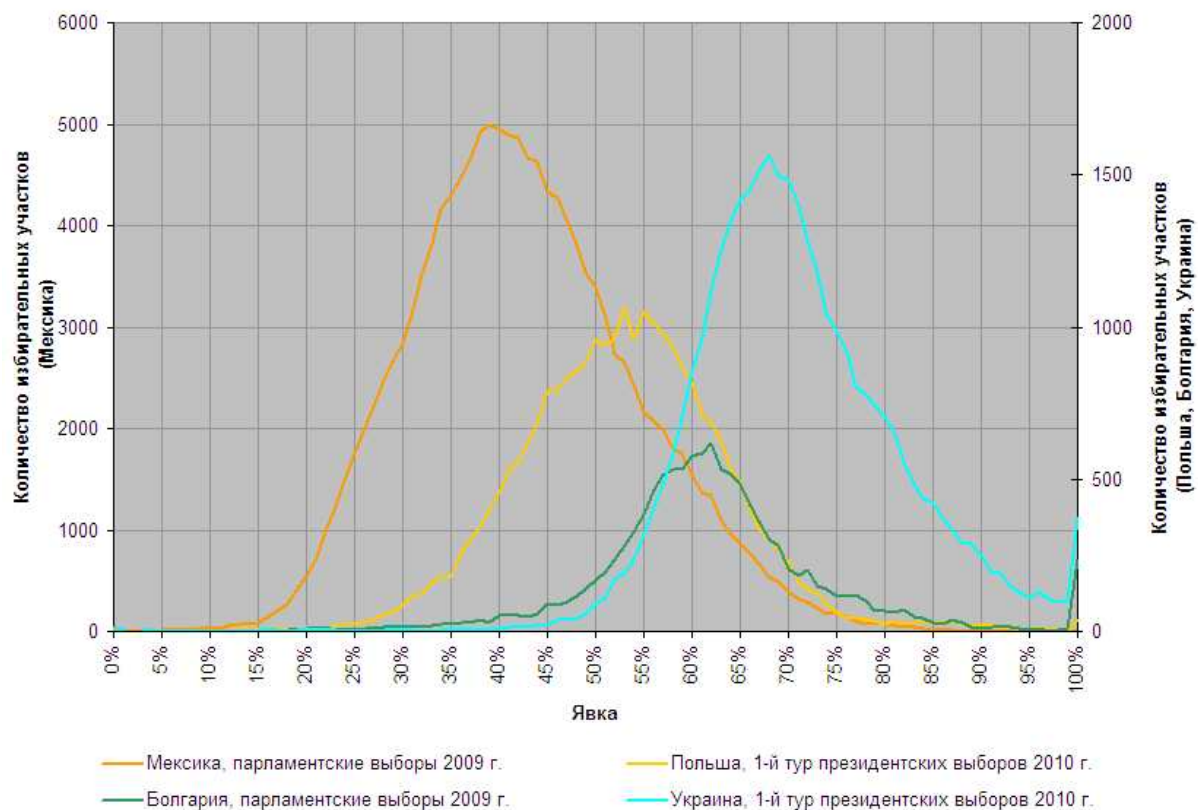
К числу таких величин относятся в том числе и явка избирателей, и доля голосов, поданных за конкретную партию, если каждый избиратель принимает решение об участии в выборах и голосовании самостоятельно и независимо.

Из-за разных размеров избирательных участков, разной активности и разных предпочтений избирателей на разных участках и т. п. на практике распределение этих показателей является суммой множества Гауссовых распределений с разной шириной и разным положением центральной линии, но разумно ожидать, что его общий характер (симметричная колоколообразная форма) будет по-прежнему сохраняться.

Когда речь идет об анализе десятков тысяч измерений, диаграммы рассеяния становятся неудобны. В этой ситуации удобнее применять другой стандартный статистический инструмент – гистограммы распределения данных по различным показателям. При построении гистограммы диапазон значений интересующего нас показателя разбивается на интервалы и для каждого интервала подсчитывается количество измерений, в которых показатель попадает в данный интервал. Например, медицинские статистики строят гистограммы заболеваемости простудными заболеваниями по неделям, демографы – гистограммы распределения численности населения по возрасту и т. п.

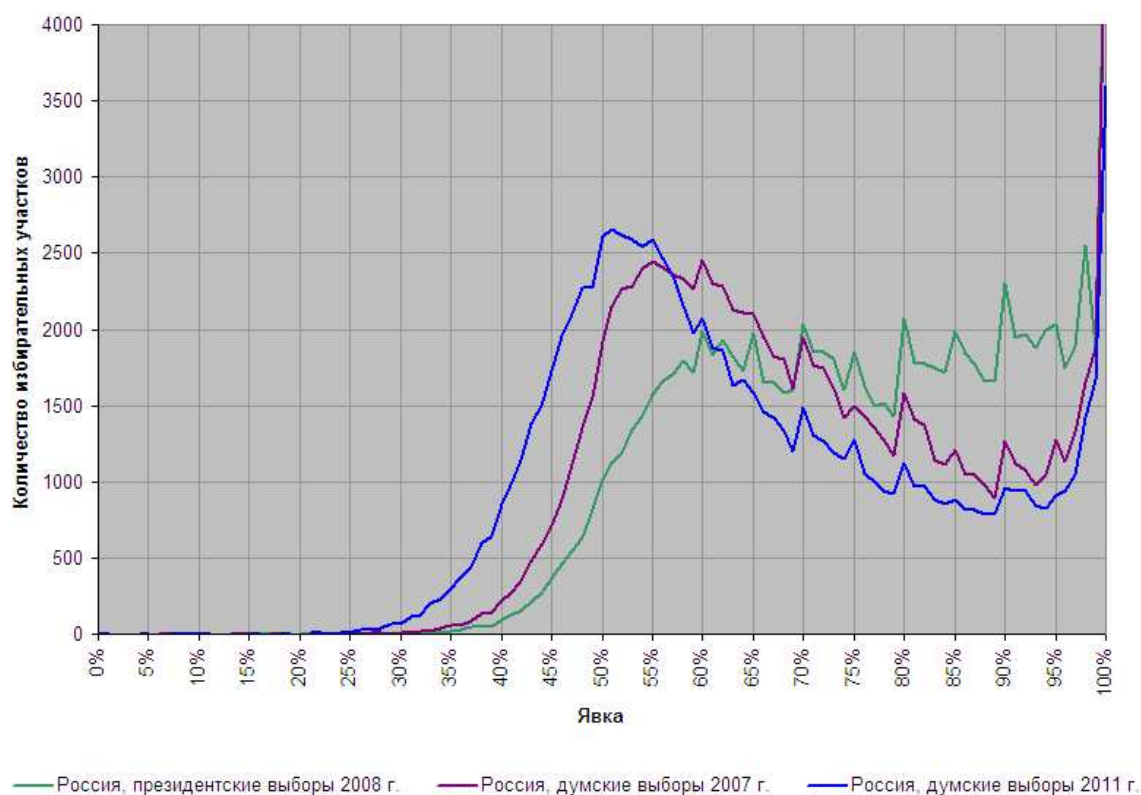
При анализе выборов удобно строить гистограммы распределения избирательных участков или голосов избирателей по явке и по процентам голосования за различные партии. Для начала, в качестве примера, гистограммы распределения избирательных участков по явке в нескольких зарубежных странах. По горизонтальной оси отложена явка, по вертикальной – количество избирательных участков в интервале явки шириной 1%.





Для всех показанных стран распределения достаточно гладкие и более или менее симметричные, похожие на Гауссовы.

Совсем не так обстоит дело в случае российских выборов.



Бросаются в глаза две примечательные особенности.

1. Распределения участков по явке не имеют колоколообразной формы, как на приведенных выше графиках для зарубежных стран. В первую очередь это касается области высоких явок, где распределение не снижается симметрично начальному участку, а остается на высоком уровне вплоть до явок в 100%. Другими словами, на наших выборах завышено количество участков с высокой явкой.

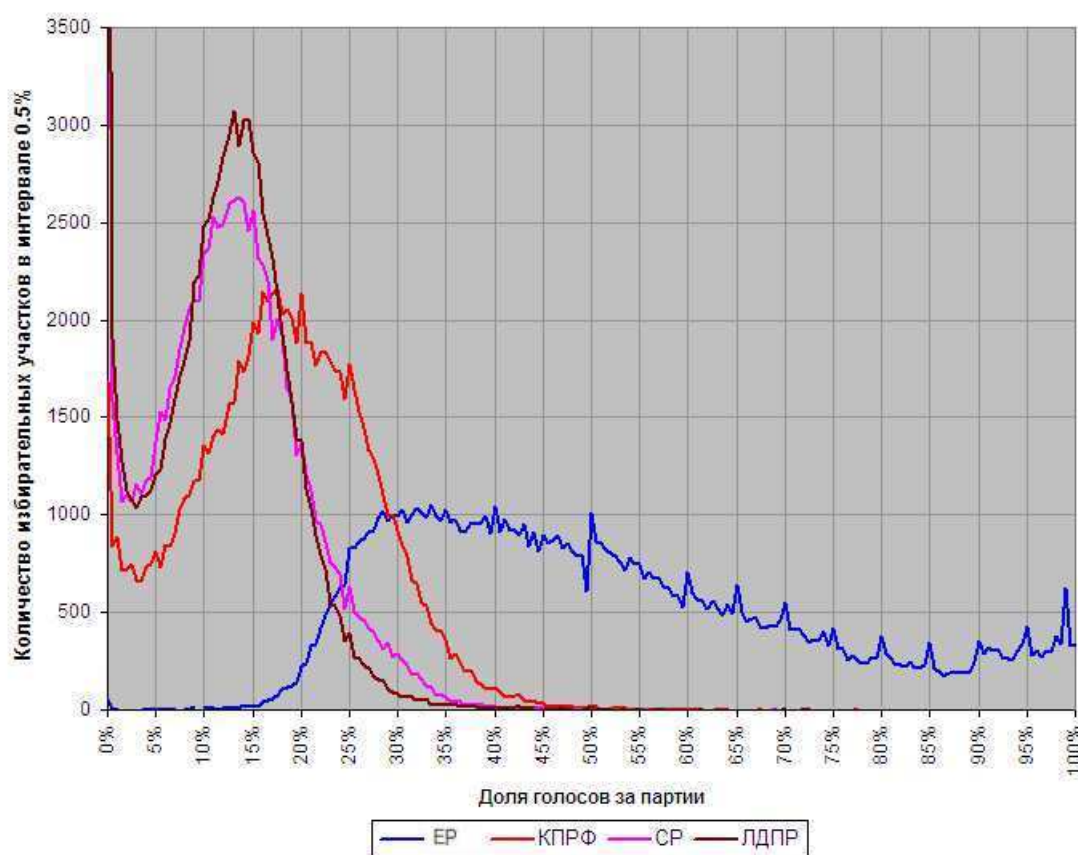
2. Распределения имеют резкие пики на значениях явки, кратных 5%. Например, на президентских выборах 2008 года участков с явкой 79% было 1429, с явкой 80% – 2069, 81% – 1787.

Как мы уже видели при рассмотрении диаграмм рассеяния, большое количество участков с повышенной явкой образуется при искусственном добавлении голосов в пользу какой-либо партии, хотя можно пытаться объяснить его, например, высокой явкой в республиках Кавказа или в сельской местности.

Но что с трудом поддается объяснению естественными механизмами – это склонность результатов выборов к круглым значениям явки.

Единственное разумное объяснение состоит в том, что такая явка получена в результате ручного вмешательства, то есть фальсификацией результатов выборов в соответствии с заданными сверху показателями.

Кстати, меньше всего пики на красивых явках выражены на распределении для думских выборов 2011 года. Однако это не означает, что с ними все в порядке. На следующем графике показано распределение избирательных участков на выборах 2011 года по проценту голосов, полученных различными партиями. По горизонтальной оси отложена доля голосов, полученных партиями, по вертикальной – количество избирательных участков с такой долей голосов за партии в интервале 0.5%.



Здесь все еще интереснее.

График для «Единой России» немного похож на нормальное распределение разве что в своей начальной левой части, до вершины.

Правая часть распределения по голосам за ЕР неестественно растянута вплоть до 100%, а пики на кратных 5% значениях доли голосов однозначно показывают, что на процент голосования за ЕР, как и на явку, был заказ. Надо отметить, что если довести явку до нужного уровня еще теоретически можно, хватая избирателей за рукав на улице и приводя их на участок, то сделать красивый процент голосования можно только после вскрытия избирательных урн и подсчета бюллетеней, то есть только путем прямой подделки результатов.

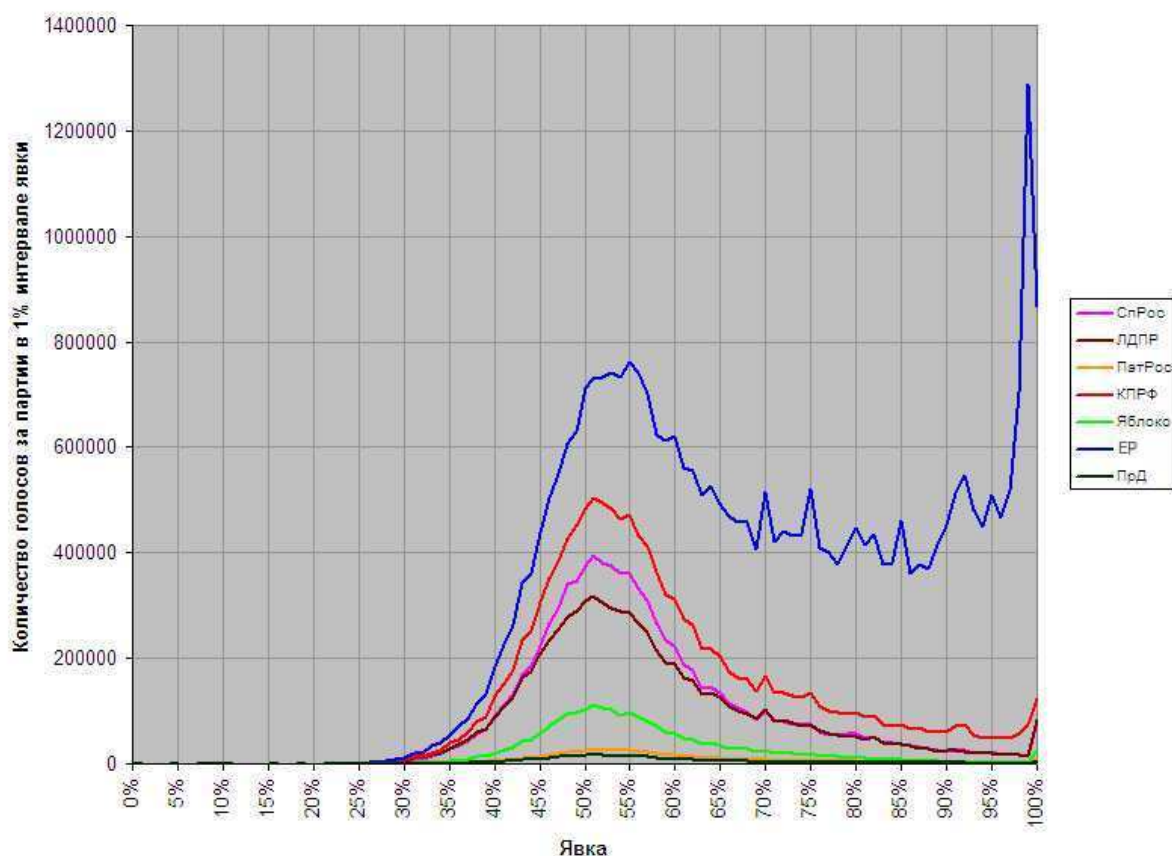
С другой стороны, графики для КПРФ, «Справедливой России и ЛДПР вполне были бы похожи на Гауссовы распределения, если бы не завышенное количество участков с низкими долями голосов за эти партии (левая



сторона графиков) и не «объединенная» вершина у графика распределения по доле голосов за КПРФ. Это дает статистическое подтверждение многочисленным сообщениям очевидцев о переписывании бюллетеней, с завышением голосов за ЕР и занижением за другие партии.

Итак, мы видим, что единственная партия, распределение голосов за которую радикально отличается от Гауссова – это «Единая Россия».

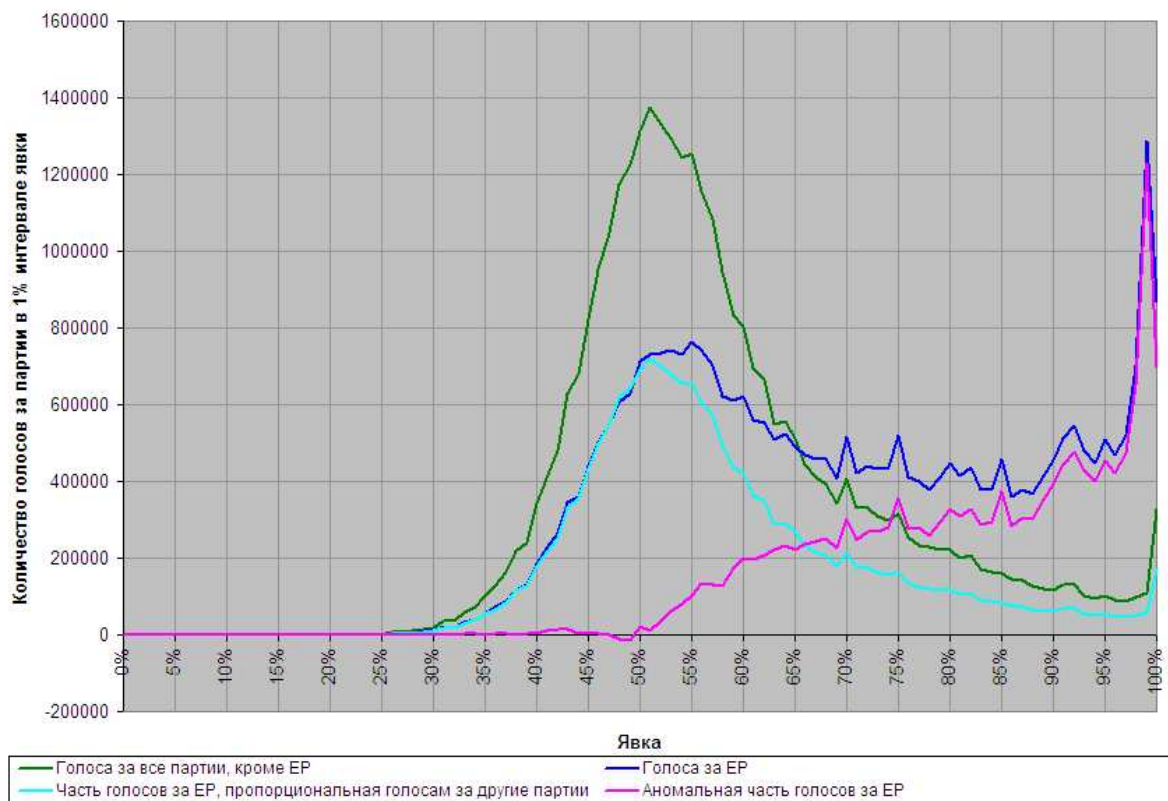
Более того, есть явные свидетельства, что показатели голосования за эту партию подделывались. Это еще один аргумент в пользу того, что странность российских распределений по явке – результат именно добавления голосов в пользу «Единой России». Посмотрим, как это отражается на результатах голосования количественно. На следующем графике показано распределение голосов, полученных разными партиями на участках с различной явкой. По горизонтальной оси отложена явка, по вертикальной – количество голосов на участках в интервале явки 1%.



Распределения голосов, полученных всеми партиями, очень похожи между собой, и только распределение голосов за «Единую Россию» отличается от них: чем выше явка, тем большую относительную долю голосов получает ЕР.

Именно такую картину мы видели на диаграммах рассеяния: при добавлении голосов за ЕР голоса за остальные партии просто размазываются по оси явки, сохраняя неизменное соотношение между собой, а относительная доля ЕР растет.

Предполагая, что различие в распределении голосов за ЕР и за другие партии в зависимости от явки обусловлено именно искусственным завышением голосов за ЕР, можно попробовать оценить вклад этого завышения. Для этого попытаемся выделить из распределения голосов за ЕР часть, пропорциональную сумме голосов за другие партии.



Как видно из графика, из распределения голосов за «Единую Россию» действительно удалось выделить часть, пропорциональную голосам за другие партии, таким образом, что до явки в 50 – 52% остаток голосов за ЕР после вычитания этой части практически равен нулю. В рамках наших предположений это означает, что в этом диапазоне явок добавленных за ЕР голосов практически нет. Это согласуется с расположением нижнего облака на диаграмме рассеяния на рис. 4 и в некоторой степени оправдывает нашу методику.

Оставшуюся («аномальную») часть голосов за ЕР в рамках этого подхода следует рассматривать как результат искусственного завышения голосов за эту партию. После того как разделение на «нормальную» (пропорциональную голосам за другие партии) и «аномальную» части выполнено, можно оценить их количественно и попытаться восстановить «скорректированные» результаты голосования в отсутствие такого завышения.

На момент написания статьи данные голосования на выборах в Госдуму 4 декабря 2011 года еще не были окончательными и охватывали 108,6 млн избирателей (примерно 98% полной численности).

При вычислении на этих данных из 32,1 млн голосов, поданных за ЕР, «нормальными» оказываются примерно 16,8 млн., а «аномальными», т.е. искусственно добавленными – примерно 15,2 млн.

Официальные и скорректированные итоги голосования при этом выглядят следующим образом:

|                     | Официально     |           | Скорректир.    |           |
|---------------------|----------------|-----------|----------------|-----------|
|                     | Кол-во голосов | % голосов | Кол-во голосов | % голосов |
| Справедливая Россия | 8 653 720      | 13.49%    | 8 653 720      | 17.69%    |
| ЛДПР                | 7 628 916      | 11.89%    | 7 628 916      | 15.60%    |
| Патриоты России     | 635 566        | 0.99%     | 635 566        | 1.30%     |
| КПРФ                | 12 535 119     | 19.54%    | 12 535 119     | 25.63%    |
| Яблоко              | 2 228 672      | 3.47%     | 2 228 672      | 4.56%     |
| Единая Россия       | 32 071 784     | 50.00%    | 16 837 903     | 34.43%    |
| Правое Дело         | 390 204        | 0.61%     | 390 204        | 0.80%     |
| Всего               | 64 143 981     |           | 48 910 100     |           |
| Явка                | 59.65%         |           | 45.40%         |           |

Автор: Сергей Шпилькин    © Газета.Ru    ПОЛИТИКА, РОССИЯ    8580    15.12.2011, 20:20    679

URL: <https://babr24.com/?ADE=100735>    Bytes: 14216 / 13476    Версия для печати    Скачать PDF

👍 Пореккомендовать текст

Поделиться в соцсетях:

Также читайте эксклюзивную информацию в соцсетях:

- [Телеграм](#)

- [ВКонтакте](#)

Связаться с редакцией Бабра:

[newsbabr@gmail.com](mailto:newsbabr@gmail.com)

Автор текста: **Сергей  
Шпилькин.**

#### НАПИСАТЬ ГЛАВРЕДУ:

Телеграм: [@babr24\\_link\\_bot](#)

Эл.почта: [newsbabr@gmail.com](mailto:newsbabr@gmail.com)

#### ЗАКАЗАТЬ РАССЛЕДОВАНИЕ:

эл.почта: [bratska.net.net@gmail.com](mailto:bratska.net.net@gmail.com)

#### КОНТАКТЫ

Бурятия и Монголия: Станислав Цырь

Телеграм: [@bur24\\_link\\_bot](#)

эл.почта: [bur.babr@gmail.com](mailto:bur.babr@gmail.com)

Иркутск: Анастасия Суворова

Телеграм: [@irk24\\_link\\_bot](#)

эл.почта: [irkbabr24@gmail.com](mailto:irkbabr24@gmail.com)

Красноярск: Ирина Манская

Телеграм: [@kras24\\_link\\_bot](#)

эл.почта: [krasylar.babr@gmail.com](mailto:krasylar.babr@gmail.com)

Новосибирск: Алина Обская

Телеграм: [@nsk24\\_link\\_bot](#)

эл.почта: [nsk.babr@gmail.com](mailto:nsk.babr@gmail.com)

Томск: Николай Ушайкин

Телеграм: [@tomsk24\\_link\\_bot](#)

эл.почта: [tomsk.babr@gmail.com](mailto:tomsk.babr@gmail.com)

[Прислать свою новость](#)

#### ЗАКАЗАТЬ РАЗМЕЩЕНИЕ:

Рекламная группа "Экватор"

Телеграм: [@babrobot\\_bot](#)

эл.почта: [equatoria@gmail.com](mailto:equatoria@gmail.com)

#### СТРАТЕГИЧЕСКОЕ СОТРУДНИЧЕСТВО:

эл.почта: [babrmarket@gmail.com](mailto:babrmarket@gmail.com)

[Подробнее о размещении](#)

[Отказ от ответственности](#)

[Правила перепечаток](#)

[Соглашение о франчайзинге](#)

[Что такое Бабр24](#)

[Вакансии](#)

[Статистика сайта](#)

[Архив](#)

[Календарь](#)

[Зеркала сайта](#)