
МНОГОМЕРНЫЙ АНАЛИЗ ДАННЫХ

Е.С. Большов
(Киев)

ДИСКРИМИНАНТНЫЙ АНАЛИЗ В ПРОГНОЗИРОВАНИИ ПОВЕДЕНИЯ НЕОПРЕДЕЛИВШИХСЯ ИЗБИРАТЕЛЕЙ

В статье анализируются возможности метода дискриминантного анализа для прогнозирования поведения неопределившихся избирателей в предвыборных опросах. Описываются формальные и содержательные условия применимости метода. Приводятся результаты прогнозирования в ситуациях проведения выборов Президента Украины, Верховной Рады и мэра Киева.

Ключевые слова: дискриминантный анализ, предвыборный опрос, неопределившиеся избиратели, прогнозирование результатов выборов.

Постановка исследовательской задачи

Типичной задачей электоральных исследований является предсказание поведения избирателей, которые на момент последнего предвыборного опроса не определились с кандидатом, за которого собираются голосовать. Например, украинский опыт показывает, что доля таких избирателей может составлять 30-35%. Ошибки прогнозирования результатов выборов зачастую превышают ожидаемые вследствие выборочного характера социологических исследований, что отмечается в ряде работ [1; 2]. Во мно-

Евгений Сергеевич Большов – аналитик Киевского центра политических исследований и конфликтологии, аспирант Национального университета «Киево-Могилянская Академия». E-mail: eugen.bolshov@gmail.com.

гом эти ошибки связаны с неверным прогнозированием поведения неопределившихся избирателей, т.е. тех, кто на вопросы социологов ответил, что он еще не определился с выбором [3; 4].

В практике электоральных исследований используется, по крайней мере, два способа предсказания поведения таких избирателей. *Первый*, самый распространенный, *способ* состоит в том, что все рейтинги следует рассчитывать только по отношению к определившимся избирателям, т.е. неопределившиеся избиратели вообще не учитываются. Неявно это означает, что исследователь предполагает, что голоса неопределившихся избирателей распределены в тех же пропорциях, что и голоса избирателей, сделавших свой выбор и сообщивших о нем исследователю. Предполагается, что электоральные симпатии неопределившихся избирателей совпадают с симпатиями тех, кто уже сделал свой выбор. Вместе с тем, ряд авторов [5; 6] указывают, что такое допущение зачастую оказывается неверным, что и приводит к значительным ошибкам в прогнозах результатов выборов.

Второй способ опирается на прогнозирование поведения неопределившихся избирателей с помощью построения статистических прогностических моделей, параметрами которых являются различные характеристики респондентов (политические, социально-демографические и т.д.). В широком смысле первый способ можно считать частным случаем второго способа, если ответы всех неопределившихся избирателей считать пропущенными данными. Примером простейшей модели прогнозирования поведения неопределившихся избирателей может выступать распределение их голосов на основе партийной идентификации. Например, для США это определяется посредством ответов на вопрос: «*Считаете ли Вы себя республиканцем (демократом?)*». Заметим, что этот способ может быть использован только в странах с четким партийным структурированием и стабильной партийной системой, которые предполагают четкую идентификацию избирателей с партиями, которые они поддерживают (например, Англия

и США). В постсоветских странах вследствие слабой идентификации избирателей с партиями этот способ вряд ли может быть эффективным.

Возможны и другие модели. Как показали исследования [3; 4; 6] и, в частности, наш опыт построения подобного рода моделей, использование широкого спектра характеристик избирателей при статистическом моделировании их вероятного поведения приводит к значительному улучшению качества прогнозов результатов выборов. Инструменты моделирования могут быть различными. В большинстве известных нам публикаций используется дискриминантный анализ, однако, применяется также логистическая регрессия и алгоритмы построения дерева решений, например, CHAID¹.

Наша *исследовательская задача* состояла в изучении возможностей использования статистического моделирования, а именно, дискриминантного анализа для улучшения качества прогнозирования результатов выборов посредством моделирования и предсказания поведения неопределившихся избирателей. При этом основной методологической предпосылкой служило предположение о том, что неопределившиеся избиратели голосуют так же, как и определившиеся избиратели, близкие к ним по переменным, влияющим на электоральный выбор. Например, логично предположить, что

¹ CHAID (CHi-squared Automatic Interaction Detector) – один из алгоритмов построения дерева решений, который используется как для прогнозирования (отнесения объекта к тому или иному классу на основе его характеристик), так и для поиска взаимосвязей между переменными. С помощью CHAID исследователь может обнаружить статистические закономерности попадания объекта в тот или иной класс и, соответственно, использовать эти закономерности для классификации объектов, принадлежность которых к тому или иному классу неизвестна. В нашем случае такими объектами выступают неопределившиеся избиратели, которых нужно классифицировать – отнести к сторонникам какой-либо партии (кандидата).

неопределившиеся избиратели старшего возраста, проживающие в Восточном регионе Украины и выступающие за возврат к плановой экономике и укрепление отношений Украины и России, с высокой вероятностью отдадут свой голос коммунистам, как и их «двойники» из числа определившихся респондентов.

Тогда определив, какие переменные влияют на электоральный выбор и как по распределению этих переменных отличаются электораты разных кандидатов, становится возможным отнесение каждого неопределившегося избирателя к той части электората, «расстояние» до которой по включенным в анализ переменным меньше.

Основания выбора переменных

Для отбора переменных, разделяющих электораты, и отнесение неопределившихся избирателей к тому или иному из них нами был использован метод дискриминантного анализа, описанный в работах [7; 8, с. 346–368; 9; 10].

Для построения искомой модели необходимы переменные, которые статистически значимо связаны с электоральным выбором избирателей. Рассмотрим имеющийся опыт и попытаемся оценить, использование каких переменных является наиболее эффективным.

Прежде всего, остановимся на анализе переменных, связанных с *отношением избирателей к кандидатам в президенты и оценкой их способности решать важные проблемы страны*. В частности, в работе [3] использовалась оценка респондентами способности каждого из кандидатов не допустить вступления США в войну и уменьшить темпы инфляции. Аналогичный подход использовался нами для прогноза результатов внеочередных выборов мэра Киева в 2008 г.¹ При этом мы сравнивали (контролиро-

¹ Проводился телефонный опрос 5-11 мая 2008 г., репрезентативный относительно жителей Киева, объем выборки 800 респондентов. Опрос проведен Киевским международным институтом социологии и Центром социальных и по-

вали) число неотвеченных и неопределенных ответов (вида «*трудно сказать*», «*ничего не могу сказать*») для двух групп. Первая – те, кто определился (обозначим их как *группу О*) с выбором кандидата, за которого они будут голосовать. Вторая – те, кто еще не определился с выбором (обозначим их как *группу Н*). Результаты сравнения представлены в табл. 1 и свидетельствуют о том, что эти группы незначительно отличаются по количеству отказов от ответа и по числу неопределенных ответов на интересующие нас вопросы (переменные). Так, в *группе Н* доля дающих неопределенные ответы была выше, но незначительно – от 4 до 7 процентных пунктов.

Таблица 1

ЧИСЛО ПРОПУЩЕННЫХ ЗНАЧЕНИЙ В ГРУППАХ
ОПРЕДЕЛИВШИХСЯ И НЕОПРЕДЕЛИВШИХСЯ ИЗБИРАТЕЛЕЙ
(ВЫБОРЫ МЭРА КИЕВА В 2008 г.)

Варианты ответов на вопрос: «Скажите, пожалуйста, как, по Вашему мнению, следующие кандидаты в мэры могут справиться с такими задачами?» о Л. Черновецком	Доля респондентов, которые дали ответы вида «трудно сказать», «ничего не могу сказать», % к числу опрошенных	
	Группа О	Группа Н
Борьба с земельными спекуляциями, незаконными выделениями земли	26	33
Борьба с транспортными заторами, «пробками», строительство новых дорог	30	34
Помощь малообеспеченным людям	29	35
Повышение качества коммунальных услуг, улучшение работы жексов	32	37
Сохранение исторического облика города	29	36

литических исследований «СОЦИС» по заказу Киевского центра политических исследований и конфликтологии.

Наш опыт показывает, однако, что не всегда подобные вопросы (переменные) могут быть использованы для моделирования поведения неопределившихся избирателей, так как зачастую возникает ситуация, когда доля неответов на соответствующие вопросы может быть значительной, что и наблюдается по данным табл. 1.

Дополнительно проиллюстрируем эту ситуацию на примере анализа второго тура выборов Президента в 1999 г. в Украине, в котором участвовали действующий на тот момент Президент Леонид Кучма и его оппонент, лидер Коммунистической партии Украины, Петр Симоненко. На подвыборке определившихся респондентов нами была построена модель электорального выбора. В нем зависимая переменная, имеющая номинальный уровень измерения, принимала два значения в зависимости от голосования за одного из кандидатов. В качестве независимых переменных служили различные характеристики респондентов, которые гипотетически влияют на выбор кандидата. Были использованы рекомендуемые в работе [3] переменные, а именно: удовлетворенность возможной победой того или иного кандидата на выборах и оценка их способности навести порядок в стране.

Удовлетворенность оценивалась по семибалльной порядковой шкале от «крайне удовлетворен» до «крайне неудовлетворен» как ответ на вопрос:

«Если победит Леонид Кучма (Петр Симоненко), в какой мере Вы были бы удовлетворены или не удовлетворены его победой?»

Оценка способности также определялась по семибалльной порядковой шкале от «крайне низко» до «крайне высоко» как ответ на вопрос:

«Скажите, пожалуйста, насколько высоко или низко Вы оцениваете способность Леонида Кучмы (Петра Симоненко) навести порядок в стране?»

Модель дискриминантного анализа с использованием этих переменных оказалась статистически значимой для подвыборки определившихся респондентов. Для оценки возможностей класси-

фикации с помощью построенной модели была реализована процедура кросспроверки. Дискриминантные и классификационные функции оценивались на одной части выборки (так называемая «учебная» выборка), а качество классификации – на другой части выборки, которая не принимала участие в построении функций, – тестовой. Это делалось для того, чтобы качество классификации оценивалось на новых объектах, а не на тех, которые использовались для вычисления значений дискриминантных и классификационных функций. В противном случае качество классификации с помощью модели было бы переоценено. На тестовой выборке качество классификации было высоким – 96% определившихся респондентов были верно классифицированы в «свои» группы.

Вместе с тем, оказалось, что не все объекты изучаемой совокупности имеют значения по этим переменным. Так, неопределившиеся респонденты в большинстве своем не могли ответить на эти вопросы. Для этой группы (*группа Н*) доля пропущенных значений (см. табл. 2), в число которых включены отказы от ответа и ответы вида «трудно сказать», «ничего не могу сказать», превосходила аналогичный показатель по группе определившихся респондентов (*группа О*). В зависимости от характера вопроса эти доли колеблются от 45 до 64 процентных пунктов.

Таблица 2

ЧИСЛО ПРОПУЩЕННЫХ ЗНАЧЕНИЙ В ГРУППАХ
ОПРЕДЕЛИВШИХСЯ И НЕОПРЕДЕЛИВШИХСЯ ИЗБИРАТЕЛЕЙ
(ВТОРОЙ ТУР ВЫБОРОВ ПРЕЗИДЕНТА УКРАИНЫ В 1999 г.)

Переменная	Доля отказов от ответа, % к числу опрошенных		Доля ответов «трудно сказать», «ничего не могу сказать», % к числу опрошенных	
	Группа О	Группа Н	Группа О	Группа Н
Удовлетворенность победой Кучмы	1	17	17	47

Окончание табл. 2

Переменная	Доля отказов от ответа, % к числу опрошенных		Доля ответов «трудно сказать», «ничего не могу сказать», % к числу опрошенных	
	Группа О	Группа Н	Группа О	Группа Н
Удовлетворенность победой Симоненко	4	23	13	58
Сможет ли Кучма навести порядок в стране?	2	13	29	33
Сможет ли Симоненко навести порядок в стране?	10	28	19	33

Эти результаты показывают, что такие переменные не могут быть использованы для прогнозирования поведения неопределившихся респондентов. Причина в том, что пропущенные значения переменных, входящих в дискриминантные, а следовательно, и в классификационные функции, делают невозможным оценку вероятностей попадания избирателя в ту или иную группу. Очевидно, следует использовать другие, статистически значимо связанные с электоральным выбором переменные, которые не имели бы такого большого числа пропущенных данных в группе неопределившихся избирателей.

Таковыми переменными оказались: *оценка текущей ситуации в стране, оценка экономической ситуации в стране, предпочтение плановой или рыночной экономики, возраст респондента и регион проживания.*

Мы использовали эти переменные для проведения пошагового дискриминантного анализа с исключением (это процедура, в ходе которой на первом этапе в анализ включаются все интересующие исследователя переменные и потом исключаются те, которые вносят статистически незначимый вклад в дискримина-

цию). В результате была построена новая дискриминантная функция. Качество классификации на тестовой выборке было ниже, чем при использовании первой функции – 74% определившихся респондентов были верно классифицированы в «свои» группы. Но анализ количества пропущенных значений по переменным, использованным во второй функции, показал, что в отличие от первого набора (отношения к победе того или иного кандидата) *группа Н* и *группа О* имеют примерно одинаковую долю пропущенных значений и ответов вида «*трудно сказать*», «*ничего не могу сказать*».

Таким образом, в случае, когда *группы Н* и *О* значительно отличаются по доле пропущенных значений и неопределенных ответов на вопросы об удовлетворенности возможной победой того или иного кандидата и оценке их способности навести порядок в стране, следует использовать переменные, которые напрямую не измеряют отношение к кандидатам, но влияют на электоральный выбор.

Перейдем к рассмотрению переменных, связанных с *отношением к ряду ведущих политиков*. Мы использовали эти переменные при прогнозировании результатов выборов в Парламент Украины в 2002 г. и выборов Президента Украины в 2004 г. В обоих случаях был проведен факторный анализ вопросов об отношении к 8-10 ведущим политикам для уменьшения числа переменных.

В 2002 г. нами было получено факторное решение из четырех факторов: первый фактор был проинтерпретирован как отношение к правым политикам, второй – к центристским провластным политикам, третий – к левым политикам, четвертый определял позицию респондента на шкале «власть – оппозиция». Что касается числа пропущенных значений по таким переменным, то в *группе Н* их было на 5-15 процентных пунктов больше, чем в *группе О*, но этот показатель не достигал 50-70%, как в случае выборов 1999 г.

Следующая группа переменных, представляющих интерес с позиции влияния на электоральное поведение, связана с *информацией о дискуссионной сети респондента*. Как указывается в ряде работ [10; 11; 12; 13], избирателей можно рассматривать не как изолированных актеров, решение которых зависит не только от собственных характеристик и политических предпочтений, но и от контекста, в котором они находятся. В качестве такового, например, может служить разный состав социальных сетей, в которые включен избиратель [10]. В работах [11; 14] продемонстрировано, что на электоральный выбор избирателя сильно влияет позиция тех людей, с которыми он обсуждает политические проблемы.

В этой связи имеет смысл обратиться к исследованию, проведенному перед первым туром выборов Президента Украины в 2004 г.¹ В нем респондента просили вспомнить четырех людей, с которыми он чаще всего за последний месяц обсуждал важные для него проблемы. После этого задавался вопрос о том, будут ли эти люди принимать участие в первом туре голосования, и если да, то кому они собираются отдать свой голос². В процессе анализа данных *число сторонников кандидата в президенты* (среди собеседников респондента) использовалось как некоторая *прокси-переменная* для измерения уровня социального влияния в пользу кандидата. Такие переменные оказались статистически значимыми в модели дискриминантного анализа, и, по нашему мнению, их следует применять при статистическом моделировании поведения неопределившихся респондентов.

¹ Репрезентативное для Украины исследование, проведенное методом face-to-face в октябре 2004 г. Объем выборки – 2000 респондентов. Опрос был проведен Киевским международным институтом социологии по заказу Киевского центра политических исследований и конфликтологии.

² Как указывается в некоторых работах [см., например, 12] данные о голосовании собеседников респондента, полученные от него, хорошо согласуются с результатами опроса самих собеседников. Тем самым, эта информация достаточно валидна.

Результаты оценки точности прогноза поведения неопределившихся избирателей

Так как инструментом моделирования поведения неопределившихся избирателей в нашем случае является дискриминантный анализ, то в качестве первого шага оценки точности метода логично подсчитать стандартный показатель для оценки качества построенных дискриминантных и классификационных функций на подвыборке определившихся респондентов (*группа О*). А именно, сравнить число объектов, верно классифицированных с помощью дискриминантных функций, с числом верно классифицированных объектов в случае, если бы отнесение объектов к группам осуществлялось случайным образом. Посредством этого показателя можно оценить, насколько классификация с помощью дискриминантного анализа лучше (или хуже), чем классификация, полученная случайным образом.

Если точность первой из классификаций хотя бы на 25% лучше второй, есть смысл использовать дискриминантный анализ для прогнозирования поведения неопределившихся избирателей [7]. Вместе с тем это неполная оценка точности метода, поскольку она рассчитана на подмассиве избирателей, чей выбор известен. Очевидно, необходимо оценить, как хорошо построенная модель дискриминантного анализа предсказывает поведение неопределившихся избирателей, что существенно сложнее. Рассмотрим некоторые из возможных подходов к получению такой оценки.

Панельный поствыборный опрос. Один из возможных источников информации о том, как проголосовали неопределившиеся, – это поствыборный опрос на той же выборке избирателей, в котором респондентов спрашивают о том, участвовали ли они в голосовании и если да, то кому отдали свой голос. Отметим, что информация, полученная в панельных исследованиях на политическую тему, имеет ряд недостатков, и на это указывают многие авторы [например, 15]. Во-первых, действует эффект участия в

первой волне исследования, что приводит к активизации потребления политической информации в перерыве между исследованиями и влияет на поведение респондента (например, повышается вероятность принятия участия в выборах). Во-вторых, выбывание из панели происходит неслучайным образом (селективный характер неответов) и зависит от уровня интереса к политике. И, главное, в поствыборных электоральных исследованиях часто наблюдается феномен присоединения к победителю – большее число избирателей говорят о поддержке кандидата на прошедших выборах, чем действительно отдали ему свой голос [15].

Однако более валидных данных, чем результаты панельных исследований, в распоряжении исследователей зачастую нет. Мы оценивали качество статистического прогнозирования поведения неопределившихся избирателей с использованием данных панельного исследования после первого тура выборов Президента Украины в 2004 г. Наш анализ показал, что 67% (см. табл. 3) неопределившихся респондентов были верно классифицированы с помощью модели дискриминантного анализа (кандидат, за которого они должны были бы с наибольшей вероятностью проголосовать по результатам работы модели, был назван ими как тот, кому они отдали свой голос в ходе поствыборного исследования). Это позволяет считать статистическое моделирование поведения неопределившихся избирателей полезным инструментом улучшения качества прогнозов.

Таблица 3

ПОКАЗАТЕЛИ ТОЧНОСТИ ПРОГНОЗИРОВАНИЯ ПОВЕДЕНИЯ
ИЗБИРАТЕЛЕЙ (ПО ДАННЫМ ПАНЕЛЬНОГО ИССЛЕДОВАНИЯ
ВЫБОРОВ ПРЕЗИДЕНТА УКРАИНЫ В 2004 г.)

Группа	% верно классифицированных
Определившиеся респонденты (предвыборное исследование)	83
Определившиеся респонденты (поствыборное исследование)	75

Окончание табл. 3

Группа	% верно классифицированных
Неопределившиеся респонденты (поствыборное исследование)	67

Еще один способ оценки качества статистического прогнозирования поведения неопределившихся избирателей – сравнение прогноза результатов выборов без попыток предсказать поведение таких респондентов с прогнозом, полученным с учетом моделирования. Другими словами, сравнивается точность прогноза, полученная с помощью модели, с точностью прогноза, который опирается на гипотезу о том, что неопределившиеся избиратели поддерживают кандидатов в тех же пропорциях, что и определившиеся респонденты.

По результатам сравнения можно сделать вывод о том, что моделирование поведения неопределившихся избирателей улучшило прогноз, но это относится к случаю, когда их симпатии распределяются по-другому, чем среди определившихся респондентов. Например, моделирование возможного поведения неопределившихся избирателей перед выборами мэра Киева в 2008 г. показало, что среди *группы Н* потенциальных избирателей оппозиционных кандидатов (В. Кличко, А. Турчинов, Н. Катеринчук) было больше, чем возможных избирателей действующего мэра Л. Черновецкого (тогда как среди определившихся избирателей число его сторонников было больше, чем у трех главных оппозиционных кандидатов). Результаты выборов (см. табл. 4) показали, что без предварительного применения дискриминантного анализа прогноз был бы хуже. Произошла бы переоценка количества голосов, отданных за Л. Черновецкого и за кандидатов с незначительным «числом голосов», и недооценка сторонников трех оппозиционных политиков. Таким образом, учет отличий особенностей группы неопределившихся респондентов позволил сделать прогноз выборов более точным.

Таблица 4
 МОДЕЛЬНОЕ И РЕАЛЬНОЕ РАСПРЕДЕЛЕНИЕ ГОЛОСОВ (ВЫБОРЫ МЭРА КИЕВА В 2008 г.)

Электорат	Распределение голосов		Результаты выборов
	Голоса неопределившихся избирателей делятся пропорционально голосам определившихся избирателей	Для прогноза поведения неопределившихся избирателей используется дискриминантный анализ	
Леонида Черновецкого	39,8	35,1	37,7
Александра Турчинова, Николая Катеринчука и Виталия Кличко	35,4	44,4	43,7

Экзит-поллы. Еще одним способом косвенной оценки качества статистического прогнозирования поведения неопределившихся избирателей является использование данных экзит-поллов. В инструментарии экзит-поллов в Украине зачастую присутствовал вопрос о том, когда респондент принял решение о том, кому отдать свой голос. Тогда можно выделить группу избирателей, которые отвечают, что они «определились прямо на избирательном участке», «в последний день перед голосованием» и «в последнюю неделю до голосования», считая, что в последнем предвыборном исследовании эти респонденты были неопределившимися. Тем самым можно сопоставить результаты статистического прогнозирования поведения этой группы избирателей в предвыборном опросе с их ответами в экзит-полле. Такой способ был использован на парламентских выборах 2002 г. в Украине. Как видно из табл. 5, прогнозные расчеты по модели хорошо совпадают с результатами экзит-полла.

Необходимые условия корректности статистического моделирования

Для корректного использования статистического моделирования электорального поведения необходимо, чтобы выполнялось второе условие однородности, сформулированное Ю.Н. Толстой [16; 17; 18], – для всех групп рассматриваемой совокупности (в нашем случае для определившихся и неопределившихся респондентов) математическая зависимость должна быть одинаковой. Очевидно, что если факторы, которые влияют на принятие решения определившимися избирателями, отличаются от факторов, формирующих решение неопределившихся (на момент опроса, конечно) избирателей, построение подобных моделей неоправданно.

Чтобы проверить, выполняется ли это условие, перед выборами Президента в 2004 г. было построено два уравнения логисти-

Таблица 5
РАСПРЕДЕЛЕНИЕ ГОЛОСОВ НЕОПРЕДЕЛИВШИХСЯ ИЗБИРАТЕЛЕЙ – МОДЕЛЬНОЕ И ПО РЕЗУЛЬТАТАМ ЭКЗИТ-ПОЛЛА (ВЫБОРЫ ВЕРХОВНОЙ РАДЫ УКРАИНЫ В 2008 г.)

Группа партий	Прогнозируемое распределение голосов по модели	Распределение голосов в экзит-полле
«Коммунистическая партия Украины», «Блок Натальи Витренко»	21	18
«Наша Украина», «Социалистическая партия Украины», «Блок Юлии Тимошенко»	39	40
«За единую Украину», «Социал-демократическая партия Украины (объединенная)», «Команда озимого поколения», «Партия зеленых Украины», «Женщины за будущее»	41	42

ческой регрессии¹, в которых зависимой переменной было голосование за В. Януковича (по данным поствыборного опроса), независимыми – все те переменные, которые включались в модель дискриминантного анализа.

Обе модели, как для избирателей, которые определились на момент последнего опроса, так и для неопределившихся респондентов, оказались статистически значимыми². Если сравнить модели по показателям, аналогичным проценту объясненной дисперсии в линейной регрессии³, то модель для респондентов, которые уже определились, выглядит лучше (см. табл. 6).

Таблица 6

ЗНАЧЕНИЯ КОЭФФИЦИЕНТА (В) В МОДЕЛЯХ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ (ЗАВИСИМАЯ ПЕРЕМЕННАЯ – ГОЛОСОВАНИЕ ЗА В. ЯНУКОВИЧА)

Независимая переменная	Группа О	Группа Н
Количество сторонников Януковича в дискуссионной сети респондента	0,47	0,4
Отношение к России (по пятибалльной шкале)	0,41	0,47
Отношение к провластным политикам (значение соответствующего фактора)	0,91	0,89
Отношение к В. Ющенко и Ю. Тимошенко (значение соответствующего фактора)	-1,13	-1,21
Константа	-1,3	-1,3

¹ Это разновидность модели нелинейной регрессии, где независимые переменные могут иметь любой уровень измерения, а зависимая – номинальный и если она принимает только два значения, то имеет место бинарная логистическая регрессия.

² Chi-square первой модели 687, второй – 115, значимость обеих моделей – 0.000.

³ Например, такие показатели, как Cox & Snell R Square и Nagelkerke R Square.

Однако важна степень отличия моделей. В случае, когда коэффициенты при каких-то переменных значительно разнятся в моделях для определившихся и неопределившихся респондентов, очевидно, что эти переменные неодинаково влияют на электоральный выбор в различных группах избирателей. Сравнение моделей электорального выбора показывает, что значения коэффициента B достаточно близки для определившихся (*Группа O*) и неопределившихся (*Группа H*) избирателей. В этом случае можно предположить, что на электоральный выбор влияют одинаковые факторы с примерно равной силой. И, соответственно, условие корректного использования дискриминантного анализа для прогнозирования поведения неопределившихся респондентов можно считать выполненным. Вместе с тем из этого не следует, что на каждых выборах такая ситуация будет повторяться, поэтому одной из задач исследователя является проверка условия однородности перед применением статистического моделирования поведения неопределившихся респондентов.

ЛИТЕРАТУРА

1. *Freedman P., Goldstein K.* Building a Probable Electorate from Preelection Polls // *Public Opinion Quarterly*. 1996. Vol. 60. P. 574–587.
2. *Buchanan W.* Elections Predictions: An Empirical Assessment // *Public Opinion Quarterly*. 1986. Vol. 50. Iss. 2. P. 222–227.
3. *Fenwick J., Wiseman F., Becker J.F., Heiman J.R.* Classifying Undecided Voters in Preelection Polls // *Public Opinion Quarterly*. 1982. Vol. 46. Iss. 3. P. 383–391.
4. *Kim T.* Discriminant Analysis as a Prediction Tool for Uncommitted Voters in Preelection Polls // *International Journal of Public Opinion Research*. 1995. Vol. 7. Iss. 2. P. 110–127.
5. *Myers D.J., O'Connor R.E.* The Undecided Respondent in Mandatory Voting Settings: A Venezuelan Exploration // *The Western Political Quarterly*. September 1983. Vol. 36. No. 3. P. 420–433.
6. *Хмелько В.Е.* Об опыте прогнозирования // *Политические настроения накануне выборов*. Киев, 1998.
7. *Klecka W.R.* Discriminant Analysis. Beverly Hills: Sage Publication, 1980.
8. *Бююль А., Цефель П.* SPSS: искусство обработки информации. СПб: ДиаСофт, 2001.

9. Бессокирная Г.П. Дискриминантный анализ для отбора информативных переменных // Социология: методология, методы, математические модели. 2003. № 16. С. 25–35.

10. Pattie C., Johnston R. Context, Conversation, and Conviction: Social Networks and Voting at the 1992 British Election // Political Studies. 1999. P. 877–899.

11. Huckfeldt R., Sprague J. Discussant Effects on Vote Choice: Intimacy, Structure, and Interdependence // Journal of Politics. February 1991. Vol. 53. No. 1.

12. Sprague J. Is There a Micro Theory Consistent with Contextual Analysis // Strategies of Political Inquiry. London: Sage, 1982.

13. Huckfeldt R., Levine J., Morgan W., Sprague J. Election Campaign, Social Communication, and the Accessibility of Perceived Discussants' Preferences // Political Behavior. 1998. Vol. 20. No. 4.

14. Huckfeldt R., Sprague J. Discussant Effects on Vote Choice: Intimacy, Structure, and Interdependence // The Journal of Politics. February 1991. Vol. 53. No. 1.

15. Voogt R., Kempen V. Nonresponse Bias and Stimulus Effects in the Dutch National Election Study // Quality & Quantity. 2002. Vol. 36. P. 325–345.

16. Толстова Ю.Н. Логика математического анализа социологических данных. М., 1991.

17. Толстова Ю.Н. Обеспечение однородности социологических данных // Социологические исследования. 1986. № 3. С. 149–154.

18. Толстова Ю.Н. Принципы анализа данных в социологии // Социология: методология, методы, математические модели. 1991. № 1.