

## ВЛИЯНИЕ СОЦИАЛЬНЫХ ФАКТОРОВ НА СТАТИСТИКУ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ: БАНКИ КРИСТАЛЛИЧЕСКИХ СТРУКТУР

Словохотов Ю.Л.

Химический факультет МГУ, Институт элементоорганических соединений РАН  
119991, Москва, Воробьевы горы, 1, стр. 3, химический факультет МГУ  
slov@phys.chem.msu.ru

Хотя численные параметры и математические модели используют во многих науках об обществе, степень объективности и воспроизводимость *измерений* для социальных систем остаются не вполне ясными. Из-за многообразия разнородных факторов, влияющих на состояние социума, возможность разделить результаты измерений на содержательный *сигнал* и случайный, поддающийся минимизации *шум* в общественных дисциплинах неочевидна и требует отдельного анализа. Ключ к решению этой проблемы могут дать большие выборки объективно установленных физических данных, сформированные под влиянием социальных факторов.

Нами проанализированы распределения для выборок количественных параметров кристаллов (геометрия типичных фрагментов, симметрия), достоверно установленных для 900 тыс. индивидуальных кристаллических веществ. Большинство этих веществ было получено искусственно в ответ на определенные запросы человеческого общества и в ходе взаимодействия исследователей. Поэтому выборки, извлекаемые из компьютерных *банков структурных данных*, подвержены сильному влиянию "человеческих" факторов, а их статистика косвенно отражает экономические и социальные процессы.

В докладе обсуждаются типичные распределения количественных данных для "неживых" физических и для "живых" социальных систем. Показано, что крупные выборки (сотни и тысячи точек) результатов структурных исследований похожи на массивы "социальных" данных (негауссовы распределения с протяженными "хвостами", кластеризация точек). Это отражает такие физические свойства социума, как микро- или мезоскопический характер систем, нестационарность случайных процессов, корреляции последовательных событий. Подобные выборки, несмотря на точность и достоверность составляющих их чисел, часто не поддаются рутинным методам обработки результатов измерений. В эмпирических статистических исследованиях их обычно считают "плохими" и подвергают произвольному редактированию, что искажает содержащуюся в них информацию.