63-й Всемирный статистический конгресс Международного статистического института: новые тренды в мировой статистике

Елена Викторовна Зарова

ГБУ «Аналитический центр» Правительства Москвы; Российский экономический университет имени Г.В. Плеханова, г. Москва, Россия

doi: <https://doi.org/10.34023/2313-6383-2021-28-04-121-127>.

Для цитирования: Зарова Е.В. 63-й Всемирный статистический конгресс Международного статистического института: новые тренды в мировой статистике. Вопросы статистики. 2021;28(4):121–127.

63rd ISI World Statistical Congress: New Trends in Global Statistics

Elena V. Zarova

Analytical Center by Moscow City Government; Plekhanov Russian University of Economics, Moscow, Russia

doi: <https://doi.org/10.34023/2313-6383-2021-28-04-121-127>.

For citation: Zarova E. V. 63rd ISI World Statistical Congress: New Trends in Global Statistics. *Voprosy Statistiki*. 2021;28(4):121–127. (In Russ.)

Всемирные статистические конгрессы, организуемые Международным статистическим институтом (МСИ), проводятся один раз в два года в разных странах. Всемирный статистический конгресс и связанные с ними события являются кульминацией деятельности МСИ, поскольку программа конгресса объединяет мероприятия ассоциаций, советов и комитетов, членов МСИ, выступления ученых и практиков в области статистики из разных стран. Совсем недавно, с 11 по 16 июля 2021 г., состоялся очередной 63-й Всемирный статистический конгресс МСИ. Впервые он проводился в виртуальном формате из-за ситуации, связанной с пандемией коронавируса.

Каждый конгресс имеет разнообразную научную программу, которая включает десятки тематических сессий, круглых столов, докладов в формате постеров, лекций и мастер-классов. Систематизация и анализ тематики Все-

мирного статистического конгресса МСИ необходимы для получения информации об актуальных трендах развития мировой статистической мысли и наиболее важных, представляющих всеобщий интерес прикладных статистических задачах и методах их решения. Большое значение для развития российской статистики имеет мировое признание работ отечественных авторов, что подтверждается в том числе их включением в программу Всемирного статистического конгресса МСИ. Тематический анализ сессий конгресса позволяет определить место российских ученых в общей новаторской статистической деятельности. В статье приводятся результаты тематического анализа сессий конгресса с авторским переводом и ссылками, содержащими разъяснение ряда терминов и указывающими на сопутствующие литературные источники.

Ключевые тематические направления 63-го Всемирного статистического конгресса МСИ: результаты текстового частотного и контентного анализа

Работа Конгресса началась 11 июля с традиционной церемонии открытия и онлайн-приема. С приветственным словом к участникам конгресса выступил президент МСИ профессор Джон Бейлер (John Bailer). Также был представлен ролик с видеообращениями статистиков из разных стран и разного возраста, которые связали свои пожелания конгрессу с ответами на заранее заданные вопросы: «Сколько лет вы занимаетесь статистикой?» и «В каком Всемирном статистическом конгрессе МСИ вы принимали участие в первый раз?». В качестве специального гостя на церемонию открытия была приглашена профессор биостатистики из США, лауреат Международной премии по статистике 2021 г. Нэн Лэрд (Nan Laird)¹ — в знак признания ее работы, посвященной эффективным методам сложных лонгитюдных² исследований.

В рамках конгресса прошли заседания 158 тематических сессий (Invited Paper Sessions — IPS) со средним числом спикеров пять человек. Кроме того, в формате устных выступлений (oral presentations), а также постерных докладов (CPS) было представлено еще более 60 докладов.

Программа конгресса отражает основные наиболее актуальные направления мировой статистической мысли, а также появившиеся недавно новые сферы научного поиска в области статистики. Для получения представления об этих векторах развития теоретической и прикладной статистики в мировом масштабе выполнена систематизация тематики сессий конгресса с выделением двух разделов:

1. Общесистемная тематика;
2. Специальная тематика.

В первом разделе «Общесистемная тематика» на основе анализа наименований и содержания IPS были выделены следующие тематические группы:

— Статистические методы анализа и прогнозирования ситуации, связанной с пандемией COVID-19;

— Математическая статистика, методы интегрирования данных;

— Развитие официальной статистики, партнерских отношений, международного сотрудничества в сфере статистики;

— Новые источники статистических данных, развитие науки о данных в официальной статистике;

— Пространственная статистика;

— Индикаторы ЦУР;

— История статистики.

Второй раздел «Специальная тематика» представлен следующими группами IPS:

— Методология и прикладные решения в социально-экономической и отраслевой статистике;

— Финансовая статистика;

— Социальная статистика;

— Статистическая грамотность, обучение статистике;

— Экологическая, медицинская, биологическая статистика.

На рис. 1 и 2 представлены результаты группировки сессий конгресса по выделенным направлениям.

Определенные выводы о смене приоритетов в направлениях статистических исследований стран — участников конгресса можно сделать, сравнив результаты частотного анализа употребления профессиональных статистических терминов в тематике сессий. В обзорном анализе результатов предыдущего 62-го Всемирного статистического конгресса МСИ (Малайзия, 2019 г.), опубликованном в журнале «Вопросы статистики», было отмечено, что «... наиболее часто встречающимися терминами в тематике его сессий и конференций были (по убыванию): официальная статистика, большие данные, Цели устойчивого развития, окружающая среда, бедность, наука о данных (data science), временные ряды, робастность» [1, с. 85–86].

Сравнение данного рейтинга с рейтингом частоты использования этих же терминов в тематике сессий 63-го конгресса (см. рис. 3) показывает, что сохранились лидирующие позиции тематического объекта «официальная статистика»

¹ Нэн Лэрд (Nan Laird) — профессор медицинской школы им. Харви В. Файнберга, заслуженный специалист по биостатистике в Гарвардской школе общественного здравоохранения им. Т.Х. Чана, в котором с 1990 по 1999 г. занимала должность заведующего кафедрой.

² Лонгитюдные исследования (longitudinal study от longitude — долговременный) — научный метод, применяемый, в частности, в социологии и психологии, в котором изучается одна и та же группа объектов (в психологии — людей) в течение времени, за которое эти объекты успевают существенным образом поменять какие-либо свои значимые признаки. URL: <https://ru.wikipedia.org/wiki/%>

Рис. 1. Распределение сессий 63-го Всемирного статистического конгресса МСИ по основным темам в рамках направления «Общесистемная тематика» (число сессий)

Рис. 2. Распределение сессий 63-го Всемирного статистического конгресса МСИ по основным темам в рамках направления «Специальная тематика» (число сессий)

Рис. 3. Сравнительный анализ рейтингов частоты употребления наиболее распространенных терминов в тематике сессий 62-го (2019 г.) и 63-го (2021 г.) Всемирных статистических конгрессов МСИ

и резко увеличилась частота применения термина «временные ряды», что привело к его перемещению с 7-го на 2-е место.

На основе частотного анализа словосочетаний в тематике 63-го конгресса можно выя-

вить основные дискуссионные области с учетом употребления наиболее популярных терминов. В частности, на рис. 4–6 представлены тематические связи терминов «большие данные», «официальная статистика» и «COVID-19».

Рис. 4. Основные тематические связи термина «большие данные» в наименованиях сессий 63-го Всемирного статистического конгресса МСИ (результаты частотного анализа)

Рис. 5. Основные тематические связи термина «официальная статистика» в наименованиях сессий 63-го Всемирного статистического конгресса МСИ (результаты частотного анализа)

Рис. 6. Основные тематические связи термина «COVID-19» в наименованиях сессий 63-го Всемирного статистического конгресса МСИ (результаты частотного анализа)

На основе систематизации математических методов, представленных в тематике раздела «Математическая статистика, методы интегрирования данных», а также математико-статистических методов, указанных в тематике других разделов, выделены следующие наиболее актуальные области развития математического обеспечения статистических исследований:

- статистика экстремальных значений;
- «невероятностные» (non-probability sampling)³, малые выборки;
- непараметрические методы анализа линейных и нелинейных связей;
- временные ряды: обнаружения точки перелома нескольких временных рядов, анализ последовательности событий, сетевой анализ;
- работа с пропущенными и неполными данными;
- методы машинного/статистического обучения;
- методы интегрирования исходных и гранулированных данных статистических наблюдений и альтернативных источников данных.

Актуальная тематика в фокусе персоналий

Традиционно в рамках Всемирных статистических конгрессов МСИ проходят мероприятия по популяризации работ ученых-статистиков, внесших в прошлом значимый вклад в развитие статистической науки. Прошедший конгресс также не явился исключением. Одной из таких площадок стала сессия IPS 33 – «Бетти, Марио, Пилар, Дани и Эстер: в числе известных статистиков». Организаторы этой сессии отметили, что большая часть статистических исторических исследований сосредоточена на изучении наследия статистиков из Европы и Северной Америки, задача же данной сессии – привлечь внимание к ученым других континентов, внесших большой вклад в теорию и методологию статистики. Кто же такие Бетти, Марио, Пилар, Дани и Эстер? Из содержания докладов, представленных на данной сессии, а также согласно другим источникам – это известные в конкретных профессиональных областях статистики ученые:

Бетти – Фрэнсис Элизабет Аллан (Frances Elizabeth Allan) (1905–1952), австралийский статисти-

стик, математик, работавшая в сфере применения статистических методов в сельском хозяйстве [2].

Марио – Марио Шеборг (Mario Wscheborg) (1939–2011), уругвайский ученый, развивавший в своих трудах теорию стохастических процессов, стохастической геометрии, методы обработки статистических выбросов.

Пилар – Пилар Иглесиас (Pilar Iglesias) (1960–2007), чилийский ученый и деятель в сфере государственной статистики [3], внесшая вклад в развитие методов «предиктивистской характеристики распределения» (predictivistic characteristics of distributions) на основе теоремы Де Финетти (De Finetti's theorem), а также методов построения регрессионных моделей в случае несоответствия совокупностей нормальному закону распределения, байесовских непараметрических методов (Bayesian nonparametrics) [4].

Дани – Дани Герхардус Криг (Danie G. Krige) (1919–2013), южноафриканский статистик и горный инженер, один из основателей геостатистики [5].

Эстер – Эстер Самуэль-Кан (Ester Samuel-Cahn) (1933–2015), израильский статистик и педагог, внесшая вклад в теорию сложных решений (compound decision problems) в области математической статистики [6].

Всемирный статистический конгресс МСИ предоставляет известным ученым уникальную возможность выступить с лекцией на одной из самых значимых площадок конгресса – Gosset Lecture (Лекция в честь Уильяма С. Госсета – британского ученого-статистика, более известного под псевдонимом Стьюдент).

Выбор лектора проводится в период подготовки к конгрессу коллегиально с учетом актуальности и оригинальности его работ, определяющих новые тренды развития мировой статистической мысли. В апреле 2021 г. на заседании Совета МСИ, в котором участвовала автор статьи, было принято единогласное решение пригласить для чтения «Лекции Госсета» известного ученого Галит Шмуэли (Galit Shmueli) – заслуженного профессора и директора Института сервисных наук Тайваньского Национального университета Цин Хуа.

Д-р Шмуэли ведет курсы по интеллектуальному анализу данных (Data mining), прогнозной

³ «Невероятностная» выборка (non-probability sampling) определяется как метод выборки, при котором исследователь отбирает образцы на основе собственного субъективного суждения, а не случайного выбора. Этот метод выборки в большой мере зависит от опыта исследователей; широко используется для качественных исследований. URL: <https://www.questionpro.com/blog/non-probability-sampling/>.

аналитике, интерактивной визуализации, методам исследования и другим темам бизнес-аналитики [7].

Исследования д-ра Шмуэли сосредоточены на методологии статистического анализа и интеллектуального анализа данных в информационных системах, в здравоохранении, а также в исследовании поведения человека. Она является автором нескольких книг, в том числе популярного учебника «Data Mining for Business Analytics» («Методы интеллектуального анализа данных в бизнес-аналитике») [8], а также монографии «Information Quality: The Potential of Data and Analytics to Generate Knowledge» («Качество информации: потенциал данных и аналитики для генерации знаний») [9].

Тема лекции д-ра Шмуэли «The Language of Statistics (and What's Lost in Translation)» [«Язык статистики (и что теряется при переводе)»] раскрывает три проблемы: 1) отличие объясняющих статистических моделей от предсказывающих; 2) качество информации; 3) повышение надежности прогноза.

Автор предложила решение этих проблем для конкретных статистических исследований в аспекте выявления и снятия противоречий между терминами: Theory – Data (теория – данные); Causation – Association (причинная связь – ассоциация); Retrospective – Prospective (ретроспектива – перспектива); Bias – Variance (статистическая ошибка – дисперсия); Average Unit – Individual Unit (средняя величина – индивидуальная единица). Предложения автора по комплексному решению этих противоречий составляют значимый вклад в развитие науки о данных, что было подчеркнуто участниками конгресса при обсуждении лекции.

Российские участники 63-го Всемирного статистического конгресса МСИ внесли существенный вклад в развитие общей актуальной тематики дискуссий, организацию и проведение трех сессий, в том числе:

– IPS 157 – Poverty and Inequality: New Challenges and New Statistical Responses («Бедность и неравенство: новые вызовы и новые статистические ответы») под руководством д-ра экон. наук, профессора, члена-корреспондента РАН, избранного члена Международного статистического института И.И. Елисеевой;

– IPS 118 – Soviet Statistics as a Historical Artefact – («Советская статистика как исто-

рический артефакт») под руководством канд. экон. наук, профессора, руководителя Российской ассоциации статистиков, избранного члена Международного статистического института А.Н. Пономаренко;

– IPS 244 – Developing Machine Learning Methods in Experimental Official and Analytical Statistics («Развитие методов машинного обучения в экспериментальной официальной и аналитической статистике») под руководством д-ра экон. наук, профессора, избранного члена и члена Совета Международного статистического института Е.В. Заровой.

* *
*

Представленные результаты анализа тематических направлений 63-го Всемирного статистического конгресса МСИ в сопоставлении с результатами предыдущих конгрессов [1 и 10] создают возможность для ориентации в актуальной перспективной тематике статистических исследований в мировом масштабе, сравнения проблематики отечественной статистической науки с ведущими мировыми трендами, выявления новых направлений поиска для целей развития статистической теории и методологии.

Литература

1. Зарова Е.В. 62-й Всемирный статистический конгресс Международного статистического института. Вопросы статистики. 2019. Т. 26. № 9. С. 85–87.
2. Heyde C.C. Frances Elizabeth Allan // Australian Dictionary of Biography. 1993. Vol. 13. URL: <https://adb.anu.edu.au/biography/allan-frances-elizabeth-betty-9330>.
3. Del Pino G., Icaza G. Pilar Iglesias: A Life of Contributions to Statistics // Chilean Journal of Statistics. 2010. Vol. 1. No. 1. P. 5–15. URL: [http://soche.cl/chjs/volumes/01/01/delPino_Icaza\(2010\).pdf](http://soche.cl/chjs/volumes/01/01/delPino_Icaza(2010).pdf).
4. Arellano-Valle R.B., Bolfarine H., Iglesias P.L. A Predictivistic Interpretation of the Multivariate Distribution // Test. 1994. Vol. 3. Iss. 2. P. 221–236. doi: <https://doi.org/10.1007/BF02562703>.
5. Chilès J.-P., Desassis N. Fifty Years of Kriging // Daya Sagar B.S., Cheng Q., Agterberg F. (eds) Handbook of Mathematical Geosciences: Fifty Years of IAMG. Springer, Cham., 2018. P. 589–612. doi: https://doi.org/10.1007/978-3-319-78999-6_29.
6. Samuel E. The Compound Statistical Decision Problem // Sankhyā: The Indian Journal of Statistics. Series A (1961–2002). 1967. Vol. 29. No. 2. P. 123–140. URL: <https://www.jstor.org/stable/25049459>.
7. Galit Shmueli. Bio. URL: <https://www.galitshmueli.com/content/bio>.

8. **Shmueli G., Bruce P.C.** *Data Mining for Business Analytics. Concepts, Techniques, and Applications with XLMiner.* New Jersey: John Wiley & Sons, 2013.

9. **Kenett R.S., Shmueli G.** *Information Quality: The Potential of Data and Analytics to Generate Knowledge.* Wiley, 2016.

10. **Зарова Е.В.** Актуальные вопросы статистической науки и статистического образования в выступлениях специалистов на 61-м Всемирном статистическом конгрессе // *Вопросы статистики.* 2017. № 8. С. 91–92.

Информация об авторе

Зарова Елена Викторовна — д-р экон. наук, профессор, заместитель руководителя проектного офиса ГБУ «Аналитический центр» Правительства Москвы; профессор кафедры статистики РЭУ имени Г.В. Плеханова. 119019, г. Москва, ул. Новый Арбат, д. 11, стр. 1. E-mail: Zarova.ru@gmail.com. ORCID: <https://orcid.org/0000-0003-0375-2534>.

References

1. **Zarova E.V.** The 62nd ISI World Statistics Congress. *Voprosy Statistiki.* 2019;26(9):85–87. (In Russ.)

2. **Heyde C.C.** Frances Elizabeth Allan. In: *Australian Dictionary of Biography.* 1993. Vol. 13. Available from: <https://adb.anu.edu.au/biography/allan-frances-elizabeth-betty-9330>.

3. **Del Pino G., Icaza G.** Pilar Iglesias: A Life of Contributions to Statistics. *Chilean Journal of Statistics.* 2010;1(1):5–15. Available from: [http://soche.cl/chjs/volumes/01/01/delPino_Icaza\(2010\).pdf](http://soche.cl/chjs/volumes/01/01/delPino_Icaza(2010).pdf).

4. **Arellano-Valle R.B., Bolfarine H., Iglesias P.L.** A Predictivistic Interpretation of the Multivariate Distribution. *Test.* 1994;3(2):221–236. Available from: doi: <https://doi.org/10.1007/BF02562703>.

5. **Chilès J.-P., Desassis N.** Fifty Years of Kriging. In: Daya Sagar B.S., Cheng Q., Agterberg F. (eds) *Handbook of Mathematical Geosciences: Fifty Years of IAMG.* Springer,

Cham.; 2018:589–612. Available from: doi: https://doi.org/10.1007/978-3-319-78999-6_29.

6. **Samuel E.** The Compound Statistical Decision Problem. *Sankhyā: The Indian Journal of Statistics. Series A (1961–2002).* 1967;29(2):123–140. Available from: <https://www.jstor.org/stable/25049459>.

7. **Galit Shmueli.** *Bio.* Available from: <https://www.galitshmueli.com/content/bio>.

8. **Shmueli G., Bruce P.C.** *Data Mining for Business Analytics. Concepts, Techniques, and Applications with XLMiner.* New Jersey: John Wiley & Sons; 2013.

9. **Kenett R.S., Shmueli G.** *Information Quality: The Potential of Data and Analytics to Generate Knowledge.* Wiley; 2016.

10. **Zarova E.V.** Relevant Issues of Statistical Science and Statistical Education in Speeches of Specialists at the 61st World Statistics Congress. *Voprosy Statistiki.* 2017;(8):91–92. (In Russ.).

About the Author

Elena V. Zarova — Dr. Sci. (Econ.), Professor, Deputy Head, Project Office, Analytical Center by Moscow City Government; Professor, Department of Statistics, Plekhanov Russian University of Economics. 11, New Arbat Ave., Bldg. 1, Moscow, 119019, Russia. E-mail: Zarova.ru@gmail.com. ORCID: <https://orcid.org/0000-0003-0375-2534>.