

Automatic positioning of application for patent in global patent space

Kravets A.G., Doctor of Technical Sciences, Professor of SAPR and Personal Computer chair

Dmitriy Mikhaylovich Korobkin, Candidate of Technical Sciences, Associate Professor of SAPR and Personal Computer chair

Dykov D.A., graduate student of SAPR and personal computer chair

*Ukustov A.S., graduate student of SAPR and personal computer chair
Volgograd State Technical University*

The authors describe the result of the pilot project «E-patent Examiner» within the frame of the system prototype which was created and it was hosted in the «cloud» of the server cluster. For the experiments in the prototype knowledge base a gallery of patents from subject areas such as «electricity», «physics», «mechanics» of the Russian Federation and the United States patent arrays was loaded.

Keywords: Natural Language «Cloud» Processing, Semantic analysis, Topic modeling, statistic analysis, global prior-art search.

УДК 303.447

О ЗАВИСИМОСТИ РАСПРЕДЕЛЕНИЯ ГОЛОСОВ ОТ ЯВКИ

*Андрей Юрьевич Бузин, канд. физ.-мат. наук, канд. юрид. наук,
доц. кафедры нелинейного анализа и оптимизации факультета
физико-математических и естественных наук
Тел.: +7(916) 115-75-15, e-mail: abuzin@votas.ru
Университет дружбы народов РУДН
<http://www.rudn.ru>*

Зависимость распределения голосов от явки является предметом острых дискуссий среди исследователей электоральной статистики. В статье сравниваются некоторые показатели такой зависимости. Особое внимание уделено показателю Собянина-Суховольского, применяемому во многих исследованиях для оценки уровня фальсификаций. На примере конкретных выборов проиллюстрировано, что сам этот показатель варьируется в зависимости от выборки избирательных комиссий.

Ключевые слова: выборы, явка, критерий Собянина-Суховольского elections, turnout, Sobyenin-Suhovolskiy criterium

Суть обсуждаемой проблемы

Изучение качества голосования и подсчета голосов часто заключается в исследовании распределения на декартовой плоскости с координатами (τ, u) точек, соответствующих участковым избирательным комиссиям (УИК), где τ – значение явки в УИК, а u – доля голосов (от общего числа избирателей УИК), полученных определенным претендентом в этой УИК [Собянин, 1995; Муягков, 2009; Кобак, 2012; Klimek, 2012; Бузин, 2008]. Так, например, А. Собянин и В. Суховольский, а вслед за ними и многие другие исследователи, изучают наклон a регрессионной прямой, проведенный через эти точки. Если наклон a существенно отличается от доли голосов (от числа проголосовавших избирателей) v , набранных претендентом, то обычно делается вывод о фальси-

фикации итогов голосования путем незаконного приписывания дополнительных голосов одному из претендентов.

При этом достаточное условие отклонения a от v используется как необходимое: действительно, при фальсификации итогов голосования a будет отклоняться от v в большую сторону, однако приписывание голосов не является единственной причиной такого отклонения. Любая зависимость распределения голосов от явки будет давать такое отклонение. В связи с этим представляется важным исследование вопроса о том, в какой мере зависимость распределения голосов от явки определяется прямыми фальсификациями, а в какой – другими причинами.



А.Ю. Бузин

Цель работы

В работе дан сравнительный анализ различных показателей зависимости распределения голосов от явки. Показано, что хотя показатели ведут себя по-разному, у них есть общие черты. При этом показатель Собянина-Суховольского удобен не только для исследования наличия прямых фальсификаций, но и для исследования зависимости голосования за определенного кандидата от явки, которая в случае конкурентных выборов проявляет немонотонное поведение. Показано, что при использовании коэффициента Собянина-Суховольского следует обращать внимание на его зависимость от выборки комиссий.

Показатели зависимости распределения голосов от явки

Коэффициент a является не единственным показателем, который характеризует усредненную зависимость распределения голосов от явки. Для описания этих показателей будем использовать следующие обозначения.

Явку в i -й избирательной комиссии будем обозначать как

$$(1) \quad \tau_i = \frac{V_i}{W_i}, i = \overline{1, N}$$

где N – рассматриваемое число избирательных комиссий, V_i – число избирателей, принявших участие в выборах в i -й избирательной комиссии (в России это число определяется по числу подписей избирателей за получение избирательного бюллетеня), W_i – общее число избирателей, внесенное в список избирателей i -й избирательной комиссии (списочный состав избирателей).

Доля голосов, полученная j -м претендентом, вычисляется либо как результат деления полученных им голосов на списочный состав избирателей:

$$(2) \quad u_{ij} = \frac{v_{ij}}{W_i}, i = \overline{1, N}, j = \overline{1, M}$$

либо как результат деления полученных им голосов на число избирателей, принявших участие в выборах (обычно незначительно отличающегося от числа бюллетеней, извлеченных из избирательных ящиков):

$$(3) \quad v_{ij} = \frac{V_{ij}}{V_i}, i = \overline{1, N}, j = \overline{1, M}$$

Здесь M – число претендентов в избирательном бюллетене, V_{ij} – число голосов, полученных j -м претендентом в i -й комиссии.

Упорядочим УИК по увеличению явки, то есть, перенумеруем УИК таким образом, чтобы $\tau_1 \leq \tau_2 \leq \dots \leq \tau_N$.

Ранее упомянутый показатель – коэффициент регрессии a_j – определяется для значения τ_K , стоящего на K -м месте по формуле [Минько, 2004, с.91]

$$(4) \quad a_j(\tau_K) = \frac{K \sum_{i=1}^K u_{ij} \tau_i - (\sum_{i=1}^K \tau_i)(\sum_{i=1}^K u_{ij})}{K \sum_{i=1}^K \tau_i^2 - (\sum_{i=1}^K \tau_i)^2}$$

В силу (1)-(3) величину $a_j(\tau_K)$ можно представить как

$$(4') \quad a_j(\tau_K) = \frac{K \sum_{i=1}^K v_{ij} \tau_i^2 - (\sum_{i=1}^K \tau_i) (\sum_{i=1}^K v_{ij} \tau_i)}{K \sum_{i=1}^K \tau_i^2 - (\sum_{i=1}^K \tau_i)^2}$$

Доля голосов, полученных j -м претендентом в первых K упорядоченных по явке УИК вычисляется следующим образом:

$$(5) \quad v_j(\tau_K) = \frac{\sum_{i=1}^K v_{ij}}{\sum_{i=1}^K v_i} = \frac{\sum_{i=1}^K v_{ij} V_i}{\sum_{i=1}^K V_i} = \frac{\sum_{i=1}^K v_{ij} \tau_i W_i}{\sum_{i=1}^K \tau_i W_i}$$

Для характеристики зависимости распределения от явки можно также использовать так называемый Т-линейный фильтр, то есть, следующее осреднение:

$$(6) \quad q_j^T(\tau_K) = \frac{1}{T} \sum_{i=K-T+1}^K v_{ij}$$

В работе [Kobak, 2012] для осреднения зависимости распределения от явки используется разбиение всего интервала явок на 1-процентные интервалы $[0;1\%), [0;2\%), \dots, [0;100\%)$ точками $t_0=0, t_1=1\%, \dots, t_{100}=100\%$. Каждая УИК попадает в одну из 100 групп в соответствии с тем, какому интервалу принадлежит ее явка. УИК со 100-процентной явкой попадают в отдельную 101-ю группу. После этого можно получить осреднённое распределение голосов по таким группам:

$$(7) \quad r_j(t_k) = \frac{\sum_{i: t_k \leq \tau_i < t_{k+1}} v_{ij}}{\sum_{i: t_k \leq \tau_i < t_{k+1}} v_i} = \frac{\sum_{i: t_k \leq \tau_i < t_{k+1}} v_{ij} \tau_i W_i}{\sum_{i: t_k \leq \tau_i < t_{k+1}} \tau_i W_i}$$

Из формул (4'),(5)-(7) видно, что если v_{ij} не зависит от i и равно просто v_j , то все четыре показателя зависимости будут равны v_j . На практике v_{ij} довольно сильно различается для разных УИК. Естественно, что никакой прямой зависимости v_{ij} от явки не наблюдается, и если мы построим график $v_{ij}(\tau_k)$, то получим сильно колеблющуюся кривую.

Именно поэтому имеет смысл изучать различного вида осреднения v_{ij} . Все четыре перечисленных осреднения ведут себя по-разному, что проиллюстрировано ниже на примере московских выборов (рисунок 1).

Были также исследованы зависимости указанных показателей для других выбо-

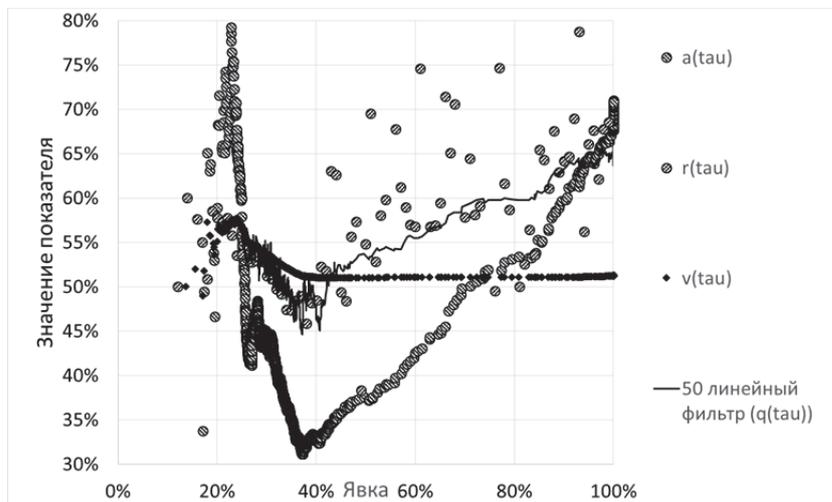


Рис. 1. Различные показатели зависимости доли голосования за Собянина на выборах Мэра Москвы в 2013 году совпадают (исключение составляет показатель $v(\tau)$). Можно с уверенностью говорить о возможной зависимости распределения голосов от явки.

Во-вторых, разные показатели могут быть использованы для разных целей: если $q(\tau)$ и $v(\tau)$ достаточно близки к реальным значениям доли голосов, набранных претендентом (также, как и $r(\tau)$ на интервале концентрации комиссий), то $a(\tau)$ иллюстрирует устойчивые тенденции зависимости доли голосов, набранной претендентом, от явки.

В-третьих, само по себе значение $a(\tau)$, исчисленное по всем комиссиям, не даёт возможности однозначно заявлять о фальсификации результатов. Имеет смысл иссле-

ров, в том числе для тех, для которых с большой долей уверенности можно говорить о наличии или отсутствии прямых фальсификаций.

Выводы

Во-первых, различные показатели зависимости распределения голосов от явки имеют некоторые общие черты, а именно: их промежутки монотонности, в основном,

довать поведение $a(\tau)$ в зависимости от явки. Подозрительным является постоянный рост этого коэффициента (для поддерживаемого администрацией претендента), начиная с областей явки, в которой концентрируются УИК.

В-четвертых, реальная конкуренция приводит к значительным отклонениям $a(\tau)$ от монотонного поведения. По виду кривых $a(\tau)$, а также $q(\tau)$ и $v(\tau)$ можно судить о том, какие области явки были более или менее выгодны для определенного претендента.

Литература

1. Собянин А.А., Суховольский В.Г. Демократия, ограниченная фальсификациями: Выборы и референдумы в России в 1991-1993 гг. – М., 1995
2. Myagkov M., Ordeshook P.C., Shakin D. The forensics of election fraud: Russia and Ukraine. - Cambridge, Cambridge university press, xiv, 2009
3. Kobak D., Shpilkin S., Pshenichnikov M.S. Statistical anomalies in 2011–2012 Russian elections revealed by 2D correlation analysis, arXiv:1205.0741v2 [physics.soc-ph] 17 May 2012
4. P. Klimek, Y. Yegorov, R. Hanel, and S. Thurner. Statistical detection of systematic election irregularities// Proceedings of the National Academy of Sciences of the United States of America, 2012, vol. 109 no. 41. [Электронный ресурс]. URL: www.pnas.org/cgi/doi/10.1073/pnas.1210722109
5. Бузин А.Ю., Любарев А.Е. Преступление без наказания: Административные избирательные технологии федеральных выборов 2007-2008 годов. – М.: ЦПК «НИККОЛО М»; Центр «Панорама», 2008. – 284с.
6. Минько А.А. Статистический анализ в MS Excel. – М.: ИД «Вильямс», 2004.

Votes distribution dependence on turnout

*An drey Yurevich Buzin, Candidate of Physical and Mathematical Sciences chair, Candidate of Jurisprudence, Associate Professor of Nonlinear analysis and optimization chair
Physical and Mathematical and Natural Sciences faculty, Associated Professor of Peoples' Friendship University of Russia*

Votes distribution dependence of turnout is the subject of intense debate among electoral statistics researchers. The article compares some indicators of such dependence. Particular attention is given to Sobyenin-Suhovolskiy indicator that is used in many studies to assess the level of fraud. Examples of specific election illustrate that this indicator varies depending on the election commissions sample.

Keywords: elections, turnout, criterion Sobyenin-Sukhovolsky.

УДК 001

ИННОВАЦИОННОЕ СОТРУДНИЧЕСТВО РОССИИ И АВСТРИИ: ОПЫТ ЭЗАН

Валерий Иванович Шевченко, д-р.техн. наук, проф.

Тел. +43 1 664-13- 66- 472, e-mail: ezan.vienna@chello.at

Представительство экспериментального завода научного приборостроения со специальным конструкторским бюро (ЭЗАН) в Австрии

<http://www.ezan.ac.ru>

Статья об инновационном сотрудничестве государств между Россией и Австрией. Обсуждается деятельность ЭЗАН инновационного центра делового сотрудничества в Вене, передачи технологий и развития совместных проектов. Анализируются перспективы направления двустороннего сотрудничества в разработке новых технологий и приводятся некоторые примеры сотрудничества.

Ключевые слова: наука, технологии, инновации, международное сотрудничество, программы, проекты.

«Стратегия инновационного развития Российской Федерации на период до 2020 года» предусматривает переход экономики страны на инновационный путь разви-