От редакции

Ниже приводится статья Ю.Н. Тюрина и Д.С. Шмерлинга – известных отечественных специалистов в области непараметрической статистики. Ее публикация является продолжением реализации описанного в № 16 журнала замысла Редакционного Совета: регулярной публикации сведений об известных, актуальных для социологии математических методах, трудно доступных в настоящее время для широкого круга читателей.

Ю.Н. Тюрин, Д.С. Шмерлинг
(*Москва*)

НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ СТАТИСТИКИ

Данная публикация подготовлена в жанре энциклопедической статьи. В ней описаны основные достижения в области непараметрической статистики – совокупности методов, привлекательных для исследователя тем, что они позволяют решать часто встречающиеся в социологии задачи: анализ связи между признаками, проверка статистических гипотез и т.д. В то же время они не опираются на трудно проверяемые и далеко не всегда выполняющиеся предположения о характере распределения исходных признаков. Дается подробнейшая библиография по непараметрической статистике.

Ключевые слова: непараметрические методы статистики, распределение вероятностей, параметрический закон распреде-

Юрий Николаевич Тюрин – доктор физико-математических наук, профессор МГУ.
Дмитрий Семенович Шмерлинг – кандидат физико-математических наук, профессор ВШЭ ГУ .

ления, выборка, генеральная совокупность, статистическая гипотеза, непараметрический критерий, ранговый критерий, коэффициент ранговой корреляции, ранговый метод, критерий Колмогорова, критерий Уилкоксона, критерий Стьюдента, точечное оценивание параметра, интервальное оценивание параметра, доверительный интервал, доверительная вероятность, мощность критерия, устойчивость статистического вывода, таблицы распределений, непараметрический регрессионный анализ, непараметрический дискриминантный анализ.

Непараметрическая статистика – непараметрические методы статистики, непараметрика – часть математической статистики, комплекс методов обработки статистических данных, не требующих, чтобы распределение вероятностей было описано каким-либо параметрическим законом распределения (например, нормальным). Она опирается на более широкие и менее ограничительные свойства распределений вероятностей: статистическую независимость распределений (ошибок наблюдений), непрерывность этих распределений; часто – на ту или другую симметрию распределений и т.п.

Отрицание, содержащееся в названии этого направления, имеет исторические корни: в прошлом (30-е гг.) оно возникло как альтернатива господствовавшей тогда системе обработки данных, основанной на гауссовском (нормальном) распределении. Совокупность одномерных гауссовских распределений образует двупараметрическое семейство. Существуют и другие параметрические семейства распределения вероятностей, например, показательное, логнормальное, распределение Парето и т.д. «Непараметрические» как название для нового метода подчеркивало его универсальную применимость к непрерывным одномерным распределениям.

Первоначально непараметрические методы предназначались для проверки статистических гипотез (об одномерных распределениях вероятностей). Наиболее известные непараметрические критерии – это критерии Колмогорова-Смирнова, изобретенные в 1930-х гг., ранговые критерии Уилкоксона (Вилкоксона, Wilkoxon) и Манна-Уитни (Mann, Whitney) 1940–1950-х гг. и, конечно, коэффициенты ранговой корреляции (1904–1930-х гг.) Спирмена и Кенделла. Они породили целые научные направления [1; 8; 10; 12; 13; 14; 17; 24; 26; 32; 37].

Критерий Колмогорова предназначен для проверки гипотезы о том, что данная выборка $x_1, \ldots, x_n$ извлечена из распределения с данной функцией $F(*)$, которая предполагается непрерывной. Определим эмпирическую функцию распределения, положив $F(*) = \frac{1}{n}\{\text{число } x_i < x\}$. Функция $F_n(x)$ – ступенчатая; скачки вверх величиной $\frac{1}{n}$ происходят в точках выборки. Критерии обсуждаемого типа основаны на сопоставлениях функций $F_n(*)$ и $F(*)$. А.Н. Колмогоров рассмотрел статистику

$$D_n = \sup_x \left|F_n(x) - F(x)\right|.$$

Для случайной величины $\sqrt{n}D_n$ он нашел предельное распределение при $n \to \infty$, в случае, когда гипотеза H верна. Если же H неверна, то $\sqrt{n}D_n \to \infty$.

Обратим внимание на примечательное свойство статистики D_n: распределение этой статистики одинаково для всех непрерывных распределений $F(*)$. Это свойство, называемое свободой от распределения (*distribution – free*), позволило вычислить распределения D_n для каждого n и одинаково пользоваться ими при решении любых задач.

Пример использования критерия Колмогорова для проверки H. Вычисляем значение D_n. Выбираем уровень значимости ε. Извлекаем из таблицы распределения D_n величину $(1 - \varepsilon)$ квантили, которая выступает далее как критическое значение для D_n: гипотезу H мы отвергаем, если найденное значение D_n превосходит критическое значение.

Еще одно теоретически важное свойство критерия Колмогорова – этот критерий состоятелен против всех альтернатив. Это означает, что если подлинное распределение выборки отличается от гипотетического $F(*)$, то при достаточно большом числе наблюдений n с помощью критерия Колмогорова это отличие будет обнаружено. К сожалению, эта универсальность оборачивается тем, что мощность[1] критерия Колмогорова (и других родственных ему универсально состоятельных критериев) оказывается довольно низкой против некоторых важных альтернатив, например, альтернативы сдвига. В этих случаях лучше применять упомянутые ранее ранговые критерии типа критерия Уилкоксона.

Эти критерии основаны не на самих наблюдениях, а на их рангах. Это могут быть также ранги остатков (разностей между величиной наблюдения и значением функции регрессии), если для описания наблюдений принята модель более сложная, чем просто выборка, например, модель регрессии. Ранги наблюдений – это те номера, которые присваиваются наблюдениям при их расстановке в определенном порядке (обычно в порядке возрастания). Свойство рангов элементов выборки из непрерывного распределения – все возможные последовательности рангов имеют равные вероятности для появления. Именно это свойство обеспечивает всем статистикам, основанным на рангах элементов выборки, свободу от распределения.

[1] Вероятность отвергнуть неверную гипотезу – важная характеристика статистического критерия.

Так как ранговые методы опираются на упорядочение наблюдений, они применимы только к числовым данным. Для многомерных наблюдений не существует естественного способа сравнения и упорядочивания. Именно поэтому непараметрические методы используются только для одномерных наблюдений. (Разработка непараметрических методов для многомерного статистического анализа – приоритетная область научных исследований в настоящее время.)

Критерий Уилкоксона (критерий ранговых сумм) дает представление о специфике ранговых методов. Он предназначен для проверки гипотезы о том, что две выборки $x_1, \ldots, x_m$ и $y_1, \ldots, y_n$ извлечены из общей генеральной совокупности. (Далее это гипотеза H.) Распределение генеральной совокупности предполагается непрерывным.

Упорядочим объединенную совокупность $x_1, \ldots, x_m$ и $y_1, \ldots, y_n$ в порядке возрастания и найдем, таким образом, ранги наблюдений. Обозначим через $s_1, \ldots, s_n$ ранги элементов второй выборки и рассмотрим их сумму $w_{m,n} = s_1 + \ldots + s_n$ (таким же образом можно действовать с рангами первой выборки). Основной результат, относящийся к статистике $w_{m,n}$, – при гипотезе H случайная величина $w_{m,n}$ распределена свободно, т.е. одинаково для всех непрерывных генеральных совокупностей. Поэтому для каждой пары чисел (m, n) распределение $w_{m,n}$ можно вычислить, используя таблицы распределений, которые даны во многих руководствах и справочниках [8; 47; 48; 50; 51].

Гипотезу H отвергают, если полученное опытным путем значение статистики $w_{m,n}$ значимо отклоняется (влево или вправо) от своего ожидаемого значения $\frac{1}{2} = n\,(m+n+1)$. Критические значения для этих отклонений содержатся в упомянутых таблицах.

Переход от точных значений исходных наблюдений к их порядковым номерам (рангам), очевидно, сопровождается некоторой потерей информации. Нередко эта потеря невелика. Так, по сравнению с критерием Стьюдента – наиболее подходящим статистическим правилом при сопоставлении двух гауссовских выборок (отличающихся средним) – критерий Уилкоксона оказывается лишь слегка менее мощным: его асимптотическая эффективность в этом случае составляет относительно критерия Сьюдента $3/\pi \approx 0{,}95$ [9; 11].

Это значит, что, если мы (по осторожности или по незнанию) при сравнении двух гауссовских выборок объемом 100 и 100 воспользуемся не критерием Стьюдента, а критерием Уилкоксона, мы получим в результате такую же точность, как если бы наши выборки имели объемы 95 и 95 (и мы использовали критерий Стьюдента).

Для негауссовских же выборок применять критерий Стьюдента вообще не рекомендуется.

Такая небольшая потеря эффективности в достаточной степени компенсируется еще одним важным достоинством ранговых методов: их устойчивостью по отношению к отступлениям от предложений статистической модели. (Надо помнить, что математическая модель всегда описывает реальность только приближенно.) Это свойство статистических правил сейчас называют робастностью (от англ. *robust* – крепкий, грубый, дюжий). Поясним один из аспектов робастности на примере тех же выборок.

Как правило, выборки, которые мы называем гауссовскими, лишь приближенно следуют нормальному распределению. Это проявляется в том, что среди наблюдений, если их несколько десятков (или сотен), находится несколько таких, которые нельзя получить из гауссовского закона – слишком далеко отстоят они от центра распределения (и центра выборки). Присутствие таких «выбросов» (англ. *outliers*, *gross errors* и т.п.), если их не исключить, может сильно исказить результаты статистической обработки, если пользоваться гауссовскими статистическими процедурами. Для ранговых же методов наличие небольшого числа пусть даже очень больших выбросов влияет на результат незначительно. Таким образом, ранговые методы более устойчивы к выбросам, засорениям и прочим несовершенствам статистического материала.

В более широком аспекте исследованием и разработкой устойчивых статистических правил занимается особый раздел математической статистики, называемой робастной статистикой [55; 56; 57; 70; 71; 72; 73].

В 60-х гг. было обнаружено (Hodges, Lehmann), что ранговые статистические критерии можно применять и для оценки неизвестных параметров статистических моделей. Так, с помощью статистики ранговых сумм Уилкоксона Ходжес и Леман построили свободные от распределения доверительные интервалы для параметра сдвига одной выборки относительно другой. Подчеркнем, что закон распределения, сформировавший эти выборки, при этом подходе остается неопределенным (и неизвестным исследователю). Единственное требование – этот закон распределения должен быть непрерывным.

Располагая системой вложенных, доверительных интервалов для интересующего нас параметра (доверительный интервал тем шире, чем выше его доверительная вероятность), уже нетрудно указать для него и точечную оценку (приближенное значение): такой оценкой для него может быть число, общее для всех доверительных интервалов. Так, в задаче о сдвиге одной выборки относительно другой, «ранговой» оценкой сдвига может служить медиана совокупности попарных разностей $y_j - x_i$, где $i = \overline{1, m}$ и $j = \overline{1, n}$. Мы взяли термин ранговый в кавычки, поскольку ранги в конечной формуле не присутствуют, а участвуют лишь на стадии ее вывода [8; 10; 14; 29].

К настоящему времени непараметрические методы (в первую очередь – ранговые методы и их модификации и расширения) образуют систему обработки статистических данных, по своим возможностям не уступающую классическому методу наименьших квадратов (теория которого базируется на гауссовском распределении ошибок).

Помимо ранговых, созданы и другие непараметрические системы обработки, например, основанные на знаках наблюдений или остатков (знаковый критерий Фишера, см. также [29]). По непараметрической статистике есть много хороших учебников [8; 10; 12; 13; 14; 15; 18; 19; 20; 37; 38; 39; 59; 65; 68 и др.]. Несколько в стороне мы оставляем анализ категоризированных наблюдений [59 и др.].

Потеря информации, которая происходит при переходе к рангам или знакам наблюдений, компенсируется широтой применений и устойчивостью статистических выводов относительно грубых ошибок, неточностей моделей и т.п. К важным достоинствам непараметрических методов относится математическая простота большей части статистических правил. Таблицы распределений и процентных точек (квантилей), которыми необходимо снабдить каждый статистический критерий, теперь отчасти заменяют пакеты статистических программ. В лучших из них непараметрические методы представлены широко.

Обсуждавшиеся до сих пор статистические правила были точными в том смысле, что доверительные вероятности, уровни значимости и т.д. выполняются точно (если соблюдаются условия модели). Правила эти применимы к выборкам (к статистическому материалу) любого объема, в том числе и малого. Чтобы подчеркнуть это важное обстоятельство, статистические методы такого рода иногда выделяют под названием «статистики малых выборок». В этом есть смысл, так как для больших выборок можно применять асимптотические методы их анализа. Эти методы приближенные, но точность их возрастает с увеличением объема наблюдений. Выборка считается «большой», если точность статистического правила оказывается достаточной для поставленных целей исследования. Понятие «большая выборка» не вполне четкое и зависит от целей исследования и выбранных математических средств.

С большими выборками также связывают некоторые непараметрические задачи: непараметрическое оценивание регрессии (непараметрический регрессионный анализ) [4; 5; 6], непараметрическое оценивание плотности [1; 5; 49], непараметрический дискриминантный анализ [4], адаптивные методы [15; 64] и т.д. Ответы имеют асимптотический характер.

Задача непараметрического оценивания регрессии дает о них представление. Она ставится вполне традиционно: по наблюдениям (x_i, y_i), где переменные связаны зависимостью $y_i = f(x_i) + \varepsilon_i$, $i = 1, \ldots n$, случайные ошибки независимы, необходимо оценить неизвестную функцию $f(*)$. Особенность задачи – число наблюдений n неограниченно возрастает. Это условие придает задаче асимптотический характер. Термин непараметрический имеет здесь двойной смысл: ни распределение ошибок, ни сама функция $f(*)$ не определяется какими-либо формулами с параметрами.

Поскольку ответ следует дать для произвольной функции $f(*)$, необходимо, чтобы точки $x_1, \ldots, x_n$ со все возрастающей плотностью заполняли бы всю область определения $f(*)$. Если $f(*)$ непрерывна, то приближенное к $f(x)$ значение можно получить, усредняя наблюдения $y_i = f(x_i) + \varepsilon_i$ по окрестности точки x. Чтобы систематическая ошибка при этом убывала с ростом n, должно неограниченно увеличиваться число точек x_i, которые попадают в эту окрестность x (и, тем самым, участвуют в усреднении). Чтобы с ростом n убывала систематическая ошибка, ширина окрестности должна стремиться к нулю. Чем, по условиям задачи, более гладкой (и, следовательно, меняющейся медленнее) считается функция $f(*)$, тем более широкой может быть окрестность x и тем более точной оказывается статистическая оценка. Конечно, эти окрестности должны «стягиваться» к точкам $x_{i.}$, но, чем более гладкая функция, тем шире могут оставаться окрестности. Видно, что задача поставлена чисто математически. Экономические применения этой техники нам неизвестны.

Приложения непараметрических методов все чаще появляются в экономических и социально-политических зарубежных и отечественных журналах, чрезвычайно распространены в экспериментальной и социальной психологии, а через них – в маркетинге, социологии, теории надежности (в широком смысле), в политических исследованиях, планировании, изучении рисков, анализе категоризированных (классифицированных) данных.

ЛИТЕРАТУРА

1. Вероятность и математическая статистика: Энциклопедия / Гл. ред. Ю.В. Прохоров. М.: Большая российская энциклопедия, 1999.

2. *Ибрагимов И.А., Хасьминский Р.З.* Асимптотическая теория оценивания. М.: Наука, 1979.

3. *Фукунага К.* Введение в статистическую теорию распознавания образов / Пер. с англ. М.: Мир, 1979.

4. *Хардле В.* Прикладная непараметрическая регрессия / Пер. с англ. М.: Мир, 1993.

5. *Деврой Л., Дьерфи Л.* Непараметрическое оценивание плотности. L_1-переход / Пер. с англ. М.: Мир, 1988.

6. Алгоритмы и программы восстановления зависимостей / Под ред. В.Н. Вапника. М.: Наука, 1984.

7. Классификация и кластер / Ред. Дж. Вэн Райзин / Пер. с англ. М.: Мир, 1988.

8. *Холлендер М., Вулф Д.* Непараметрические методы статистики / Пер. с англ. М.: Финансы и статистика, 1983.

9. *Никитин Я.Ю.* Асимптотическая относительная эффективность непараметрических критериев. М.: Наука; Физматлит, 1985.

10. *Хеттманспергер Т.* Статистические выводы, основанные на рангах / Пер. с англ. М.: Финансы и статистика, 1987.

11. *Pitman E.J.G.* Notes on Non-parametric Statistical Inference. N.Y.: Columbia Univ., 1948 (dublicated).

12. *Кэндел М.* Ранговые корреляции / Пер. с англ. М.: Статистика, 1975.

13. *Gibbons J.D., Chakraborty S.* Nonparametric Statistical Inference. N.Y. e.a.: Marcel Dekker, 1992.

14. *Lehmann E.L.*, With spec. assist. of H.Y.M.D. Abrera. Nonparametric Statistical Methods Based on Ranks. S.F., N.Y. e.a.: Holden-Day – McGraw – Hill, 1975.

15. *Randles R.H., Wolfe D.A.* Introduction to the Theory of Nonparametric Statistics. N.Y. e.a.: Wiley, 1979.

16. *Puri M.L., Sen P.K.* Nonparametric Methods in Multivariate Analysis. N.Y. e.a.: Wiley, 1971.

17. *Puri M.L., Sen P.K.* Nonparametric Methods in General Linear Models. N.Y. e.a.: Wiley, 1985.

18. *Нискина Н.П., Тейман А.И., Шмерлинг Д.С.* Непараметрические методы статистики, основанные на рангах, и их применения: Препринт. М.: ВНИИ системн. иссл., 1986.

19. *Нискина Н.П., Тейман А.И., Шмерлинг Д.С.* Двухфакторный непараметрический дисперсионный анализ для неполных данных: Препринт. М.: ВНИИ системн. иссл., 1986.

20. *Нискина Н.П., Тейман А.И., Шмерлинг Д.С.* Специальные задачи двухфакторного дисперсионного анализа: Альтернативы упорядоченности обработок. М.: ВНИИ системн. иссл., 1986.

21. *Noether G.E.* Elements of Nonparametric Statistics. N.Y. e.a.: Wiley, 1967.

22. *Walsh Y.E.* Handbook of Nonparametric Statistics. Princeton, N.Y.: Van Nostvand, 1962, 1965, 1968. Vol. 1–3.

23. *Рунион Р.* Справочник по непараметрической статистике / Пер. с англ. М.: Финансы и статистика, 1982.

24. *Hollander M., Wolfe D.A.* Nonparametric Statistical Methods. 2nd ed. N.Y. e.a.: Wiley, 1999.

25. *Мартынов Г.В.* Критерии омега-квадрат. М.: Наука, 1978.

26. *Pratt J.W., Gibbons J.D.* Concepts of Nonparametric Theory. N.Y.: Springer – Verlag, 1981.

27. Goodness-of-Fit Techniques / Ed. by R.B. D'Agostino, M.A. Stephens. N.Y. e.a.: M. Dekker, 1986.

28. *Greenwood P.E., Nikulin M.S.* A Guide to Chi-squared Testing. N.Y.: Wiley, 1996.

29. *Болдин М.В., Симонова Г.И., Тюрин Ю.Н.* Знаковый статистический анализ линейных моделей. М.: Наука; Физматлит, 1987.

30. *Hettmansperger T.P., McKean J.W.* Robust Nonparametric Statistical Methods. L.: Arnold, 1998.

31. *Jureckova J., Sen P.K.* Robust Statistical Procedure: Asymptotics and Interrelations. N.Y. e.a.: Wiley, 1996.

32. *Гаек Я., Шидак З.* Теория ранговых критериев / Пер. с англ. М.: Наука, 1971.

33. Nonparametric Methods: Handbook of Statistics / Ed. by P.R. Krishnaiah, P.K. Sen. N.Y. e.a.: North-Holland Elsevier,1984. Vol. 4.

34. *Sen P.K.* Sequential Nonparametrics: Invariance Principles and Statistical Inference. N.Y. e.a.: Wiley, 1981.

35. *David H.A.* The Method of Paired Comparisons. 2nd ed. L., N.Y.: Griffin, Oxford Univ. Press,1988. (С 1-го издания 1963 г. (1970) есть русский перевод: Дэвид Г. Метод парных сравнений / Пер. с англ. М.: Статистика, 1978.)

36. Probability Models and Statistical Analysis for Ranking Data / Ed. by M.A. Flinger, Y.S. Verducci. N.Y. e.a.: Springer – Verlag, 1993.

37. *Тюрин Ю.Н., Макаров А.А.* Статистический анализ данных на компьютере / Под ред. В.Э. Фигурнова. М.: Инфра-М, 1998.

38. *Siegel S.* Nonparametric Statistics for the Behavioral Sciences. N.Y.: McGraw – Hill, 1956.

39. *Fraser D.A.S.* Nonparametric Methods in Statistics. N.Y. – L.: Wiley, 1963.
40. *Shorack G.R., Wellner Y.A.* Empirical Processes with Applications to Statistics. N.Y., 1986.
41. *Леман Э.* Проверка статистических гипотез / Пер. с англ. 2-е изд. М.: Наука, 1979.
42. *Королюк В.С., Боровских Ю.В.* Теория *U*-статистик. Киев: Наукова Думка, 1989.
43. *Питмен Э.* Основы теории статистических выводов / Пер. с англ. М.: Мир, 1986.
44. *Рао С.Р.* Линейные статистические методы и их применения / Пер. с англ. М.: Наука, 1968.
45. *Бикел П., Доксам К.* Математическая статистика / Пер. с англ. М.: Финансы и статистика, 1983. Вып. 1–2.
46. *Кокс Д., Хинкли Д.* Теоретическая статистика / Пер. с англ. М.: Мир, 1978.
47. *Ликеш И., Ляга Й.* Основные таблицы математической статистики / Пер. с чешск. М.: Финансы и статистика, 1985.
48. *Большев Л.Н., Смирнов Н.В.* Таблицы по математической статистике. 3-е изд., перераб. и доп. М.: Наука, 1983.
49. *Silverman B.D.* Density Estimation for Statistics and Data Analysis. L.: Chapman and Hall, 1986.
50. *Оуен Д.Б.* Сборник статистических таблиц / Пер. с англ. 2-е изд., испр. М.: ВЦ АН СССР, 1973.
51. *Мюллер П., Нейман П., Шторм Р.* Таблицы по математической статистике / Пер. с нем. М.: Финансы и статистика, 1985.
52. Справочник по прикладной статистике: В 2 т. / Под ред. Э. Ллойда, У. Ледермана; Пер. с англ. М.: Финансы и статистика, 1990.
53. *Groeneboom P., Wollner Y.A.* Information Bounds and Nonparametric Maximum Likelihood Estimation. Birkthauser, 1982.
54. *Green P.J., Silverman B.W.* Nonparametric Regression and Generalized Linear Models: Roughness Penalty Approach. Charman and Hall, 1994.
55. *Huber P.J.* Robust Statistical Procedures. 2nd ed. SOAM, 1996.
56. *Wilcox R.* Introduction to Robust Estimation and Hypothesis Testing. Academic, 1997.
57. Robust Statistics, Data Analysis, and Computer Intensive Methods / Ed. by H. Rieder // In Honor of Peter Huber's Birthday. N.Y. e.a.: Springer – Verlag, 1996.
58. *Rieder H.* Robust Asymptotic Statistics. N.Y. e.a.: Springer – Verlag, 1996.
59. *Agresti A.* Categorical Data Analysis. N.Y. e.a.: Wiley, 1990.
60. *Schmidli H.* Reduced Rank Regression: With application to Quantitative Structure – Activity Relationships. N.Y. e.a.: Springer – Verlag, 1995.
61. *Marden Y.I.* Analysing and Modelling Rank Data. Chapman and Hall, 1989.

62. *Janssen A., Mason D.M.* Non-standard Rank Tests. N.Y. e.a.: Springer – Verlag, 1990.

63. *Kendall M., Gibbons Y.D.* Rank Correlation Methods. 5th ed. Edward Arnold, 1990.

64. *Benhen K., Neuhaus G.* Rank Tests with Estimated Scores and their Applications. B.G. Teubner, 1989.

65. *Sprent P.* Applied Nonparametric Statistical Methods. Chapman and Hall, 1989.

66. *Buning H.* Robust and Adaptive Tests (German: Robusty und Adaptive Tests). Berlin: Walter de Gruyter, 1991.

67. *Denker M.* Asymptotic Distribution Theory in Nonparametric Statistics. Wieweg, Friedr., & son Verlag, 1985.

68. *Martiz J.S.* Distribution – Free Statistical Methods. 2nd ed. Chapman and Hall, 1995.

69. *Pagan A., Ullah A.* Nonparametric Econometrics. N.Y.: Cambr. Univ. Press, 1999.

70. Устойчивые статистические методы оценки данных / Под ред. Р.Л. Лонераи, Г.Н. Уилкинсона; Пер. с англ. М.: Машиностроение, 1984.

71. *Хампель Ф. и др.* Робастность в статистике: Подход на основе функций влияния / Пер. с англ. М.: Мир, 1989.

72. *Хьюбер П.* Робастность в статистике / Пер с англ. М.: Мир, 1989.

73. *Смоляк С.А., Титаренко В.П.* Устойчивые методы оценивания (Статистическая обработка неоднородных совокупностей). М.: Статистика, 1980.

74. *Колмогоров А.Н.* Об эмпирическом определении закона распределений (1933), Об одном новом подтверждении законов Менделя (1940) // *Колмогоров А.Н.* Теория вероятностей и математическая статистика. М.: Наука, 1986. С. 134–141; С. 209–214.