

## **Включение анализа данных в сферу математики: необходимость и последствия**

**// *Философия математики. Актуальные проблемы. Тез. II междунар. Научной конф.***

**М.: Макс пресс, 2009. С. 257-259**

**Толстова Юлиана Николаевна, *д.с.н., профессор***

*Государственный университет – Высшая школа экономики*

**Включение анализа данных в сферу математики: необходимость и последствия  
(может ли математика содержать элементы постнеклассичности?)**

Исторически сложилось так, что анализ данных (АД) (рассматриваемый нами как ветвь науки; о его специфике будем говорить на примере социологии) обычно воспринимается математическим истеблишментом как нечто для него чуждое. С одной стороны, это оправдано. АД явно не удовлетворяет тем критериям строгости, которые считаются само собой разумеющимися для современной математики. Он включает в себя массу эвристических алгоритмов, для многих методов не имеет средств переноса результатов с выборки на генеральную совокупность, требует постоянного диалога исследователя с компьютером. Но, с другой стороны, существуют аргументы в пользу того, что АД должен развиваться «под крылом» математики.

1. Алгоритмы, входящие в АД, зачастую довольно серьезно «начинены» математикой: профессиональное математическое образование требовалось для их создания, оно же нужно и для их квалифицированного использования. И то, что математики как бы отторгают от себя все, что связано с АД, приводит к целому ряду дефектов и в развитии наук, стимулирующих развитие АД (например, социологии), и в соответствующих образовательных курсах. Это проявляется, во-первых, в том, что математики, желающие развивать АД, зачастую «двигают» его не в ту сторону, которую диктует практика; математики, преподающие АД нематематикам, не пользуются успехом у слушателей; во-вторых, напротив, исследователи-нематематики просто в силу недостатка математического образования зачастую некорректно применяют методы, не могут достаточно профессионально преподавать АД, эффективно развивать методный арсенал.

2. Можно сделать предположение, что невнимание к запросам практики негативно сказывается и на развитии самой математики: потенциальные возможности для развития новых направлений игнорируются. Математики перестают быть математиками в широком смысле

слова, замыкаются в своей «башне из слоновой кости». Вспомним деление графом Витте (русским министром финансов и Председателем совета министров в конце XIX – начале XX, по образованию математиком) всех математиков на математиков-философов и математиков-исчислителей [1, с.78]. В математическом мире процветают последние и не хватает первых (мысли Витте продолжает Арнольд [2]).

Основным источником нестрогости алгоритмов служат трудности формализации тех явлений, для изучения которых разрабатывается метод. Однако если бы математики всегда полагали, что попытки изучения, казалось бы, не поддающихся формализации явлений – не их удел (т.е. если бы среди математиков не было бы математиков-философов), то никогда не родилась бы, например, теория вероятностей. Но Паскаль и Ферма были математиками-философами, и математика не потеряла одну из своих замечательных ветвей. И хотя с середины XVII века, когда в переписке этих ученых родилась теория вероятностей, почти до середины XX века было очень много споров о том, стоит ли теорию вероятностей считать вполне математической дисциплиной [3], в конце концов теория вероятностей завоевала всеми признанное место в математике (конечно, потому, что существовали и такие математики-философы, как Колмогоров). Вспомним также, например, отношение Стеклова к творчеству Лузина (Лузин, говоря о пока не доказанных, но интуитивно ему ясных утверждениях писал: «мне кажется», а Стеклов отмечал на полях: «ему кажется, а мне не кажется» [4]). Рискнем провести аналогию между упомянутыми фактами и допущением тех неопределённостей, которые несет с собой АД. Допуск же в математику нестрогостей, связанных с «ухватыванием» исследователем неформализованных (до поры, до времени) аспектов реальности можно расценивать как признание того, что исследователь есть часть познаваемой им системы, что является одной из фундаментальных черт постнеклассической науки.

Потенциальных «точек роста» математики применительно к тем «толчкам», которые может дать для развития этой науки современная социология, довольно много. Например, развитие статистики объектов нечисловой природы, изучение устойчивости огромного количества методов относительно нарушений условий их применимости (и тем, и другим занимаются, но мало по сравнению с запросами практики); изучение частично упорядоченных множеств, решеток (имеются важные для социологии разработки, связанные с использованием решеток для анализа понятий, но это – только одно из возможных направлений; скажем, еще в 50-х годах XX века известный американский ученый Кумбс призывал создать аналог факторного анализа для частично упорядоченных множеств, но этого пока нет).

Особое внимание обратим на необходимость изучения ситуации, когда понятие переменной перестает играть содержательную роль, переменная из реальной становится номинальной (логический номинализм-реализм), как бы распадается на отдельные альтернативы, а социолога интересуют их сочетания, детерминирующие определенное поведение респондента. Эти сочетания получили название взаимодействий (понятие близко к понятию взаимодействия в классической математической статистике). В свое время автор предложила обобщить это понятие [5]. Методов поиска (обобщенных) взаимодействий существует много. Часть из них имеет довольно строгий характер, часть – типичные «неточные» методы АД. Стоит задача объединения этих методов в единый комплекс, аккумулирующий в себе достоинства каждого метода и, напротив, приглушающий соответствующие недостатки.

Подчеркнем, что в описанной ситуации речь идет о довольно кардинальном изменении привычного вида исходных данных, о переходе к другой парадигме: объекты задаются не как точки признакового пространства, а как наборы дискретных свойств (или логических функций от таких свойств, что предполагается принципами обобщения понятия взаимодействия), не обязательно являющихся значениями принимаемых во внимание признаков. Такой вид первичных моделей изучаемых объектов отвечает человеческому мышлению. Поэтому методы поиска обобщенных взаимодействий относят к интеллектуальному АД, претендующему на моделирование работы человеческого интеллекта.

Описание объектов с помощью (обобщенных) взаимодействий можно расценивать как своеобразный подход к пониманию нелинейностей, фундаментальный характер которых считается общепринятым в постнеклассической науке. Это – второй элемент постнеклассичности, который, на наш взгляд, должна вобрать в себя математика, по крайней мере, с точки зрения потребностей моделирования социальных явлений.

### **Литература.**

1. Витте С.Ю. Воспоминания. Т.1. М.: Изд-во социально-экономической литературы, 1960
2. Арнольд В.И. «Жесткие» и «мягкие» математические модели // Математическое моделирование социальных процессов. Вып. 1. М.: Социологический факультет МГУ, 1998. С.29-51
3. Крамер Х. Полвека с теорией вероятностей: наброски воспоминаний // современные проблемы математики. М.: Знание, 1979

4. М.А.Лаврентьев. Николай Николаевич Лузин // Успехи математических наук, 1974. Т. XXIX. Вып. 5 (179). С. 177-182
5. Толстова Ю.Н. Анализ социологических данных. М.: Научный мир, 2000