Д.Д. Козлов, К.Ю. Доронина

Обобщенные регрессионные модели в психологии: пример исследования фиксированной установки[1]

Аннотация: В статье анализируется место математического моделирования в системе методов психологии. Приводятся основные идеи метода наименьших квадратов, лежащего в основании большинства методов математического моделирования в общественных науках. Раскрывается сущность метода обобщенных регрессионных моделей (ОРМ) как объединения методов регрессионного и дисперсионного анализов. На примере экспериментального исследования фиксированной установки показывается, как использование метода ОРМ позволяет обнаружить дополнительные эмпирические факты и проверить гипотезы в тех случаях, в которых методы сравнения центральных тенденций и корреляционный анализ не позволяют этого сделать.

Ключевые слова: математическое моделирование, обобщенные регрессионные модели, метод наименьших квадратов, установка.

1. Методологические и математические основания обобщенных регрессионных моделей.

1.1. Математическое моделирование в системе психологического знания.

В основании любого научного знания[2] лежат факты. Как любой научный факт, психологический факт представляет собой некоторое

---

[1] Авторы выражают признательность д.пс.н., профессору каф. общей психологии СамГУ Агафонову А.Ю. и к.т.н., доценту каф. высшей математики СГАУ Белашевскому Г.Е. за ценные комментарии и замечания в процессе подготовки статьи.

[2] Мы исходим из естественнонаучного представления о знании, соответствующего англоязычному термину science. Мы полагаем, что академическая психология не является гуманитарной дисциплиной, в которой изучаются феномены, а не явления, и используются герменевтические методы, а не наблюдение и эксперимент.

явление, для которого определены условия его возникновения и методы его обнаружения при соблюдении этих условий. Как следует из сказанного, любой научный факт не существует вне теории, которая характеризует эти условия и методы. Тогда развитие научного знания можно описать как последовательное развитие и конкуренцию теорий, претендующих на возможно более полное описание и объяснение наблюдаемой фактологии. Механизм развития любой теории в науке – это поиск новых и опровержение существующих фактов. Для доказательства закономерности и неслучайности явлений, обнаруживаемых в опыте при использовании валидных методов измерения (т.е. для превращения явления в научный факт), используются математические методы. Традиционные методы психологического исследования (методы сравнения центральных тенденций и корреляционный анализ) вполне пригодны для этой цели. Установление научных фактов достаточно для проверки содержательной модели (т.е. теоретических построений). Зачем же тогда дополнительно выводить математическую модель, в чем смысл такого занятия для психолога-исследователя? Отметим по крайней мере пять причин в пользу использования математических моделей в психологии.

1. Математическое моделирование позволяет исследовать совместное влияние и взаимодействие нескольких переменных одновременно, т.е. позволяет дать целостное объяснение и интерпретацию для нескольких фактов одновременно.

2. Математические модели позволяют оценить величину взаимосвязи переменных, т.е. дают основания не только для качественного, но и количественного описания и дальнейшей интерпретации.

3. Методы математического моделирования являются более мощным исследовательским инструментом в том смысле, что позволяют обнаружить более тонкие различия даже в тех случаях, где применение традиционных для психологического исследования методов не позволяет обнаружить статистически достоверные результаты (этот факт будет проиллюстрирован ниже в настоящей работе).

4. Динамическая математическая модель[1] позволяет описать некоторое явление не только как *факт*, но и как *процесс*. Только понимание логики процесса позволяет сформулировать полноценные причинные гипотезы, перейти от исследования *предикторов* к пониманию иногда скрытых *причин* наблюдаемых явлений (Капрара, Сервон, 2003).

5. На основании математической модели возможно построение прогноза в отношении будущего, а также оценка точности и достоверности такого прогноза.

Более подробный анализ теоретических и методологических аспектов математического моделирования можно найти в работах Мышкиса (2007), Советова и Яковлева (2001), Самарского и Михайлова (2001) и др.

1.2. Метод наименьших квадратов: ключевые идеи.

В общем виде математическая модель выражает взаимосвязь между зависимыми и независимыми переменными в виде системы уравнений вида $y=F(X)$, где y – зависимая переменная, X – вектор независимых переменных, F – функция произвольного вида. Динамические системы в большинстве случаев описываются системами дифференциальных уравнений[2]. Знание функции взаимосвязи переменных позволяет вычислить (предсказать) значение зависимой переменной по известным значениям независимых переменных, включенных в уравнение (входящих в модель).

В психологии математические модели никогда не дают совершенно точного прогноза в силу погрешностей измерения, «зашумленности» данных, влияния неучтенных переменных и т.д., а также в силу вероятностного характера самих моделей. Тогда точное значение для каждого i-го случая можно выразить в виде уравнения:

[1] Не все математические модели являются динамическими.

[2] Идея о необходимости исследования динамических систем в психологии через их описания дифференциальными уравнениями впервые высказана Куртом Левиным (Левин. 2000).

$$y_i = F(X_i) + e_i, \qquad (1)$$

где e_i – ошибка модели, индивидуальный вклад всех неучтенных в модели факторов в значение y_i.

Целью метода наименьших квадратов является нахождение такой функции $F(X)$, чтобы величина ошибки e была по возможности наименьшей. С математической точки зрения это может быть выражено как поиск такой функции $F(X)$, для которой сумма квадратов ошибок модели

$$\sum_i e_i^2 = \sum_i (y_i - \hat{y})^2, \qquad (2)$$

где $\hat{y} = F(X_i)$, была минимальной.

Метод не накладывает никаких ограничений на вид функции $F(X)$. Так, линейная модель, лежащая в основании метода множественной регрессии, может быть выражена уравнением

$$y_i = C + \sum_j b_j x_j + e_i. \qquad (3)$$

Взаимодействие двух переменных может быть описано путем добавления в модель (3) слагаемых вида

$$b_{jk} x_j x_k, \; j \neq k. \qquad (4)$$

Аналогично можно описать взаимодействия переменных более высокого порядка. Дополнительная нелинейность вводится добавлением нелинейных членов либо другими нелинейными преобразованиями, однако общая идея всех методов регрессионного анализа (РА) остается неизменной.

Для оценки качества регрессионных моделей используются коэффициент множественной детерминации (КМД) $R^2 = D_{\hat{y}} / D_y$, коэффициент множественной корреляции (КМК) $R = \sqrt{R^2}$, а также уточненные КМД и КМК, в формулы которых вносится поправка на количество переменных в модели, что позволяет снизить побочное влияние шума на оценку качества модели [*Филатов*, 2006; *Гусев*, 2000]. Значимость КМД как соотношение двух дисперсий определяется по F-критерию. Статистическая значимость каждого из *b*-коэффициентов (достоверность отличия коэффициента от нуля) определяется по одновыборочному t-критерию Стьюдента.

Расписав дисперсии в формуле КМД и сократив в них общий знаменатель, получаем:

$$R^2 = \frac{D_{\hat{y}}}{D_y} = \frac{\dfrac{\sum_{i=1}^{N}(\hat{y}_i - M_{\hat{y}})^2}{N-1}}{\dfrac{\sum_{i=1}^{N}(y_i - M_y)^2}{N-1}} = \frac{\sum_{i=1}^{N}(\hat{y}_i - M_{\hat{y}})^2}{\sum_{i=1}^{N}(y_i - M_y)^2} = \frac{SS_{\hat{y}}}{SS_y}, \qquad (5)^{1}$$

где M_y – среднее значение (математическое ожидание) для переменной y, $M_{\hat{y}}$ – среднее значение оценок переменой y по значениям функции $F(x)$ (т.е оценки по модели).

На основании свойства аддитивности дисперсии, а также учитывая формулы (2) и (5), основную модель, описываемую уравнением (1), можно также представить в виде $D_y = D_{\hat{y}} + D_e$, или, избавившись от знаменателей дисперсий, в виде

$$SS_y = SS_{\hat{y}} + SS_e. \qquad (6)$$

Последняя формула полностью совпадает с основной формулой дисперсионного анализа (ДА), в которой общая (total) дисперсия раскладывается на внутригрупповую (within group, wg) и межгрупповую (between group, bg) дисперсии:

$$SS_{total} = SS_{bg} + SS_{wg}. \qquad (7)$$

Аналогично, КМД[2] в ДА равен

---

[1] SS в данном случае – принятое сокращение для обозначения суммы квадратов (от англ. sum of squares).

[2] При описании дисперсионного анализа обычно говорят просто о коэффициенте детерминации, а не о КМД, хотя в случае многофакторного и многомерного ДА по смыслу это именно коэффициент **множественной** детерминации. Во всех случаях мы используем термин КМД для сохранения единой терминологии.

$$R^2 = \frac{SS_{bg}}{SS_{total}}, \tag{8}$$

что совпадает с выражением (5). Собственно, в основании ДА лежит все та же линейная модель (3) с включением переменных взаимодействия факторов, аналогичных (4) [*Гусев*, 2000]. Однако если в РА все переменные являются количественными, то в ДА все независимые переменные являются категориальными. Это позволяет объединить методы РА и ДА в единый класс математических моделей, получивших название обобщенных регрессионных моделей (ОРМ), и использовать показатели КМД и КМК, а также их уточненные значения, как основные показатели качества таких моделей. К тому же исходные допущения о нормальности распределения всех переменных и гомогенности дисперсий для зависимых переменных одинаковы для методов РА и ДА (более подробное обсуждение допущений, необходимых для проведения ДА, см. у Гусева (2000) и Гласса и Стэнли (1976)). Теперь основные уравнения, лежащие в основании ДА, можно переписать в форме, принятой для РА. Так, основное уравнение однофакторного ДА

$$y_i = M_x + (M_j - M_x) + e_i,$$

где M_x – среднее значение для всей эмпирической выборки, M_j – среднее значение в j-й группе, к которой принадлежит i-е наблюдение, эквивалентно выражению

$$y_i = C + b_j + e_i. \tag{9}$$

Тогда в случае одной категориальной переменной, разбивающей всю выборку на m независимых групп, и n количественных независимых переменных значения зависимой переменной y_i, могут быть выражены следующим образом:

$$y_i = C_0 + b_j + \sum_{k=1}^{n} b_k x_k + \sum_{k=1}^{n} b_{jk} x_k + e_i, j = 1..m, \tag{10}$$

где коэффициенты b_j отражают независимый вклад категориальной переменной в объяснение (дисперсии) зависимой переменной, коэффициенты b_k – независимый вклад количественной переменной, коэффициенты b_{jk} - вклад взаимодействия категориальной переменной с количественной переменной. Значимость коэффициентов может быть также оценена по t-критерию, а общая значимость всей модели – по F-критерию. Общая доля объясненной дисперсии зависимой переменной, равная КМД[1], показывает объяснительную силу модели в целом, а отношение дисперсии значений каждого слагаемого к дисперсии зависимой переменной (а также стандартизированные β-коэффициенты) – вклад каждой переменной или их взаимодействия в отдельности.

Аналогичные соображения можно обобщить на любое количество категориальных переменных, а также включить в модель произвольную нелинейность для количественных переменных.

1.3. Область применения ОРМ

ОРМ позволяют исследовать зависимость количественной переменной от набора переменных любой природы (как количественных, так и категориальных), выраженной в произвольной форме.

В случае, если зависимая переменная является категориальной, а независимые – количественными, то в психологии традиционно используется метод дискриминантного анализа. Однако с математической точки зрения дискриминантный анализ аналогичен многомерному ДА, поскольку в его основании лежит все та же линейная модель (3), в которой зависимые и независимые переменные меняются местами [StatSoft, 2001]. Это означает, что дискриминантный анализ, по сути, является методом, обратным многомерному ДА, и, соответственно, после проведения дискриминантного анализа его результаты можно также представить в форме, традиционной для ОРМ. Таким образом, единственный класс задач, для которых ОРМ принципиально не применимы, – это случай, в котором **все** переменные являются категориальными. Однако для случая бинарных переменных существуют методы логистической регрессии, в которых переменная откли-

---
[1] более корректным будет использование значения уточненного КМД

ка (зависимая переменная) выражает вероятность того, что бинарная переменная примет то или иное значение, и которая основывается на той же линейной модели (3).

Во всех остальных случаях ОРМ являются гибким исследовательским инструментом. Значения КМД, F и t-критериев позволяют обосновать несколько методов для поиска наилучшей возможной модели и исключения из нее всех статистически незначимых (т.е. описывающих скорее шум, чем реальные зависимости) переменных [см. подробнее *Наследов*, 2004; StatSoft, 2001]. Таким образом, исследователь, использующий методы ОРМ, в некотором смысле подобен человеку, складывающему пазл: из множества кусочков-переменных ему необходимо сложить целостную картинку, при этом отложив в сторону все ненужные элементы и не упустив ни одной значимой переменной из первоначального набора.

Дальнейшее развитие методов ОРМ заключается в возможности использования различных функций, минимализирующих потери (а не только квадратичных ошибок в соответствии с формулой (2)). Другое направление современного развития ОРМ – метод моделирования структурными уравнениями [*Митина*, 2008].

2. Метод ОРМ в исследовании фиксированной установки

2.1. Постановка проблемы

Д.Н. Узнадзе определяет установку как целостное состояние субъекта, «которое предваряет появление определенных фактов осознания или предшествует им» [*Узнадзе*, 1966]. В его экспериментах фиксированная установка формировалась путем повторного предъявления окружностей разного размера, где правая окружность всегда была больше левой. После 10-15-кратного предъявления стимульного материала испытуемому в качестве целевого стимула предъявлялись равные окружности. Приблизительно у 70% испытуемых возникала иллюзия контраста: левый круг казался больше правого. Небольшой процент испытуемых составляли те, кто продолжал воспринимать стимульный материал адекватно или демонстрировали появление иллюзии, полностью соответствующей установочной серии (там же).

Дальнейшие исследования Узнадзе показали, что установочная серия может оказывать воздействие независимо от модальности, производимый эффект сохраняется в условиях постгипнотической амнезии. Более того, установочный эффект, сформированный в одной модальности, обнаруживается также и в других модальностях (Узнадзе, 2001). Тем не менее, существуют границы подобного воздействия. Эффект восприятия «по контрасту» возникал в большинстве случаев предъявления неопределенного стимула, однако когда картинка была противоположной (контрастное соотношение стимулов установочной серии и целевого стимула), испытуемые безошибочно дифференцировали истинное соотношение размеров окружностей. При ассимилятивном соотношении целевого и установочного стимулов результаты в целом соответствовали классическим экспериментам, однако не были столь однозначными. Эти и другие экспериментальные факты, обнаруженные Узнадзе, позволили обосновать следующие гипотезы:

1. При ассимилятивном соотношении установочного и целевого стимулов наличие установочной серии приводит к большим ошибкам в восприятии целевого стимула.

2. Существует граница величины зрительного стимула, после которой влияние установочной серии перестает быть значимым.

Соответствующие эмпирические гипотезы имеют следующую формулировку:

1. При наличии установочной серии и ассимилятивного соотношения установочного и целевого стимулов меньшее количество испытуемых будут верно воспринимать различия в величине окружностей целевого стимула.

2. При увеличении разницы в величине окружностей целевого стимула влияние подобной установочной серии на осознание различий в величине окружностей целевого стимула будет уменьшаться, а после достижения некоторого «порогового» значения и вовсе исчезнет.

2.2. Процедура и методы исследования

Общая выборка исследования составила 100 человек в возрасте от 14 до 50 лет с нормальным или скорректированным до нормального зрением. Вся выборка была также разделена на экспериментальную и контрольную группы (N_1=N_2=50) по критерию наличия или отсутствия установочной серии.

Для проведения эксперимента была использована компьютерная программа Subliminal Images. В качестве установочного стимула использовались всегда одинаковые пары окружностей, в которых правая окружность всегда была больше левой на 27 пикселей, или 7,128 миллиметров. Установочная серия представляла собой предъявление установочного стимула в хаотичном порядке в разных частях монитора с длительностью экспозиции в 300 мс. с интервалом между предъявлениями также в 300 мс. Установочная серия состояла из 100 предъявлений стимула, что соответствует общей длительности серии в одну минуту. Испытуемый располагался перед экраном монитора на расстоянии приблизительно 50 см. Во время стимуляции фоном служил белый экран.

Целевой стимул также представляет собой пару окружностей с подобным соотношением размера, однако не столь явно выраженным. Различие в размерах целевых стимулов, использованных в эксперименте и представленные в таблице 1, позволяют разбить всю выборку на пять подгрупп.

Таблица 1.

Разница между окружностями в контрольном стимуле в пикселях и миллиметрах

Размер (пиксели)	Размер (миллиметры)
1	0,264
3	0,792
5	1,32
7	1,848
10	2,64

При предъявлении целевого стимула испытуемым предлагалось сравнить величину окружностей. Если они воспринимались как разные, то регистрировался ответ, какая из окружностей представляется больше другой. В контрольной группе при отсутствии установочной серии испытуемые сразу отвечали на вопрос о соотношении размеров окружностей после предъявления целевого стимула. Количество испытуемых в каждой из подгрупп в контрольной и экспериментальной группах было одинаковым (n=10).

2.3. Результаты: использование критерия χ^2 Пирсона

Исследование имеет классический двухфакторный дизайн вида 2x5. Действительно, различение стимула (зависимая переменная) зависит от наличия установочной серии (категориальная переменная, две градации) и величины различий между окружностями в целевом стимуле (категориальная переменная, 5 градаций). Однако и зависимая переменная – различение величин окружностей в целевом стимуле – является также номинативной (две градации: есть различение и нет различения). Таким образом, в данной ситуации критерий χ^2 Пирсона является безальтернативным. Однако в данном случае мы сможем проверить только независимое влияние каждого фактора, в то время как вторая сформулированная гипотеза предполагает их взаимодействие.

Полученные результаты ставят под сомнение нашу первую гипотезу. В экспериментальной группе 66% испытуемых верно указали различия в целевом стимуле, в то время как в контрольной группе таких испытуемых оказалось лишь 48% ($\chi^2(1)=3,305$; $p=0,069$). Среди всех 100 испытуемых, принявших участие в эксперименте, не оказалось ни одного (!), кто сделал бы неверный вывод о соотношении окружностей в целевом стимуле, т.е. ни одного, у кого бы проявился контрастный эффект в восприятии. Влияние разницы в диаметрах окружностей целевого стимула оказалось существенно большим ($\chi^2(4)=15,34$; $p=0,004$). Вполне возможно, что, в соответствии с нашей второй гипотезой, различия между контрольной и экспериментальной группами должны быть неодинаковыми при разных параметрах целевого стимула, однако полученные результаты не позволяют проверить данное предположение. Единственный вывод, который мы можем сделать на данном этапе – это вывод о том, что наличие предварительной установочной серии с ассимилятивным соотношением стимулов оказывает положительное влияние на различение целевого стимула.

2.4. Результаты: реформализация экспериментального дизайна

Выходом из сложившейся ситуации, на первый взгляд, будет рассмотрение зависимой переменной как количественной. Действительно, можно рассматривать в качестве случая подгруппу испытуемых, и в качестве показателя верного различения подсчитывать количество (или процент) испытуемых, верно оценивших соотношение диаметров окружностей в целевом стимуле. Тогда оказывается возможным применение традиционного двухфакторного ДА. Однако в нашем случае это не выход: зависимая переменная оказывается измеренной только на одном случае (одной подгруппе), а потому ее дисперсия не определена – ДА попросту невозможен! Чтобы данный математический метод оказался применимым, необходимо создать выборку подобных экспериментов, повторив их хотя бы 10 раз – т.е. увеличив общий объем экспериментальной выборки хотя бы до 1000 (!) испытуемых[1].

Другим возможным решением данной проблемы является рассмотрение одной из независимых переменных – различия в диаметрах окружностей в целевом стимуле – тоже как метрической. Тем не менее в данном случае независимые переменные оказываются измеренными в разных шкалах, и традиционные для отечественной психологии методы позволяют, как и в предыдущем случае, измерить лишь

[1] т.е. 10 подгрупп x 10 испытуемых x 5 разных целевых стимулов x 2 группы (экспериментальная и контрольная)

взаимосвязь пар переменных в отдельности. При этом мы теряем в информативности данных и получаем результаты с меньшим уровнем статистической значимости. Подобный анализ позволяет обнаружить сильную значимую корреляцию между величиной целевого стимула и верностью его восприятия (r=0,768; t(8)=3,388; p=0,01), однако влияние наличия предварительной установочной серии оказывается уже незначимым (t(8)=1,273; p=0,239). Таким образом, при использовании традиционных одномерных статистических методов подобная реформализация экспериментальных данных нецелесообразна. Вместе с тем как только мы переходим к ОРМ, ситуация меняется драматическим образом.

2.5. Использование ОРМ: результаты и обсуждение

Пусть *y* (*Y*) – количество испытуемых, верно решивших задачу на соотношение величин окружностей в целевом стимуле (зависимая переменная, количественная); *x* (фактор *X*) – разница в диаметре окружностей в целевом стимуле (независимая переменная, количественная); *k* (фактор *A*) – наличие предварительной установочной серии (независимая переменная, категориальная)[1]. Тогда наша задача – определить характер взаимосвязи в модели

$$Y = F(X, A), \tag{11}$$

т.е. методом наименьших квадратов найти функцию

$$y = f(x, k), \tag{12}$$

наиболее точно соответствующую экспериментальным данным.

---

[1] Здесь строчные буквы указывают на обозначение переменных в уравнениях (количественное описание, математическая модель), а заглавные (в скобках) указывают на обозначения факторов в описании дизайна модели (качественное описание, теоретическая модель). Традиционно при описании моделей экспериментов действие факторов, формализованных через категориальные переменные, обозначают преимущественно первыми буквами латинского алфавита (*A*, *B*, *C*...), а действие факторов, которым соответствуют количественные переменные – преимущественно через последние (*X*, *Y*, *Z*...)

На первом этапе необходимо определить, какой характер может принимать данная зависимость, что позволит нам определиться с возможными слагаемыми в модели (11) и в соответствующем ей уравнении (12). Иными словами, нам нужно набрать детали конструктора, из которого потом мы будем собирать модель.

Во-первых, сформулированные нами гипотезы предполагают возможность как независимого влияния каждой независимой переменной, так и взаимодействие факторов. Простейшая модель, отвечающая данным условиям – это модель вида $Y = A + X + AX$. Далее, во второй эмпирической гипотезе мы предположили, что влияние взаимодействия факторов AX нелинейно – т.е. влияние предварительной установочной серии уменьшается при уменьшении разницы в диаметрах окружностей в целевом стимуле. Наиболее простой и распространенный способ ввести подобную нелинейность в модель – это возведение переменной в степень. В нашем случае достаточно ввести только вторую степень, так как квадратичной зависимости будет достаточно, чтобы в целом (с некоторой точностью) описать подобную зависимость. Введение степеней более высокого порядка нецелесообразно: хотя в большинстве случаев это позволяет получить более точные модели, на выборках небольшого объема они становятся очень чувствительными к шуму, а потому могут быть неверно проинтерпретированы[1]. Таким образом, наша модель приобретает вид $Y = A + X + AX + AX^2$. Далее, у нас нет оснований считать, что зависимость Y от самого фактора X носит строго линейный характер, поэтому в модель следует добавить нелинейность и для него. Рассуждая аналогичным образом, приходим к окончательной формализации модели (11) в виде

---

[1] В целом, можно дать следующую рекомендацию: при построении регрессионных моделей следует стремиться к тому, чтобы количество наблюдений превышало количество коэффициентов в модели как минимум в 7-10 раз. В нашем случае это означает, что следует ограничиться рассмотрением моделей, включающих не более двух переменных. В противном случае повышается риск принять случайные особенности экспериментальной выборки за объективную закономерность.

$$Y = A + X + AX + X^2 + AX^2, \qquad (13)$$

которой соответствует уравнение

$$y = C + b_k + b_1 x + b_{2k} x + b_3 x^2 + b_{4k} x^2, k = 1..2. \qquad (14)$$

Несложно заметить, что уравнение (14) – это частный случай уравнения (10) с двумя «лишними» слагаемыми, описывающими дополнительную нелинейность модели.

Обратим внимание, что если в индексе коэффициента b стоит только число, то данный коэффициент связан с влиянием только фактора X, если в индексе стоит только k – то он связан только с влиянием фактора А, а если в индексе b-коэффициента стоит как число, так и индекс k, то такой коэффициент вносит вклад в объяснение взаимодействия двух факторов.

На втором этапе происходит поиск оптимальной модели, для чего могут использоваться методы пошагового включения, пошагового исключения и поиска лучших подмножеств.

Вывод об адекватности модели делается на основании следующих статистик:

1. статистическая значимость всей модели по F-критерию на заданном уровне;
2. статистическая значимость всех коэффициентов, входящих в модель, на заданном уровне;
3. относительно высокие показатели КМК и КМД.

Методы пошагового исключения, пошагового включения (с критическим значением для коэффициентов $p=0,1$), а также метод поиска наилучших подмножеств для моделей с разным числом переменных привели к одной и той же модели, включающей в себя только две переменные, соответствующие линейному независимому вкладу каждой переменной, т.е. в уравнении (14) статистически значимыми (на уровне $p<0,1$) оказались только коэффициенты b_k и b_1. Соответственно, итоговая математическая модель имеет вид

$$y = C + b_k + b_1 x, \ k = 1..2,$$

что соответствует теоретической модели

$$Y = A + X .$$

Для данной модели КМД=0,758, уточненный КМД=0,686, КМК=0,87, F=10,948, p=0,007. Статистические показатели для коэффициентов модели приведены в таблице 2, графическое представление модели показано на рис. 1.

Таблица 2.

Коэффициенты итоговой модели

Коэффициент	Наличие уст. серии	k	Значение коэффициента	β-коэффициент	t-критерий	p-уровень
C (константа)			2,897541		3,65745	0,008098
b_1			2,041418	0,767694	4,12676	0,004422
b_k	нет	1	-0,900000	-0,410365	-2,20592	0,063176
	есть	2	0,900000			

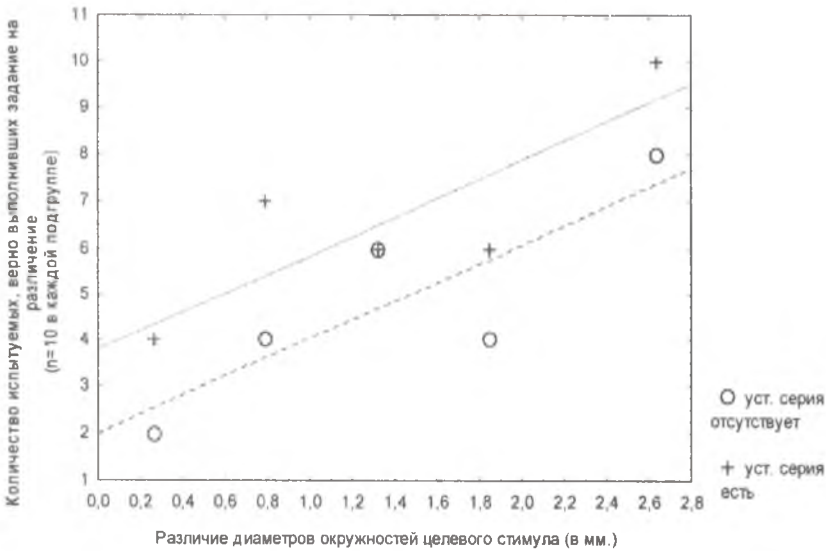

Рис. 1. Эмпирические результаты и графики регрессионных уравнений по результатам ОРМ

На первый взгляд, эти результаты всего лишь подтверждают полученные ранее. Тем не менее обратим внимание, что уровень статистической значимости для коэффициентов соответствует анализу исходных (см. п. 2.3 наст. статьи), а не агрегированных (см. п. 2.4) данных. Таким образом, на одних и тех же данных использование методов ОРМ позволяет получить результаты на более высоком уровне достоверности, т.е. эти методы являются более мощными. Фактически, при использовании методов ОРМ потеря информативности после агрегирования данных оказалась несущественной, в то время как для традиционных методов сравнения центральных тенденций и корреляционного анализа подобные действия снижают уровень статистической значимости результатов. Бо́льшая точность – это неспецифическое преимущество методов ОРМ.

Вместе с тем имеет смысл интерпретировать не только переменные, вошедшие в модель, но и те, которые были из нее исключены. Это означает, что, во-первых, взаимодействие факторов незначимо, а потому наша вторая гипотеза о различной силе влияния установочной серии в зависимости от величины различий в целевом стимуле может быть обоснованно отклонена. Во-вторых, исключение из модели нелинейных членов означает, что зависимость успешности решения задачи на различение визуальных стимулов от величины этих различий носит характер, приближенный к линейному.

В случае линейных моделей $\beta_k = r_{x_k y} r_{частн\ x_k y}$, и описывают уникальный вклад k-й переменной в объяснение дисперсии зависимой переменной, при этом выполняется условие $\sum_i \beta_i^2 = R^2$. Дополнительно учитывая свойства дисперсии и формулу (6), можно заключить, что:

1. построенная модель описывает 75,8% дисперсии зависимой переменной, которая наблюдается в полученных экспериментальных данных; при этом

2. на 58,9% разброс значений успешности решения задачи на различения целевого стимула зависит от величины различий в самом целевом стимуле и

3. на 16,8% этот разброс значений зависит от наличия предварительной установочной серии. При этом

4. 24,2% изменчивости зависимой переменной, наблюдаемые в эксперименте, остаются не объясненными.

Поскольку, как уже было отмечено, неуточненные значения КМД являются переоцененными, следует ожидать, что на уровне генеральной совокупности вклад каждого из факторов в объяснение зависимой переменой, равно как и общая доля дисперсии, объясненная всей моделью, будет несколько ниже.

Кроме того, модель может быть использована для построения прогнозов успешности решения аналогичных задач на основании полученных линейных уравнений: прогнозируемая успешность соответствует точкам, лежащим на регрессионных прямых на рис. 1.

Характер разброса данных на рис. 1 показывает, что точки, соответствующие эмпирическим результатам, в каждой из групп лежат вдоль некоторой кубической параболы, поэтому можно предположить, что введение в модель нелинейности в виде третьей степени позволит получить более точную модель. Заметим, что такие рассуждения отчасти правомерны: такая модель действительно существует, включает в себя все степени переменной x до третьей включительно, объясняет 94,2% общей дисперсии, при этом общий вклад переменной различий в целевом стимуле возрастает до 77,3% (невероятно высокий показатель!). Графически эта модель представлена на рис. 2.

Однако, как уже было отмечено выше, содержательной интерпретации такая модель не подлежит по причине большой подверженности влиянию шума (случайных характеристик) на ее параметры (вследствие поиска большого числа коэффициентов на небольшом массиве данных). Тем не менее, и в данной модели все переменные, связанные с взаимодействием факторов, оказались незначимыми, что дает дополнительную уверенность в выводе о независимом влиянии данных факторов.

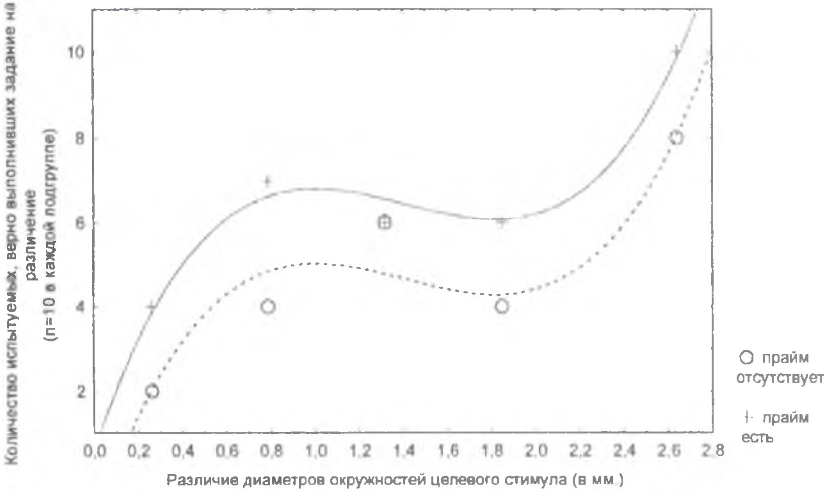

Рис. 2. Кубическая модель, построенная методами ОРМ на тех же экспериментальных данных

2.6. Общее обсуждение

Использование традиционных методов сравнения средних и корреляционного анализа позволило обосновать только один значимый результат: наличие предварительной установочной серии с ассимилятивным соотношением с целевым стимулом оказывает положительное влияние на эффективность различения разницы диаметров окружностей в целевом стимуле – факт, по большому счету, тривиальный. Использование методов ОРМ позволило обнаружить следующие эмпирические факты:

1. Влияние предварительной установочной серии с ассимилятивным соотношением с целевым стимулом на успешность решения задачи на визуальное различение величины объектов оказывается одинаково сильным независимо от объективно существующей разницы в величине сравниваемых объектов. Иными словами, такая зависимость носит характер, приближенный к линейной, однако есть основания предполагать, что существует некоторая нелинейность взаи-

мосвязи, которую можно будет обнаружить на выборках существенно большего объема[1].

2. В решении задач на визуальное различение величины объектов объективные различия в размерах сравниваемых стимулов оказывают примерно в три раза большее влияние на успешность решения такой задачи, чем наличие предварительной установочной серии.

Отметим, что для обнаружения этих двух фактов методы ОРМ (или их дальнейшее развитие в виде моделирования структурными уравнениями [*Митина*, 2008]) являются безальтернативными (т.е. другие методы не позволят их обнаружить), и в этом состоит их специфическое преимущество.

Сопоставив полученные результаты с работами Д.Н. Узнадзе (1966; 2001), можно сделать вывод, что предварительное предъявление ассимилятивной установочной серии будет приводить к эффекту восприятия по контрасту при предъявлении целевого стимула только в том случае, если целевой стимул является действительно неопределенным (в случае равных окружностей). Даже если стимул по своим параметрам остается близким к неопределенному, но на самом деле таковым не является, предварительная ассимилятивная установочная серия начинает оказывать положительный эффект и увеличивает число случаев верной оценки величины сравниваемых стимулов. А это означает, что человек способен определить и дифференцировать даже небольшие различия в размерах объектов, даже в том случае, если сам факт такого различения им не осознается, при этом сама установочная серия оказывает одинаковое влияние как на осознаваемые, так и неосознаваемые механизмы работы психики и сознания.

Более развернутая интерпретация может быть дана с учетом обширных работ по исследованию прайминга как явления имплицитной памяти. Прайминг-эффект определяется как «явление имплицитной памяти, которое представляет собой либо *(а)* изменение скорости или

---

[1] Вторая часть последнего предложения – это, конечно же, не эмпирический факт, а эмпирическая гипотеза для будущего исследования.

точности решения задачи (перцептивной, мыслительной или мнемической), наблюдаемое после предъявления информации, связанной с содержанием или с контекстом этой задачи, но не соотносящейся прямо с ее целью и требованиями, либо *(б)* повышение вероятности спонтанного воспроизведения этой информации в подходящих условиях» [*Фаликман*, *Койфман*, 2005]. В более широком понимании прайминг можно определить как «влияние стимула на последующие осознаваемые реакции испытуемого… Прайм-стимул…может осознаваться, а может предъявляться и на неосознаваемом уровне» [*Агафонов*, 2007, с. 140-141]. В соответствии с приведенными определениями проведенное исследование установки может интерпретироваться как исследование прайминг-эффекта, однако процедура проведенного экспериментального исследования не в полной мере соответствует экспериментальной парадигме исследования прайминг-эффекта, сложившейся в когнитивной психологии. Тем не менее, анализ литературы [*Агафонов*, 2003, 2007; *Аллахвердов*, 2003; *Величковский*, 2006] и полученных результатов позволяет предположить наличие некоторых общих механизмов, действие которых обнаруживается в экспериментах, направленных на изучение прайминга и влияния установки.

Более подробное обсуждение полученных результатов выходит за рамки настоящей статьи. Заинтересованный читатель может сам соотнести полученные результаты с анализом экспериментов Д.Н. Узнадзе и исследований прайминг-эффекта, обратившись к указанным выше работам, в которых дан прекрасный обзор исследований, посвященных исследованию имплицитной памяти и праймингу.

Литература

1. StatSoft, Inc. (2001). Электронный учебник по статистике. Москва, StatSoft. http://www.statsoft.ru/home/textbook/default.htm

2. *Агафонов А.Ю.* Когнитивная психомеханика сознания, или как сознание неосознанно принимает решение об осознании. – Самара: ИД «Бахрах-М», 2007.

3. *Агафонов А.Ю.* Основы смысловой теории сознания. – СПб.: Речь, 2003.

4. *Аллахвердов В.М.* Сознание как парадокс. – СПб.: Речь, 2003.

5. *Величковский Б.М.* Когнитивная наука: Основы психологии познания: в 2-х тт. Т.1. – М.: Смысл : Издательский центр «Академия», 2006.

6. *Гласс Дж., Стэнли Дж.* Статистические методы в педагогике и психологии. – М.: Прогресс, 1976

7. *Гусев А. Н.* Дисперсионный анализ в экспериментальной психологии. – М.: УМК «Психология», 2000.

8. *Капрара Дж., Сервон Д.* Психология личности. – СПб: Питер, 2003.

9. *Левин К.* Теория поля в социальных науках.: пер. с англ. – СПб: Сенсор; Речь, 2000.

10. *Митина О.В.* Моделирование латентных изменений с помощью структурных уравнений // Экспериментальная психология. – 2008. №1. – С. 131-148.

11. *Мышкис А.Д.* Элементы теории математических моделей. – М.: КомКнига, 2007.

12. *Наследов А.Д.* Математические методы психологического исследования. – СПб: Речь, 2004

13. *Самарский А.А., Михайлов А.П.* Математическое моделирование. Идеи. Методы. Примеры. – М.: Физматлит, 2001.

14. *Советов Б.Я., Яковлев С.А.* Моделирование систем: Учеб. для вузов. – М.: Высш. шк., 2001.

15. *Узнадзе Д.Н.* Психологические исследования. – М.: Наука, 1966.

16. Узнадзе Д.Н. Психология установки. - СПб.: Питер, 2001.

17. *Фаликман М.В., Койфман А.Я.* Виды прайминга в исследованиях восприятия и перцептивного внимания // Вестник Московского Университета, серия 14, Психология, №3,4. – М.: МГУ, 2005

18. *Филатов А.Ю.* Конспект лекций по эконометрике. – Иркутск: ИГУ. – 2006.