

В. ЛУШИН, Т.В. ТУЛУПЬЕВА, А.Л. ТУЛУПЬЕВ

ОТНОСИТЕЛЬНАЯ ВАЖНОСТЬ ПЕРЕМЕННЫХ-ПРЕДИКТОРОВ В АНАЛИЗЕ ДЕТЕРМИНАНТ РИСКОВАННОГО ПОВЕДЕНИЯ (КРАТКИЙ ОБЗОР)

Лушин В., Тулупьева Т.В., Тулупьев А.Л. **Относительная важность переменных-предикторов в анализе детерминант рискованного поведения (краткий обзор).**

Аннотация. Одной из центральных задач в исследованиях в области поведенческой эпидемиологии является выявление наиболее важных/весомых факторов, влияющих на изучаемое поведение. В статье предлагается краткий обзор рассмотренных J.W. Johnson'ом способов расчета оценки относительной важности переменных-предикторов. Рассматриваются примеры, связанные с выявлением детерминант рискованного поведения на основе оценок относительной важности.

Ключевые слова: относительная важность, относительный вес, регрессионный анализ, рискованное поведение.

Lushin V., Tulupyeva T.V., Tulupyev A.L. **Relative importance of predictor variables in analysis of risk behavior determinants (brief review).**

Abstract. One of the key analytic problems in behavioral epidemiology research is related to determining most important factors among those that influence the behavior of interest. The present paper offers a brief review of the method designed to estimate relative importance of predictor variables described by J. W. Johnson. Examples are discussed that illustrate the use of relative importance analysis to examine and compare impacts of determinants of sexual risk behavior.

Key words: relative importance, relative weight, regression analysis, risk behavior.

1. Введение. Одной из центральных задач в исследованиях в области поведенческой эпидемиологии является выявление наиболее важных/весомых факторов, влияющих на изучаемое поведение. Например, в [8] исследовали множественные когнитивные и аффективные предикторы сексуального поведения среди доминиканских подростков — группы населения с высоким уровнем риска заражения ВИЧ. Выявление ключевых детерминант данного поведения в исследуемом населении необходимо для создания эффективных превентивных программ с акцентом именно на ключевые детерминанты, максимально влияющие на поведение [1–3].

Похожая задача встаёт перед составителями инструментов-опросников для выявления предикторов рискованного поведения. Выявление наиболее важных предикторов предназначено для решений той же задачи — разработки эффективных превентивных программ, а также для разработки приемлемых по стоимости способов оценки,

насколько велик риск заражения, ассоциированный с поведением индивида. Примером соответствующего исследования является [5], в котором разработан опросник для раннего выявления риска заражения ВИЧем среди молодёжи в Западной Африке.

Для решения подобных задач необходим обоснованный статистический метод определения относительной важности (relative importance) или относительного веса (relative weight) предикторов в регрессионной модели.

Цель настоящей работы — описать, с учетом потребностей исследования детерминант рискованного поведения, некоторые подходы, рассмотренные и предложенные J.W. Jonson'ом [7], к диагностике регрессионных уравнений в части способов определения относительной важности переменных-предикторов.

2. Подходы к оценке относительной важности предикторов.

Достаточно легко оценить относительную важность предикторов в регрессионной модели, если они не коррелируют между собой. В таком случае относительная важность каждого предиктора может быть вычислена с помощью формулы [7, 10]:

$$w_i = \frac{b_i^2}{R^2},$$

где b_i — стандартизированный регрессионный коэффициент (standardized regression coefficient), R^2 — квадрат множественной корреляции (squared multiple correlation), w_i — показатель относительной важности -го предиктора, а i изменяется от 1 до n согласно индексации предикторов в регрессионной модели.

При этом сумма стандартизированных регрессионных компонентов составляет квадрат множественной корреляции (squared multiple correlation)

$$\sum_{i=1}^n b_i^2 = R^2,$$

что обеспечивает нормировку показателей относительной важности.

Однако в ряде социальных исследований, как, например, в работе [6], предикторы поведения в выборке населения коррелируют между собой. Например, экспектация (expectancy) – поведенческое убеждение (behavioral belief) в том, что секс принесёт удовольствие, коррелирует с аффектом радости при мысли о сексе, при том что оба предиктора отражают различные психологические процессы: аналитический и эмоциональный [6]. Интеркорреляция между предикторами

лишает вышеописанный способ сравнения детерминант поведения оснований.

Наиболее часто употребляемый, но, к сожалению, непригодный для решения обсуждаемой задачи, статистический метод определения сравнительной важности предикторов в социальных исследованиях, где интеркорреляция между предикторами неизбежна, — сравнение величины стандартизированных регрессионных коэффициентов (standardized regression coefficients) в множественной регрессии (multiple regression) [4, 13]. Многие исследователи (e.g. [9, 10]) критиковали этот метод, так как он 1) завышает показатель сравнительной важности предикторов высоко коррелируемых с результирующей (зависимой, прогнозируемой) переменной (dependent variable), 2) занижает показатели сравнительной важности всех остальных предикторов в модели, 3) нередко приводит к дискордантности знаков регрессионного коэффициента перед предиктором и корреляции этого предиктора с прогнозируемой переменной и 4) приводит к тому, что незначительные различия в выборках могут вызвать крупные различия в показателях важности предикторов в регрессионных моделях, построенных на основе таких выборок [9].

В последние годы всё больше исследователей, как, например, [5] или [8], используют альтернативный метод определения относительной важности детерминант: метод ортогональных компонент, предложенный Джонсоном [9, 10]. Этот метод предполагает переход от системы переменных-предикторов, которые могут иметь интеркорреляции, к ортогонализированной системе переменных (то есть, системе некоррелированных между собою переменных), причем близким (highly related) переменным оригинальной группы, но не коррелирующих между собой. В данном случае математически «близость» переменных оценивается с точки зрения метода наименьших квадратов: из всех возможных ортогонализаций исходной системы переменных выбирается такая, которая обеспечивает близость в указанном смысле. В [9] обсуждается, почему другие подходы к ортогонализации (метод главных компонент, метод Грамма–Шмидта) оказываются неудачными для использования в определении показателей относительной важности.

3. Описание метода ортогональных компонент, близких к исходным в смысле метода наименьших квадратов [9]. На первом этапе анализа необходимо принять, что имеется вектор значений критерия (значений зависимой переменной) y размерности $n \times 1$ и матри-

ца значений переменных-предикторов \mathbf{X} размерности $n \times p$; и все они выражены в стандартной форме (standard score form). Приняв, что \mathbf{X} имеет не меньше рядов, чем колонок, можно сформировать сингулярное разложение матрицы \mathbf{X} :

$$\mathbf{X} = \mathbf{P}\mathbf{\Delta}\mathbf{Q} \quad (1)$$

где \mathbf{P} содержит собственные векторы $\mathbf{X}\mathbf{X}'$, \mathbf{Q} содержит эйгенвекторы $\mathbf{X}'\mathbf{X}$, а $\mathbf{\Delta}$ является диагональной матрицей, содержащей сингулярные значения \mathbf{X} . Сингулярные значения являются квадратными корнями собственных значений матриц $\mathbf{X}\mathbf{X}'$ и $\mathbf{X}'\mathbf{X}$.

Диагональные значения в матрице $\mathbf{\Delta}$ определяются однозначно с точностью до знака; договоримся считать их положительными. При этом \mathbf{P} и \mathbf{Q} определяются однозначно, если ни один из диагональных элементов $\mathbf{\Delta}$ не равен нулю [11]. Если ни одна пара переменных-предикторов в матрице \mathbf{X} не связана линейной зависимостью (то есть эти переменные-предикторы не имеют коэффициента корреляции с абсолютным значением, равным единице), \mathbf{X} будет иметь максимальный ранг, а, значит, ни один из диагональных элементов $\mathbf{\Delta}$ не будет равен нулю. Это условие необходимо в дальнейшем описании.

В [11] показано, что ближайшая (best-fitting) ортогональная аппроксимация (orthogonal approximation) матрицы \mathbf{X} может быть получена согласно уравнению

$$\mathbf{Z} = \mathbf{P}\mathbf{Q}' \quad (2)$$

Колонки матрицы \mathbf{Z} представляют собой ближайшую (best-fitting) аппроксимацию колонок матрикса \mathbf{X} в смысле минимизации суммы квадратов разностей (residual sum of squares) между оригинальными переменными и ортогональными переменными [11]. Вектор коэффициентов корреляции $\mathbf{\beta}$ в уравнении регрессии \mathbf{y} относительно новой (построенной ортогональной) системы переменных \mathbf{Z} , исчисляется по формуле:

$$\mathbf{\beta} = \mathbf{Q}\mathbf{P}'\mathbf{y} \quad (3)$$

Возведенные в квадрат элементы β представляют пропорцию дисперсии в y , объяснимую с помощью (дисперсии) Z , и рассматриваются как аппроксимация относительной важности исходных переменных-предикторов, реализации которых представлены в колонках матрицы X .

Однако ортогонализированные переменные («реализации» которых заданы матрицей Z) — лишь приблизительно отражают исходные переменные; они могут не оказаться «содержательно» достаточно близкими отображениями исходных переменных в случае, если последние имеют большие по абсолютному значению корреляции между собой.

Осознавая ограничения вышеописанной процедуры, авторы работы [7] предложили метод, с помощью которого ортогональные переменные могут быть приведены в соответствие с оригинальными переменными так, чтобы достичь «большого качества приближения» оценки относительной важности оригинальных переменных. Для этого сперва строятся уравнения регрессии переменных («реализации» которых представлены векторами-колонками) Z от X :

$$\Gamma = Q\Delta^{-1}Q'. \quad (4)$$

где Γ — матрица регрессионных весомостей Z от X . Отметим, что Z получается линейным преобразованием (умножением на матрицу) из матрицы X : простая линейная трансформация матрицы X : $Z = X\Gamma$ при этом [7] относительный вклад каждой исходной переменной в ортогональную -ю переменную определяется по формуле:

$$\gamma_{jk}^{*2} = \frac{\gamma_{jk}^2}{\sum_{i=1}^p \gamma_{jk}^2}, \quad (5)$$

причём

$$\Gamma = (\gamma_{jk}).$$

В результате, относительная важность каждой исходной переменной определяется по формуле:

$$\delta_j^2 = \sum_{k=1}^p \gamma_{jk}^{*2} \beta_k^2, \quad (6)$$

причём

$$R^2 = \sum_{j=1}^n \delta_j^2.$$

Однако, как отмечается в [9], хотя предложенный выше подход к оценке показателей относительной важности неплох с точки зрения вычислительной сложности, он также несовершенен в первую очередь из-за того, что, хотя исходные переменные и ортогонализированные переменные ассоциированы друг с другом и выдвинуто требование их «близости», на самом деле различия в парах исходная переменная – ортогонализированная переменная могут быть весьма существенными, и, таким образом, оценка, базирующаяся на ортогонализированной переменной, может неадекватно отражать вклад (относительную важность) исходной переменной.

Основной результат [9] состоит как раз в новом способе оценить показатель относительной важности так, чтобы избежать указанного выше недостатка. В отличие от предыдущего подхода, в [9] предлагается строить уравнения регрессии) \mathbf{X} от \mathbf{Z} , причем

$$\mathbf{\Lambda} = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X}, \quad (7)$$

где для простоты

$$\mathbf{\Lambda} = (\lambda_{jk}) = \mathbf{\Gamma}^{-1}.$$

После ввода дополнительных обозначений (не путать со степенью Кронекера!)

$$\begin{aligned} \mathbf{\Lambda}^{[2]} &= (\lambda_{jk}^2), \\ \mathbf{\beta}^{[2]} &= (\beta_k^2), \end{aligned}$$

предлагаемый в [9] вектор показателей относительной важности запишется как

$$\boldsymbol{\varepsilon} = \mathbf{\Lambda}^{[2]}\mathbf{\beta}^{[2]}.$$

В [9] утверждается, что вычислительная сложность метода остаётся приемлемой, и, при этом, он свободен от ранее обсуждавшихся недостатков — этот аргумент был поддержан серией вычислительных экспериментов.

4. Прикладная польза метода. В результате сравнения относительной важности предикторов методом ортогональных компонентов исследователям удаётся получить достаточно аргументированные

оценки того, какие же именно факторы наиболее весомо влияют на поведение в изучаемой группе населения. Например, авторы [8], проведя такой анализ с помощью компьютерной программы MPlus, обнаружили, что из десятков исследованных когнитивных и аффективных факторов лишь 7 позитивных и негативных факторов оказали ощутимое влияние (выше 10% относительной важности) на намерение 11–14-летних доминиканских подростков заниматься сексом в обозримом будущем. Среди этих ключевых факторов оказались экспектации (expectancies) удовольствия в результате секса, увеличения собственной популярности в группе общения, увеличения восприятия собственной привлекательности, увеличения межличностной близости с партнёром и стыда в случае беременности; а также аффекты радости и испуга/нервозности при мысли о будущем сексе [8].

Данные, полученные Гильямо-Рамосом и коллегами [8], позволили сделать ряд важных выводов в отношении психологических механизмов сексуального поведения среди доминиканских подростков, расширяющих данную сферу знаний и за пределы этой группы населения. Сравнительная важность переменных-предикторов, рассчитанная на основе метода ортогональных компонентов, свидетельствует, что склонность к сексуальному риску среди доминиканской молодёжи определяется, в основном, не когнитивно-аналитическими процессами, а эвристическими процессами, относящимися к предвкушению удовольствия и увеличения собственного статуса в группе общения, а также эмоциональной окраской мыслей о сексе. Неожиданной находкой стало отсутствие экспектаций заражения ВИЧем и другими ИППП среди ключевых предикторов сексуальных намерений доминиканских подростков.

Как свидетельствуют данные, полученные Гильямо-Рамосом и коллегами [8], мысли о возможном заражении не оказывают существенного влияния на выбор сексуальной стратегии среди данной категории населения; а ведь именно акцентирование этих последствий является центральной частью превентивных мероприятий, проводимых, в том числе, и с расчётом на доминиканских подростков [8]. Исследование Гильямо-Рамоса и коллег [8] также показало, что связь между мыслительно-аффективными процессами и рискованным поведением у подростков зависит в большей степени от контекстуальной экспозиции в период перехода от детства к юности, нежели от культурно-этнической принадлежности.

Другой пример применения метода Джонсона представлен в работе [5], посвящённой составлению культурно-приспособленного ин-

струмента-опросника для раннего выявления рискованных сексуальных поведенческих стратегий среди молодёжи в Бамако, в Западной Африке. Предложив 189 испытуемым (попавшим в выборку) батарею вопросов, касающихся теоретически обоснованных детерминант сексуально рискованных поведений (преимущественно, секса без кондомов), исследователи составили регрессионные модели, отражающие влияние данных детерминант на каждый из поведенческих результатов, а затем, пользуясь методом Джонсона, преобразовали полученные регрессионные коэффициенты в ортогональные компоненты, очищенные от интеркорреляции. Результирующие ортогональные компоненты позволили выявить наиболее важные/весомые предикторы сексуально-рискованного поведения в данной категории населения [5]. Именно эти ключевые детерминанты легли в основу инструмента-опросника, адекватно приспособленного к раннему выявлению риска заражения ВИЧ среди изучаемого населения, а значит именно они обеспечивают своевременное и научно обоснованное превентивное вмешательство.

Важно отметить, что предикторы определённых видов поведения значительно различаются среди разных этнических, культурных, географических, и прочих групп населения, отличающихся по типам социо-экологических контекстов, формирующих поведенческие нормы, привычки, убеждения, и другие социогенные детерминанты [8]. Поэтому для изучения механизмов, формирующих поведение людей в разных контекстах, крайне важна опора на обоснованные статистические методы, позволяющие определить наиболее весомые детерминанты поведенческих стратегий для каждой конкретной группы населения, объединённой общими контекстуальными экспозициями. Метод ортогональных компонентов Джонсона представляется достойным методом выбора для решения такого рода исследовательских задач, особенно если учитывать необходимость кросс-культурного компаративизма как обязательной компоненты исследований угрозообразующего, рискованного или просто социально-значимого поведения.

Одна из самых важных характеристик рискованного поведения — его интенсивность. Как правило, чем меньше интенсивность рискованного поведения, тем ниже риск заражения ВИЧ-инфекцией или иной ИППП. Превентивные вмешательства, в конце концов, должны быть нацелены на снижение интенсивности рискованного поведения; кроме того, эффективность превентивных программ тоже должна оцениваться на основе того, насколько целевая группа снизила свое участие в рискованном поведении (разумеется, с учетом ряда других параметров) [3].

С указанной точки зрения, ожидается, что продуктивным будет совместное использование описанного в статье способа оценки относительной важности переменных-предикторов, на основе которых предсказывается интенсивность, которая измеряется с помощью подхода, описанного в [1, 2], где предлагается строить статистические оценки интенсивности рискованного поведения на основе ответов респондентов о последних эпизодах их рискованного поведения. Следует отметить, что в [1, 2] поведение моделируется с помощью гамма-пуассоновского случайного процесса; однако в зависимости от видов риска, особенностей групп и иных факторов может потребоваться использовать и иные статистические модели.

Литература

1. *Зельтерман Д., Тулупьев А.Л., Суворова А.В., Пащенко А.Е., Мусина В.Ф., Тулупьева Т.В., Красносельских Т.В., Гро Л., Хаймер Р.* Обработка систематической ошибки, связанной с длиной временных интервалов между интервью и последним эпизодом в гамма-пуассоновской модели поведения // Труды СПИИРАН. 2011. Вып. 16. С. 160–185.
2. *Зельтерман Д., Суворова А.В., Пащенко А.Е., Мусина В.Ф., Тулупьев А.Л., Тулупьева Т.В., Гро Л.Е., Хаймер Р.* Диагностика регрессионных уравнений в анализе интенсивности рискованного поведения по его последним эпизодам // Труды СПИИРАН. 2011. Вып. 17. С. 33–46.
3. *Пащенко А.Е., Тулупьев А.Л., Тулупьева Т.В., Красносельских Т.В., Соколовский Е.В.* Косвенная оценка вероятности заражения ВИЧ-инфекцией на основе данных о последних эпизодах рискованного поведения // Здравоохранение Российской Федерации. 2010. № 2. С. 32–35.
4. *Afifi A.A., Clarke V.* Computer aided multivariate analysis (2nd ed.). New York: Van Nostrand Reinhold, 1990. 480 p.
5. *Boileau C., Rashed S., Sylla M., Zunzunegui M.V.* Monitoring HIVs risk and evaluating interventions among young people in urban West Africa: Development and validation of an instrument // *AIDS Education and Prevention*. 2008. Vol. 20(3). P. 03–19.
6. *Gibson W.A.* Orthogonal predictors: A possible resolution of the Hoffman–Ward controversy // *Psychological reports*. 1962. No. 11. P. 32–34.
7. *Green P.E., Carroll J.D., DeSarbo W.S.* A new measure of predictor variable importance in multiple regression. *Journal of Marketing Research*. 1978. Vol. 15. P. 356–360.
8. *Guilamo-Ramos V., Jaccard J., Lushin V., Gonzalez B., Martinez B.* (2012). Emotions and Cognitions as Predictors of Early Adolescent Sexual Behavior in the Dominican Republic // *AIDS And Behavior* (under review).
9. *Johnson J.* A heuristic method for estimating the relative weight of predictor variables in multiple regression // *Multivariate Behavioral Research*. 2000. Vol. 35. P. 1–19.
10. *Johnson J., Lebreton J.* History and use of relative importance indices in organizational research // *Organizational Research Methods*. 2004. Vol. 7. P. 238–257.
11. *Johnson R.M.* The minimal transformation to orthogonality. *Psychometrica*. 1966. No. 31. P. 61–66.

12. *Neter J., Kutner M.H., Nachtsheim C.J., Wasserman W.* Applied linear statistical models. 4th ed. Boston, MS: McGraw-Hill, 1996. 1408 p.
13. *Walsh A.* Statistics for the social sciences. New York: Harper & Row, 1990. 305 p.

Лушин Виктор — MD, Ph.D. Candidate, научный сотрудник Центра здоровья и семейных проблем юношества латино-американского происхождения (CLAFH) Нью-Йоркского университета (NYU). Область научных интересов: поведенческие риски среди юношества, связанные с алкоголем, наркотиками и заражением ВИЧ/СПИД; внутри-семейные взаимоотношения подростков; прикладная психология развития. Число научных публикаций: 4. vbl206@nyu.edu; 1 Washington Square North, New York, NY, 10003, USA; ph.: 1 917-517-5082.

Viktor Lushin – MD, Ph.D. Candidate, Research Scientist at NYU Center for Latino Adolescent Family Health. Scientific Interests: risk behaviors among adolescents related to alcohol, drugs' use and HIV/AIDS transmission; adolescent family relationships; applied/developmental psychology. Number of scientific publications: 4. vbl206@nyu.edu; 1 Washington Square North, New York, NY, 10003, USA; ph.: 1 917-517-5082.

Тулупьева Татьяна Валентиновна — канд. психол. наук, доцент; старший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики Учреждения Российской академии наук Санкт-Петербургского института информатики и автоматизации РАН (СПИИРАН), доцент кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ), доцент кафедры психологии управления и педагогики Северо-Западной академии государственной службы (СЗАГС). Область научных интересов: применение методов математики и информатики в гуманитарных исследованиях, информатизация организации и проведения психологических исследований, применение методов биостатистики в эпидемиологии, психология личности, психология управления. Число научных публикаций — 75. TVT@ias.spb.su, www.tulupyev.spb.ru; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450.

Tulupyeva Tatiana Valentinovna — PhD in Psychology, associate professor; senior researcher, Laboratory of Theoretical and Interdisciplinary Computer Science, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), associate professor, Computer Science Department, Faculty of Mathematics and Mechanics, St. Petersburg State University (SPbSU), associate professor, Management Psychology and Pedagogic Department, North-West Academy of Public Administration (NWAPA). Research interests: application of mathematics and computer science in humanities, informatization of psychological studies, application of biostatistics in epidemiology, psychology of personality, management psychology. Number of publications — 75. TVT@ias.spb.su, www.tulupyev.spb.ru; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Тулупьев Александр Львович — д.ф.-м.н., доцент; заведующий лабораторией теоретических и междисциплинарных проблем информатики Учреждения Российской академии наук Санкт-Петербургского института информатики и автоматизации РАН (СПИИРАН), доцент кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ). Область научных интересов: представление и обработка данных и знаний с неопределенностью, применение методов математики и информатики в социокультурных исследованиях, применение методов биостатистики и математического моделирования в эпидемиологии, технология разработки программных комплексов с СУБД. Число научных публикаций — 250.

ALT@iias.spb.su, www.tulupyev.spb.ru; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450.

Tulupyev Alexander Lvovich — PhD in Computer Science, Dr. of Sc., Associate Professor; Head of Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS, Professor of Computer Science Department, SPbSU. Research area: uncertain data and knowledge representation and processing, mathematics and computer science applications in socio-cultural studies, biostatistics, simulation, and mathematical modeling applications in epidemiology, data intensive software systems development technology. Number of publications — 250. ALT@iias.spb.su, www.tulupyev.spb.ru; SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Рекомендовано ТиМПИ СПИИРАН, зав. лаб. А.Л. Тулупьев, д.ф.-м.н., доцент.
Статья поступила в редакцию 25.11.2011.

РЕФЕРАТ

Лушин В., Тулупьева Т.В., Тулупьев А.Л. **Относительная важность переменных-предикторов в анализе детерминант рискованного поведения (краткий обзор).**

Одной из центральных задач в исследованиях в области поведенческой эпидемиологии является выявление наиболее важных/весомых факторов, влияющих на изучаемое поведение. В качестве примера приведено исследование множественных когнитивных и аффективных предикторов сексуального поведения среди доминиканских подростков — группы населения с высоким уровнем риска заражения ВИЧ. Выявление ключевых детерминант данного поведения в исследуемом населении необходимо для создания эффективных превентивных программ с акцентом именно на ключевые детерминанты, максимально влияющие на поведение [1–3].

Похожая задача встаёт перед составителями инструментов-опросников для выявления предикторов рискованного поведения. Выявление наиболее важных предикторов предназначено для решений той же задачи — разработки эффективных превентивных программ, а также для разработки приемлемых по стоимости способов оценить, насколько велик риск заражения, ассоциированный с поведением индивида. Приведен пример исследования, в котором разработан опросник для раннего выявления риска заражения ВИЧ среди молодежи в Западной Африке.

Для решения подобных задач необходим обоснованный статистический метод определения относительной важности (relative importance) или относительного веса (relative weight) предикторов в регрессионной модели.

В работе дано описание, ориентированное на потребности исследования детерминант рискованного поведения, подходов, рассмотренных и предложенных J.W. Jonson'ом, к диагностике регрессионных уравнений в части способов определения относительной важности переменных-предикторов.

Отмечено, что предикторы определённых видов поведения значительно различаются среди разных этнических, культурных, географических и прочих групп населения, отличающихся по типам социо-экологических контекстов, формирующих поведенческие нормы, привычки, убеждения и другие социогенные детерминанты. Поэтому для изучения механизмов, формирующих поведение людей в разных контекстах, крайне важна опора на обоснованные статистические методы, позволяющие определить наиболее весомые детерминанты поведенческих стратегий для каждой конкретной группы населения, объединённой общими контекстуальными экспозициями. Метод ортогональных компонентов Джонсона представляется достойным методом выбора для решения такого рода исследовательских задач, особенно если учитывать необходимость кросс-культурного компаративизма как обязательной компоненты исследований угрозообразующего, рискованного или просто социально-значимого поведения.

SUMMARY

Lushin V., Tulupyeva T. V., Tulupyev A.L. **Relative importance of predictor variables in analysis of risk behavior determinants (brief review).**

Some of the key analytic problems in behavioral epidemiology research are related to determining most important factors among those that influence the behavior of interest in the examined population. Examples cited in this paper include a study by Guilamo-Ramos et al. examining multiple cognitive and affective predictors of sexual behavior among Dominican adolescents, a population disproportionately vulnerable to HIV. Finding key determinants of sexual risk behavior in this population by the use of relative importance/relative weight methodology provides crucial evidence to inform tailored preventive programs focusing on the most important predictors of the targeted behavior.

A similar analytic task is addressed by Boileau et al. in designing a surveillance instrument for early detection of HIV-related sexual risk behavior among adolescents in West Africa. Relative weights of theoretically plausible predictors of HIV transmission are examined in the population of interest to determine the key predictors to be used in a tailored surveillance instrument.

These and similar applied tasks require a trustworthy statistical method of determining relative importance/relative weights of predictors in regression models. The present paper describes approaches offered by J. W. Johnson to the diagnostics of regression equations in regard to the methods of determining relative importance of predictor variables.

The present description is oriented towards addressing the analytic needs of investigators examining determinants of sexual risk behaviors.

It is noted that sets of key predictors of certain behaviors vary across populations embedded in different ethnic, cultural, geographical, economic, and otherwise dissimilar ecological contexts that differently shape social norms, beliefs, habits and other sociogenic determinants of behavior. Therefore it is crucial for the researchers to examine mechanisms shaping behavior in various contexts to rely on trustworthy statistical methodologies enabling them to identify most important behavioral determinants for each population group characterized by common contextual exposures. Johnson's method of orthogonal components appears to be a worthy method of choice for addressing such intellectual problems, especially in the view of cross-cultural comparativism as a necessary component of research in risky or otherwise socially consequential behaviors.