

# ИСТОРИЯ НАУКИ

Самарская Лука: проблемы региональной и глобальной экологии.  
2020. – Т. 29. – № 2. – С. 132-154.

УДК 519.233.4+574

DOI 10.24411/2073-1035-2020-10330

## ДИСПЕРСИОННЫЙ АНАЛИЗ В ЭКОЛОГИИ (К 130-ЛЕТИЮ СО ДНЯ РОЖДЕНИЯ РОНАЛЬДА ФИШЕРА)

© 2020 Г.С. Розенберг

Институт экологии Волжского бассейна –  
филиал Самарского федерального исследовательского центра РАН, г. Тольятти (Россия)

Поступила 17.02.2020

**Розенберг Г.С. Дисперсионный анализ в экологии (к 130-летию со дня рождения Рональда Фишера).** Составляя календарь памятных дат на 2020 г. [66], я обнаружил не юбилейную, но «круглую дату» – 130 лет со дня рождения одного из крупнейших биологов и математических статистиков XX века Рональда Фишера. Разработанный им для поиска зависимостей в экспериментальных данных путём исследования значимости различий в средних значениях метод, получивший название дисперсионный анализ, до сих пор остается одним из самых востребованных методов обработки статистической информации (в том числе, и в экологии). Это и послужило поводом для написания статьи, в которой обсуждаются некоторые проблемы корректного использования данного статистического метода.

*Ключевые слова:* Рональда Фишер, дисперсионный анализ, математическая статистика, генетика, экология, основная теорема Фишера о естественном отборе.

**Rozenberg G.S. Analysis of variance in ecology (on the 130<sup>th</sup> anniversary of Ronald Fisher).** When compiling a calendar of memorable dates for 2020 [66], I discovered not a jubilee, but a “round date” – the 130<sup>th</sup> anniversary of the birth of one of the greatest biologists and mathematical statistics of the twentieth century, Ronald Fisher. Developed by him to search for dependences in experimental data by studying the significance of differences in average values, the method, called analysis of variance, is still one of the most popular methods for processing statistical information (including ecology). This was the reason for writing the article, which discusses some problems of the correct use of this statistical method.

*Key words:* Ronald Fisher, analysis of variance, mathematical statistics, genetics, ecology, Fisher's fundamental theorem of natural selection.

*...замечательные люди исчезают у нас, не оставляя по себе следов. Мы ленивы и нелюбопытны...*

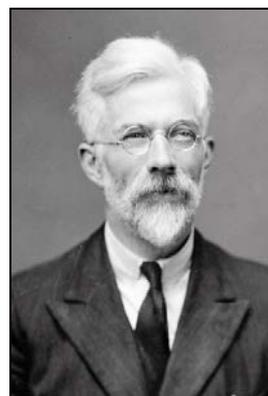
А.С. Пушкин, 1836  
[55, с. 668]<sup>1</sup>



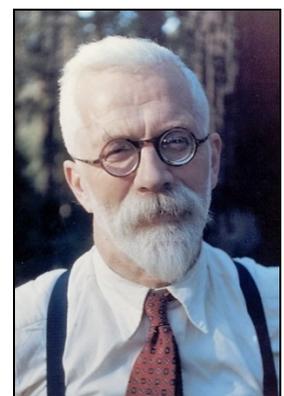
1913



1932



1943



1961

Британский зоолог, этолог и эволюционный биолог Р. Докинз (Clinton Richard Dawkins; г. р. 1941) назвал Рональда Фишера величайшим биологом, подобным Дарвину, и отцом современной статистики («Sir Ronald Fisher, the formidable English geneticist and mathematician, who could be regarded as Darwin's greatest twentieth-century successor as well as the father of modern statistics» [100, p. 39]). Не думаю, что это – завышенная оценка.

### Биографические подробности

Рональд Фишер родился 17 февраля 1890 г. в Ист-Финчли (East Finchley), в пригороде на севере Лондона. Его отец Джордж Фишер (George Fisher) был успешным аукционистом изобразительного искусства. Детство Рональда было счастливым, семья жила весьма комфортно, он был обожаем тремя старшими сёстрами, старшим братом и матерью, Кэти Фишер (Katie Fisher). Мать умерла, когда Рональду было 14 лет, а полтора года спустя отец обанкротился, проведя несколько неудачных сделок [97].

Математические способности Рон проявлял с детства. «У него было очень плохое зрение, и он изучал математику без помощи ручки и бумаги, развив способность представлять решение задач в геометрических терминах, без использования алгебраических преобразований. Способность Фишера получать результат, минуя сложную последовательность промежуточных шагов, стала легендой» (<https://cumir.ru/ronald-ejlmr-fisher>). В 1909 г. Фишер добился стипендии на обучение в Кембридже (Cambridge) и поступил в Gonville and Caius College<sup>1</sup>, который с отличием закончил в 1912 г.

Увлёкшись статистикой, Рон глубоко изучил статьи Карла Пирсона из серии «Математический вклад в теорию эволюции – Mathematical Contributions to the Theory of Evolution» (в основном в «Философских трудах» Лондонского королевского общества) и особенно заинтересовался применением в биометрии статистических методов. Еще будучи студентом, он публикует работу «Об абсолютном критерии для

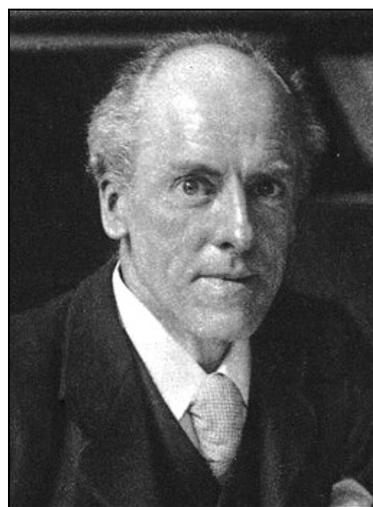
---

*Розенберг Геннадий Самуилович*, доктор биологических наук, чл.-корр. РАН, [genarozenberg@yandex.ru](mailto:genarozenberg@yandex.ru);

<sup>1</sup> Каждый раз, когда пишу работу биографического (исторического) плана, я всегда дискутирую с Александром Сергеевичем, пытаюсь доказать, что «не всё так запущено...». И всегда его цитирую.

<sup>1</sup> Колледж основан в 1348 г. Среди его выпускников – двенадцать лауреатов Нобелевской премии; рядом с колледжем состоялось прощание с астрофизиком Стивенем Хокингом (Stephen William Hawking), который проработал здесь больше 50 лет.

приближения частотных кривых» [69, 105], в которой предлагает оригинальный метод оценки параметров *метода максимального правдоподобия*. В это же время проявился интерес Фишера и к проблемам наследственности [95, 104]; причем, этот интерес сочетался у него с приверженностью к евгенике<sup>2</sup>: так, в 1911 г. Фишер принимает участие в создании Евгенического Общества Кембриджского Университета (Cambridge University Eugenics Society).



Карл Пирсон (Karl [Carl] Pearson; 1857-1936) – британский математик, статистик, биолог, философ; основатель математической статистики, один из основоположников биометрики; член Лондонского королевского общества (1896).

После окончания университета Фишер работал статистиком в лондонском Сити (City of London). Кроме того, во время Первой мировой войны (из-за крайне слабого зрения ему неоднократно отказывали в попытках попасть в армию) Фишер преподавал физику и математику в ряде государственных школ, в колледжах Thames Nautical Training College и Bradfield College; он был статистиком (актуарием) в торговом обществе, затем на маленькой сельскохозяйственной станции в Канаде.

В апреле 1917 г. Рональд на тайной церемонии (в деревне Greenhithe, графство Kent) женился на 17-летней Эйлин Гиннес (Ruth Eileen Guinness), мать которой не одобряла столь ранний брак. У них было двое сыновей и семь дочерей, одна из которых умерла в младенчестве [97].

Его интерес к генетике и статистике проявился в том, что в это же время он стал писать обзоры для журнала "Eugenic Review" и публикует

---

<sup>2</sup> Евгеника (от др.-греч. εὐγενής – хорошего рода, благородный) – учение о селекции применительно к человеку [16].

еще несколько этапных (и для него, и для науки) работ. Прежде всего, – это статьи «Распределение частот значений коэффициента корреляции в выборках из неопределенно большого населения» [106] и «Корреляция между родственниками в предположении менделевского наследования» [107]. Первая из них, заложила основы теории распределения выборок; вторая – «стала основой для целой отрасли, позже ставшей известной как *биометрическая генетика*, и внесла в науку *метод дисперсионного анализа*» (<https://cumir.ru/ronald-ejlmmer-fisher>).

Принято считать (впрочем, так оно и есть в действительности), что в математической статистике Фишер является виднейшим продолжателем классических работ и методов Карла Пирсона. В частности, Фишер, наряду с Ежи Нейманом и Эгоном Пирсоном<sup>3</sup> (сыном Карла Пирсона), разработал фундамент теории оценок параметров, статистических решений, планирования эксперимента и проверки гипотез. Однако его взаимоотношения с Пирсоном-старшим были далеки от идеальных, а с 1917 г. даже стали неприязненными [69, с. 64-65]<sup>4</sup>. Возможно именно по этому, когда после окончания войны Фишер искал новую работу и получил предложение от Карла Пирсона, возглавлявшего прославленную Galton Laboratory (лаборатория евгеники и генетики человека [13, 16, 29, 93, 25] при Университетском колледже Лондона [141], он отказался и вместо этого поступил на временную работу на сельскохозяйственную станцию, посчитав, что растущее соперниче-

ство с Пирсоном может стать своего рода профессиональным препятствием.

В 1919 г. Дж. Рассел<sup>5</sup> приглашает Фишера, чтобы посмотреть, сможет ли статистик сделать что-нибудь с массой данных по изменению урожая, собранных на Экспериментальной станции Ротамстед ([94], URL). Работа на опытной станции послужила базой для крупных исследований и предоставила Фишеру доступ к обширному собранию данных, фиксируемых в течение многих лет; это стало плацдармом для развития новых статистических методов и статистики как науки в целом. Он «задержался» на станции до 1933 г., создав там Отдел статистики (привлекший паломников со всего мира), разработав спектр новых биометрических подходов, опубликовав монографии "Statistical Methods for Research Workers" (1925; рус. пер. [83]; эта книга произвела революцию в прикладной статистике, скорректировав и даже заменив методы, которые Пирсон ввел на рубеже веков), "The Arrangement of Field Experiments" (1926) и важную статью монографического плана [110]. В этих работах Фишер предлагает новую теорию статистической оценки в противовес байесовскому подходу<sup>6</sup>, реконструирует теорию *критерия хи-квадрат Пирсона* [109] и расширяет сферу *распределения Стьюдента* [114]. Эти разработки, как и дисперсионный анализ, опирались на новую систему теории распределений, основанную на взаимосвязи нормальных распределений  $t$ ,  $\chi^2$  и  $z$ . Всё это было представлено в работе [112] и в расширенном варианте в монографии "Statistical Methods for Research Workers" (1925; рус. пер. [83], которая способствовала обучению исследователей методам, основанным на этой системе.

<sup>3</sup> Нейман, Ежи (Jerzy Sława-Neyman, при рождении Юрий Чеславович Нейман [Бендеры, Российская империя]; 1894-1981) – польский, американский математик, статистик; член Национальной академии наук США (1963), иностранный член Лондонского королевского общества (1979).

Пирсон, Эгон Пирсон (Egon Sharpe Pearson; 1895-1980) – британский статистик; президент Королевского статистического общества (1955-1956), член Лондонского королевского общества (1966).

<sup>4</sup> Фишер обвинил Пирсона в серьезной ошибке, которую тот так и не признал (Пирсон принимал рассчитанный по выборке из генеральной совокупности коэффициент корреляции за оценку истинного значения, что, как доказывал Фишер, было неправильно). В дальнейшем Фишер не раз вступал в конфликты с коллегами – с Э. Пирсоном, Е. Нейманом и др.; позднее он жаловался, что его статьи рецензируют статистики, не разбирающиеся в биологии, и биологи, не знакомые со статистикой [69, с. 82, 109]. Биографы всегда отмечали его сложный, противоречивый характер, временами проявлявшийся в грубых высказываниях или письмах, адресованным коллегам, что не раз приводило к ссорам с последними (см. [69, 97]).

<sup>5</sup> Рассел, Джон (Sir Edward John Russell; 1872-1965) – британский химик, почвовед; директор Экспериментальной станции Ротамстед (Rothamsted Experimental Station; 1912-1943).

<sup>6</sup> Байесовская вероятность противопоставляется частотной, в которой вероятность определяется относительной частотой появления случайного события при достаточно длительных наблюдениях. Разница между байесовской и частотной интерпретацией играет важную роль в практической статистике. Например, при сравнении двух гипотез на одних и тех же данных, теория проверки статистических гипотез, основанная на частотной интерпретации, позволяет отвергать или не отвергать модели-гипотезы. При этом адекватная модель может быть отвергнута из-за того, что на этих данных кажется адекватнее другая модель. Байесовские методы, напротив, в зависимости от входных данных «выдают» апостериорную вероятность быть адекватной для каждой из моделей-гипотез [19]. Байесовская теория используется как метод адаптации существующих вероятностей к вновь полученным экспериментальным данным.

Там же, на Ротамстедской станции в ходе анализа огромного количества данных об урожае, накопленных с 1842 г., он готовит и публикует серию статей [108, 113, 126] – это первое применение дисперсионного анализа (ANOVA). В 1928 г. начинается 3-х летняя стажировка у Фишера Дж. Ирвина<sup>7</sup>, который стал одним из первых, кто освоил эти методы и начал широко их пропагандировать и внедрять в биологические и медицинские исследования.

Генетические исследования Фишера на Ротамстедской станции были сосредоточены на эволюции, на интеграции менделевской теории Его первая важная теоретическая статья в этом направлении – «О коэффициенте доминирования» [111]. В 1930 г. он обобщил свои генетические и эволюционные представления в монографии "The Genetical Theory of Natural Selection" [85]. Именно в этой монографии Фишер формулирует *основную (фундаментальную) теорему естественного отбора*: скорость повышения приспособленности в любой популяции в любой момент равна её генетической дисперсии по приспособленности в этот момент. Правда, «эта теорема строго приложима только к варьированию за счет аллелей одного локуса и только при определенных условиях среды. Тем не менее понятно, что чем больше изменчивых генов и чем больше аллелей каждого гена существует, тем больше шансов для изменения частоты одних аллелей за счет других при отборе» [26, с. 467].

В 1929 г. Р. Фишер был избран членом Лондонского королевского общества.

В 1933 г. Фишер все-таки сменил К. Пирсона на должности профессора евгеники и руководителя лаборатории Гальтона в Университетском колледже в Лондоне. Фишер всегда больше восхищался Ф. Гальтоном, чем его учеником К. Пирсоном. И в дальнейшем его взаимоотношения с Э. Пирсоном, Е. Нейманом и их сотрудниками продолжали ухудшаться [69]. В эти годы выходит его книга "Планирование экспериментов – The Design of Experiments" [115], подготовленная еще на Экспериментальной станции Ротамстед, ставшая фундаментальной в этой области. Кроме того, в книге была представлена *концепция нулевой гипотезы* (принимается по умолчанию предположение о том, что не существует связи между двумя наблюдаемыми событиями, феноменами; нулевая гипотеза считается верной, пока нельзя до-

казать обратное) в контексте ставшего классическим эксперимента с дегустацией чая [84]. Небольшая подборка цитат из книги Фишера [115] дает представление о том, как «видел» эту очень важную для статистики гипотезу:

- «нулевая гипотеза никогда не доказывается и не подтверждается, но, возможно, опровергается в ходе экспериментов. Можно сказать, что каждый эксперимент существует только для того, чтобы факты могли опровергнуть нулевую гипотезу» [154, p. 19].
- «нулевая гипотеза должна быть точной, свободной от неопределенности и двусмысленности, потому что она должна служить основой для "проблемы распределения – problem of distribution", решением которой является критерий значимости» [115, p. 19].
- «Однако мы можем выбрать любую нулевую гипотезу, если пожелаем, при условии, что она точна» [115, p. 20].
- «Для этой цели, предложенный новый тест (*дама, дегустирующая чай*. – Г.Р.), был бы совершенно неуместным (inappropriate), и ни у одного экспериментатора не было бы соблазна использовать его. Однако математически он так же действителен, как и любой другой, в том смысле, что при правильной рандомизации очевидно, что он даст значимый результат с известной вероятностью, если нулевая гипотеза была верна» [15, p. 191-192].
- «Идея о том, что разные тесты значимости подходят для тестирования разных признаков одной и той же нулевой гипотезы, не представляет трудностей для работников, занятых практическими экспериментами, но стала поводом для многих теоретических дискуссий среди статистиков» [115, p. 194].
- «Этот вопрос (*выбор подходящего экспериментального измерения и нулевой гипотезы*. – Г.Р.), когда ответ на него еще не известен, плодотворно обсуждаем, только тогда, когда экспериментатор имеет в виду не одну нулевую гипотезу, а класс таких гипотез, в значении отклонений от каждой из которых он является одинаково заинтересованным» [115, p. 194].

В 1943 г. Фишер вернулся в Кембридж в качестве профессора и заведующего кафедрой генетики. Его *теория инбридинга* дала теоретическую основу экспериментам на мышах, которые он проводил еще с «лондонских времен» [117, 124, 127]. Он публикует целый спектр генетико-популяционных исследований – о коло-

<sup>7</sup> Ирвин, Джозеф (Joseph Oscar Irwin; 1898-1982) – британский статистик, который продвинул использование статистических методов в биологии и медицине.

нии молей (*Panaxia dominula* L. [124]), чешуекрылых (*Maniola jurtina* L., популяции дербенника иволистного (*Lythrum salicaria* L. [116]) и др.

Интересен и такой штрих. В конце 1948 г., после печально знаменитой «августовской сессии ВАСХНИЛ» Р. Фишер публикует короткую, всего две страницы, заметку «Что это за человек Лысенко?» [116]. Он внимательно ознакомился с брошюрой, в которой была представлена речь Т.Д. Лысенко на этой сессии [140] «Я хотел бы процитировать некоторые отрывки, которые повлияли на меня; они также помогают человеку понять его особый жаргон» [116 р. 874]. Не буду останавливаться на деталях аргументированной критики (тем более, что текст заметки есть в Интернете; см. [<https://web.archive.org/web/20090316173453/http://digital.library.adelaide.edu.au/coll/special/fisher/229.pdf>]). Приведу лишь завершающие её слова [116, р. 875]: «Нет, я не могу поверить в свете этой речи, что награда за триумфальную карьеру Лысенко, – это прогресс научных знаний или процветание бедных крестьян. Вознаграждение, которое он так охотно получает, – это Сила, сила для себя, сила угрожать, сила пытаться, сила убивать». Весьма пронизательно.

В конце 30-х годов, в США продолжал работу над концепцией *доверительного интервала* Е. Нейман и начал формироваться новый статистический подход – *последовательный анализ*, лидером которого становится А. Вальд<sup>8</sup> [150, 151]. В отличие от классических методов математической статистики, в которых число производимых экспериментов фиксируется заранее, методы последовательного анализа характеризуются тем, что момент прекращения наблюдения является случайным и определяется наблюдателем в зависимости от значения наблюдаемых данных.

Выше я уже говорил, ссылаясь на биографов, что в научном общении Фишер был сложным человеком, всегда готовым вступить в спор; его дочь говорила о нем [97]: «Он вырос, не научившись чутко относиться к обычным человеческим качествам братьев». Вот и после трагической гибели Вальда<sup>9</sup>, в 1950 г., Фишер «назначает» его своим очередным оппонентом

<sup>8</sup> Вальд, Абрахам (Abraham Wald; 1902-1950) – венгерский, американский математик, статистик.

<sup>9</sup> Вальд и его жена погибли, когда самолет Air India, в котором они путешествовали, потерпели крушение в горах Нилгири (Гималаи, Непал) во время обширной лекционной поездки по приглашению индийского правительства.

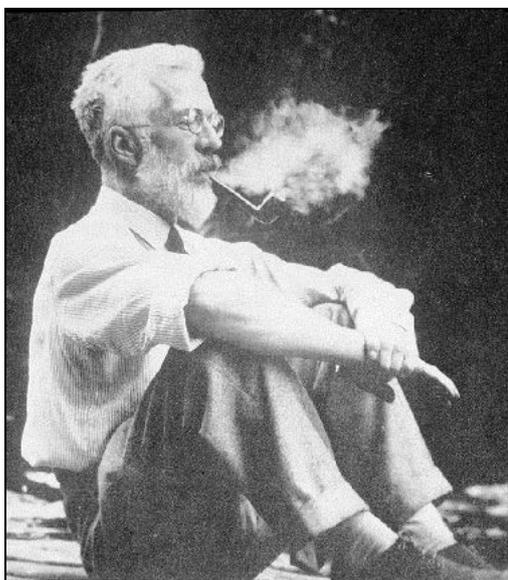
(не будем забывать, что в это же время он серьезно конфликтовал с Э. Пирсоном и Е. Нейманом; и это были лишь некоторые из его «противников»). Фишер «напал» на Вальда за то, что он считал его математиком без научного опыта, который написал некомпетентную книгу по статистике. Фишер особенно критиковал работу Вальда по планированию экспериментов, ссылаясь на незнание основных идей предмета, изложенных ранее им самим и Ф. Йейтсом<sup>10</sup> [128]. Нейман защищал Вальда как мог. Позднее, когда Нейман стал одним из ведущих статистиков в США, их личный спор перерос в спор «англо-американский» (не напоминает «московско-петербургские» противостояния по разным поводам?..). Фишеру никогда не нравилась идея сведения статистических выводов к принятию решений с помощью функций потерь (именно в этом направлении и работал Вальд); главное противоречие выразилось в том, что в Америке (в соответствии с прагматизмом Ч. Пирса<sup>11</sup>) не важно, о чем вы думаете, а важно, что вы делаете; в Англии же – наоборот [72]. Свою критику американских статистиков Фишер изложил в статье [118] и в последней монографии «Статистические методы и научный вывод» [119]. Например, вот так Фишер пишет о статистических процедурах принятия решений: «Этот акцент (*речь идет о процедурах принятия решений*. – Г.Р.) необходим в первую очередь потому, что потребности и цели работников экспериментальных наук бывают плохо поняты и искажены. Разумеется, можно предположить, что такие авторы, как Нейман и Вальд, которые рассматривали эти тесты, не обращая внимания на

<sup>10</sup> Йейтс, Фрэнк (Frank Yates; 1902-1994) – британский математик, статистик; с 1931 г. был помощником Фишера на Ротамстедской станции, а с 1933 г., после ухода Фишера возглавил отдел статистики; член Королевского статистического общества. Интересен такой факт-анекдот. Историограф королевского двора адресовал Йейтсу следующий вопрос: «Не является ли пятница роковым днем для Генрихов английских?» (в преамбуле запроса сообщалось, что короли Генрихи, принадлежавшие четырем различным правящим династиям, непременно умирали по пятницам, приводились точные даты смерти восьми английских королей). Ответ Йейтса вошел в историю науки: «Дорогой сэръ! Представленные Вами статистические данные не противоречат сформулированной Вами статистической гипотезе. Королевское статистическое общество рекомендует Вам продолжать наблюдения».

<sup>11</sup> Пирс, Чарльз (Charles Sanders Peirce; 1839-1914) – американский философ, логик, математик, основоположник прагматизма и семиотики; член Национальной АН США (1877).

их назначение в естественных науках, возможно поэтому [они] и не добились большего успеха в применении своих идей к нуждам процедур принятия решений» [119, р. 76].

Во время Второй мировой войны, в 1942 г. во время полета над Сицилией погиб его старший сын Джордж (George, 22 года), а в 1943 г. брак распался [46, 97]. Около пяти лет Фишер испытывал трудности с работой, поскольку евгеника подвергалась гонениям после нацистских опытов над людьми, но, в конце концов, он получил признание и много наград за свою работу, был удостоен рыцарского звания указом королевы Елизаветы II (Queen Elizabeth II; в первый же год своего правления) в 1952 г. С 1952 по 1954 гг. Фишер был президентом Королевского статистического общества (кстати, сменил он на этом посту Б. Хилла, с которым позднее дискутировал по корректности связи «курение – рак».



1956

Выйдя официально на пенсию (в 1957 г.), Фишер еще два года оставался в Англии. В этот период он вновь привлек к себе внимание, когда в 1957-59 гг. оспорил выводы Р. Долла и Б. Хилла<sup>12</sup> [101, 102] и статистически доказал отсутствие связи между курением и раком легких [120-123], что вызвало недоумение обществу [69, с. 128, 155-158]. Правда, следу-

<sup>12</sup> Долл, Ричард (Sir William Richard Shaboe Doll; 1912-2005) – британский врач, эпидемиолог; член Лондонского королевского общества (1966);

Хилл, Брэдфорд (Sir Austin Bradford Hill; 1897-1991) – британский эпидемиолог, статистик; президент Королевского статистического общества (1950-1952), член Лондонского королевского общества (1954); считается основоположником медицинской статистики.

ет понять, что Фишер отрицал не возможность того, что курение является причиной рака, а лишь ту уверенность, с какой поборники общественного здоровья провозглашали этот вывод.

В 1959 г. по приглашению статистика Э. Корниша<sup>13</sup> и генетика Дж. Беннетта<sup>14</sup> он переехал жить в Аделаиду (Южная Австралия; Adelaide, Australia) и стал работать в качестве научного сотрудника в Организации научных и промышленных исследований Содружества (CSIRO – Commonwealth Scientific and Industrial Research Organisation). Он умер в Аделаиде после операции по удалению раковой опухоли толстой кишки (но не лёгких...) 29 июля 1962 г.; его прах лежит в кафедральном соборе Святого Петра (St. Peter's Cathedral, Adelaide; [139]).

Фишер был человеком глубоко и искренне религиозным, патриотом, консерватором по политическим убеждениям и рационалистом в науке. Кроме того, Фишер был великолепным собеседником, а в поздние годы воплощал собой тип «рассеянного профессора», благодаря некоторой небрежности в манерах и одежде (<https://cumir.ru/ronald-ejlmr-fisher>).

50-летие со дня смерти Фишера было ознаменовано совместным заседанием в Королевском статистическом обществе при участии Британского общества истории математики и Британского общества по истории науки (A 50<sup>th</sup> Anniversary Commemoration of R.A. Fisher [7 June 2012]; [http://www.economics.soton.ac.uk/staff/aldrich/fisherguide/RSS\\_R%20A%20\\_Fisher\\_meeting.pdf](http://www.economics.soton.ac.uk/staff/aldrich/fisherguide/RSS_R%20A%20_Fisher_meeting.pdf)).

После его смерти семья пожертвовала Университету Аделаиды авторские права на издание книг, другие интеллектуальные и личные материалы Р. Фишера; это способствовало переизданию его основных монографий и научной переписки (профессор Дж. Беннетт выступил редактором 5-томного издания сборника статей Фишера [1971-1974 гг., тираж 1000 экз.]; он же отредактировал сборники с подборками из его переписки; все это в 1981 г. поступило в Библиотеку Барра Смита [главная библиотека Университета Аделаиды]) и позволило создать веб-сайт, с которого можно загрузить большинство его опубликованных работ (<https://www.adelaide.edu.au/library/special/mss/fisher/>).

## ДИСПЕРСИОННЫЙ АНАЛИЗ

Большинство методов Фишера имеют общий характер и применяются в естественных науках, в экономике и в других областях дея-

<sup>13</sup> Корниш, Эдмунд (Edmund Alfred Cornish; 1909-1973) – австралийский математик, статистик.

<sup>14</sup> Беннетт, Джон Генри (John Henry Bennett; 1926-2015) – австралийский генетик.

тельности (в том числе, и в экологии). То, что статистик в наши дни может, по образному выражению Дж. Тьюки<sup>15</sup>, «играть в любом саду», – в немалой степени заслуга Фишера (цит. по: [69, с. 159]).

В монографии "Statistical Methods for Research Workers" (1925; рус. пер. [121]) и в ряде более ранних статей [108, 113, 126, 120] предложил новую статистическую методику, известную во всем мире как дисперсионный анализ<sup>16</sup> или по английской аббревиатуре ANOVA (англ. ANalysis Of VAriance). На языке теории вероятностей основная задача дисперсионного анализа может быть сформулирована следующим образом: по результатам наблюдений некоторой случайной величины оценить зависимость её математического ожидания и дисперсии (или ковариационной матрицы в случае векторной случайной величины) от выбранных категориальных (неметрических) переменных – факторов. Фактически, это метод, направленный на поиск зависимостей в экспериментальных данных путём исследования значимости различий в средних значениях.

В отличие от *t-критерия Стьюдента*<sup>17</sup>, позволяет сравнивать средние значения трёх и более групп. Суть дисперсионного анализа сводится к изучению влияния одной или нескольких независимых переменных, обычно именуемых факторами, на зависимую переменную. Зависимые переменные представлены значениями абсолютных шкал (шкала отношений). Независимые переменные являются номинативными (шкала наименований), то есть отражают групповую принадлежность, и могут иметь два

<sup>15</sup> Тьюки, Джон (John Wilder Tukey; 1915-2000) – американский математик; член Национальной АН США (1961), иностранный член Лондонского королевского общества (1991).

<sup>16</sup> Может показаться странным, что процедура сравнения средних называется дисперсионным анализом. В действительности, это связано с тем, что при исследовании статистической значимости различия между средними двух (или нескольких) групп, мы на самом деле сравниваем (т.е. анализируем) выборочные дисперсии. Возможно, более естественным был бы термин «анализ суммы квадратов» или «анализ вариации», но в силу традиции употребляется термин *дисперсионный анализ* [<http://statsoft.ru/home/textbook/modules/stanman.html#basic>].

<sup>17</sup> Госсет, Уильям (William Sealy Gosset; 1876-1937) – британский статистик, более известный под псевдонимом Стьюдент (Student); сотрудник пивоваренного завода Arthur Guinness Son & Co в Дублине. Наиболее частые случаи применения *t-критерия* связаны с проверкой равенства средних значений в двух выборках.

или более значения (типа, градации или уровня).

Публикаций по ANOVA, даже монографического плана, – огромное количество; назову лишь некоторые из них, которые в той или иной степени использованы при написании этого, весьма конспективного раздела [24, 35, 37, 39, 48, 49, 51-53, 58, 61, 65, 73, 87, 148].

В зависимости от типа и количества переменных (для качественных и количественных) различают:

- однофакторный и многофакторный дисперсионный анализ (одна или несколько независимых переменных);
- одномерный и многомерный дисперсионный анализ (одна или несколько зависимых переменных);
- дисперсионный анализ с повторными измерениями (для зависимых выборок);
- дисперсионный анализ с постоянными факторами, случайными факторами, и смешанные модели с факторами обоих типов.

Исходными положениями для применения дисперсионного анализа являются [22, 87]:

- нормальное распределение значений изучаемого признака в генеральной совокупности;
- равенство дисперсий в сравниваемых генеральных совокупностях;
- случайный и независимый характер выборки.

Нулевой гипотезой в дисперсионном анализе является утверждение о равенстве средних значений.

Так как дисперсионный анализ служит для количественного исследования влияния входных переменных (факторов) на наблюдаемую переменную (результат опытов, результативный признак), то в зависимости от характера выборки (изложенным выше исходным положениям для применения дисперсионного анализа), кроме классического критерия Фишера, рекомендуется [22, с. 448-514) использовать другие критерии и поправки:

| Выборки  | Критерии  |
|--|---|
| Независимые выборки, необязательно равные дисперсии, приближенно одинаковые виды распределения | <i>H</i> -критерий Краскела – Валлиса, критерий Фридмана (2-х факторный дисперсионный анализ) |
| Равенство числа наблюдений в градациях фактора   | критерий Хартли   |
| Для выборок разного объема   | критерий Кокрена  |
| Однородность дисперсий   | критерий Бартлета   |
| Нормальное распределение, равенство дисперсий и  | «Быстрые критерии»:   |

|   |   |
|---|---|
| равные объемы отдельных выборочных групп    | критерий Линка – Уоллеса, критерий Немени |
| Множественное сравнение независимых выборок | критерий Уилкоксона – Уилкокса            |
| При множественных сравнениях                | поправка Бонферрони                       |

Естественно, эта таблица может быть продолжена. Но вернемся к ANOVA Фишера. Очень важным этапом дисперсионного анализа является начальный, подготовка дисперсионного комплекса – определение действующих факторов (один, два, три и больше), числа градаций, количественные или качественные признаки, равномерные или неравномерные комплексы и пр. Фактически, метод ANOVA основан на разложении общей дисперсии дисперсионного комплекса на составляющие её компоненты, сравнивая которые с помощью  $F$ -критерия, можно определить, какая доля варьирования может быть связана с организованными в комплекс данными, а какая является случайной, неорганизованной. «Так, если регулируемый фактор (например, доза удобрений) оказывает существенное влияние на результативный признак (урожай культуры), оно непременно скажется на величине групповых средних, которые будут заметно отличаться друг от друга. Таким образом, здесь происходит варьирование групповых средних, причиной которого является влияние регулируемого фактора» [35, с. 156].

Я не буду останавливаться на конкретных алгоритмах, реализующих ANOVA, так как сегодня существует множество пакетов статистических программ, способствующих эффективному использованию дисперсионного анализа (в том числе и в экологии); назову лишь широко распространенные пакеты STATISTICA 6.0 (см., например, [8]) или IBM SPSS Statistics 22 [11]. Однако отмечу несколько сравнительно новых (и, пока, не очень широко распространенных) подходов к проведению дисперсионного анализа.

**Прямой градиентный анализ.** Учитывая непреходящую важность прямого градиентного анализа, который широко используется не только в биоиндикационных, но и в ординационных исследованиях, напомним его как «классические», так и некоторые «современные» составляющие [63, 65].

В начале 50-х гг. XX в. Р. Уиттекер [79] предложил «интуитивно-статистический» метод прямого градиентного анализа применительно к растительности горных систем, сущность которого сводилась к следующему. Горная система представлялась в двухмерной схеме, где по одной оси откладывался комплексный градиент

высоты над уровнем моря (гидротермический градиент), а по другой – топографический (различные экспозиции склона). В этих осях координат строились распределения отдельных популяций растений и некоторые характеристики видового разнообразия и продуктивности, что позволило подтвердить *концепцию континуума Раменского – Глисона*. Следует заметить, что эти пионерские работы по прямому градиентному анализу встретили достаточно острую критику (см. обзор: [41]), которая в первую очередь касалась произвольности (субъективности) выбора осей ординации. Уиттекер [152] убедительно показал, что, «субъективно» выбирая эти оси, исследователь опирается на значительный багаж знаний об экологии видов и осуществляет ординацию по бесспорно ведущим градиентам.

Другой недостаток метода Уиттекера (а именно его интуитивно-статистический характер) был исправлен в количественной модификации *прямого градиентного анализа* [40, 42, 43]. Построение графиков изменения наблюдаемого признака (вес, встречаемость, обилие и пр.) для отдельных видов в зависимости от изменения выбранного для ординации видов фактора сопровождается расчетом некоторых статистических характеристик:

- *средневзвешенная напряженность фактора:*

$$\bar{X}_i = \sum_{j=1}^r p_{ij} X_j ;$$

- *средневзвешенная дисперсия:*

$$\bar{D}_i = \sum_{j=1}^r p_{ij} (X_j - \bar{X}_i)^2 ,$$

где  $X_j$  – середина  $j$ -ой градации исследуемого фактора;  $r$  – число этих градаций,

$$p_{ij} = \frac{m_{ij}}{n_j \sum_k (m_{ik}/n_k)} , \quad \sum_{j=1}^r n_j = N ,$$

где  $n_j$  – число наблюдений в  $j$ -й градации фактора;  $m_{ij}$  – число встреч вида  $i$  в этой градации,  $N$  – общее число наблюдений;

- кроме того, рассчитывается *сила влияния исследуемого фактора на данный вид*, получаемая из однофакторного дисперсионного анализа ( $\eta^2$ ).

С помощью средневзвешенных значений фактора каждый вид «привязывается» к определенному отрезку градиента, а малое значение  $D_i$  и большое  $\eta^2$  свидетельствуют о высокой индикаторной роли вида  $i$ , причем, преимущество следует отдавать оценке по значениям  $D_i$ , (Миркин, 1971) – чем меньше  $D_i$ , тем вид занимает более узкий участок по оси данного фактора и, следовательно, является хорошим его индикатором.

В 60-70-х гг. прошлого столетия прошла дискуссия по вопросу об оценке показателя си-

лы влияния фактора  $\eta^2$  при дисперсионном анализе. Сравнению подвергались в основном три показателя – *квадрат корреляционного отношения Пирсона* ( $P$ ;  $\eta^2$ ), рекомендованный для биологических исследований Н.А. Плохинским<sup>18</sup> [51, 52], *показатель Миллса – Лукомского*<sup>19</sup> (ML; [36, 38]), пропагандируемый В.Ю. Урбахом<sup>20</sup> [80], и *коэффициент Снедекора* ( $S$ ), отстаиваемый Э.Х. Гинзбургом<sup>21</sup> [14, 15]. Аргументы «за» и «против» в этой дискуссии были подробно проанализированы [62, 67, 68]. Признавая состоятельность всех трех оценок силы влияния фактора в дисперсионном анализе, участники дискуссии стремились в большей степени доказать «неработоспособность» противного подхода и часто «уходили» от ответов на вопрос о той или иной особенности «своего» показателя. Еще одной отличительной особенностью этой дискуссии было привлечение в качестве основного аргумента моделирования различных случайных дисперсионных комплексов на ЭВМ вместо попытки строгого аналитического анализа предлагаемых показателей.

Нами было доказано [67], что показатели Плохинского и Миллса–Лукомского полностью эквивалентны, так как оказываются связанными линейной зависимостью – их различие состоит только в границах изменений:  $P$  меняется от 0 до +1, а ML – от  $-(r-1)/(N-r)$  также до +1. Отсюда следует, что показатель ML неизбежно должен давать отрицательные значения в тех случаях, когда  $P < (r-1)/(N-r)$ . Иными словами, показатель ML полностью дублирует  $P$ ; границы изменения последнего от 0 до +1 делают его предпочтительнее в силу удобства. Преимуществом показателя  $S$  является разработанная Гинзбургом оптимальная схема планирования структуры дисперсионного комплекса. При малых  $n$  величина и достоверность показателей существенно зависят от структуры исследуемого материала, что обуславливает необходимость вообще осторожно использовать любые биометрические методы;

<sup>18</sup> Плохинский, Николай Александрович (1899-1988) – зоолог, генетик, статистик.

<sup>19</sup> Миллс, Фредерик (Frederick [Fred] Cecil Mills; 1892-1964) – американский математик, статистик, экономист;

Лукомский, Яков Ильич (1906-1961) – статистик, экономист.

<sup>20</sup> Урбах, Виктор Юльевич (1929-1975 ?) – отечественный математик, статистик.

<sup>21</sup> Гинзбург, Эмиль Хаимович (1936-2003) – отечественный, израильский генетик, биофизик.

для больших  $n$  все показатели практически близки.

При расчете силы влияния фактора необходимо учитывать следующий факт. В эксперименте [12, с. 8], трудно добиться «изменения одного фактора при постоянстве других». Таким образом, более корректным представляется использование 2-х, 3-факторных и т. д. комплексов, которые построить на практике довольно сложно. Поэтому некоторые авторы при оценке  $P$  ( $\eta^2$ ) прибегают к ряду эвристических процедур. Один из способов состоит в сглаживании влияния неконтролируемых факторов на изучаемый фактор [4, 59, 60, 68].

Пусть нам задан однофакторный равномерный дисперсионный комплекс для количественных признаков. Количество градаций одного комплекса равно  $r$ , а число наблюдений в каждой градации –  $n$ . Анализ дисперсионного комплекса начинается с рассмотрения ряда частных средних для каждой градации фактора:

$$\bar{X}_j = (\sum_{i=1}^n X_{ij})/n,$$

где  $X_{ij}$  – значение исследуемого признака при  $i$ -м наблюдении в  $j$ -й градации ( $j = \overline{1, r}$ ).

Ряд частных средних может быть задан линейной регрессии. Следует отметить, что «течение» эмпирической линии регрессии почти никогда не бывает плавным, так как большое влияние на изменение частных средних оказывают всевозможные «шумы», в роли которых выступают как случайные (неучтенные) факторы, так и тесно скоррелированные с изучаемым. Таким образом, кривые встречаемости признака в зависимости от изучаемого фактора оказываются многовершинными.

Наблюдаемая многовершинность противоречит представлениям о существовании только одного экологического оптимума распределения вида по градиенту среды (см., например, [79]) и резко завышает значение силы влияния фактора  $P$  ( $\eta^2$ ) за счет увеличения дисперсии, объясняющей влияние фактора ( $C_x$ ). Наличие многовершинности для кривой встречаемости признака можно объяснить двумя причинами: либо большим числом градаций фактора, либо большей степенью зависимости изучаемого фактора от неучтенных показателей. Замечу, что это понимал и П. Грейг-Смит [(18, с. 182): «С этой трудностью (*распределение значений фактора среды для всей совокупности описаний значительно отличается от нормального.* – Г.Р.) пришлось бы столкнуться, если бы кривая значений фактора имела бы два отчетливо выраженных максимума, что указывало бы на воз-

можное наличие в выборке образцов из двух совершенно разных местообитаний».

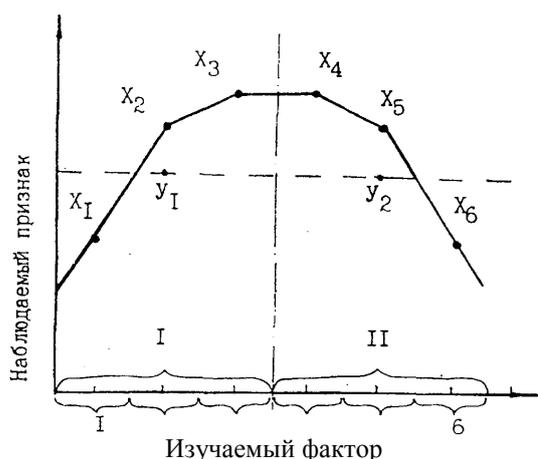


Рис. 1. Теоретическая модель зависимости частных средних от изучаемого фактора.

В первом случае путем укрупнения градаций фактора можно добиться одновершинной кривой частных средних (действительно, для двух градаций фактора кривая частных средних всегда будет одновершинной, а для трех – в восьми случаях из девяти). Однако слишком крупные градации фактора не позволяют вскрыть его «тонкое» влияние, тем самым огрубляя и заметно понижая значение  $R$  и соответственно снижая общую эффективность дисперсионного анализа. На рис. 1 показана некоторая теоретическая модель зависимости частных средних от изучаемого фактора, разбитого на шесть градаций. Если теперь представить, что фактор разбит на две градации, то нетрудно убедиться, что для этого случая значения частных средних совпадут с общей средней и, следовательно,  $C_x = 0$ , а отсюда и  $R = 0$ , что заведомо неверно.

В случае зависимости изучаемого фактора от других, применяя схему многофакторного дисперсионного анализа, можно представить общую дисперсию как сумму дисперсий, вызванных воздействием отдельных составляющих и их сочетаний. Однако если эксперимент не планировался для многофакторного дисперсионного анализа, то отсутствуют исходные данные для построения дисперсионного комплекса и, следовательно, отсутствует возможность воспользоваться результатами этого анализа.

Одним из простейших методов, позволяющих элиминировать случайные колебания эмпирической линии регрессии, вскрыть зависимость признака от изучаемого фактора, так сказать, в чистом виде и получить плавную и одновершинную теоретическую кривую, является способ выравнивания *методом скользящей*

*средней*. Будем определять средние после выравнивания по следующим формулам:

$$\bar{\bar{X}}(i) = [\bar{X}(i-1) + \bar{X}(i) + \bar{X}(i+1)]/3$$

для всех  $1 < i < r$ ; крайние значения ряда выровненных частных средних определим следующим образом:

$$\bar{\bar{X}}(1) = [2\bar{X}(1) + \bar{X}(2)]/3 \quad \text{и}$$

$$\bar{\bar{X}}(r) = [\bar{X}(r-1) + 2\bar{X}(r)]/3,$$

где  $\bar{X}(i)$  – частные средние исходного дисперсионного комплекса,  $\bar{\bar{X}}(i)$  – выровненные значения частных средних.

Эти же данные можно представить как некоторый новый теоретический однофакторный дисперсионный комплекс, где в качестве наблюдений, соответствующих  $i$ -й градации фактора, выступают частные средние исходного комплекса (по три значения, приведенных выше, для каждой градации). Легко убедиться, что общие средние исходного и теоретического дисперсионного комплексов равны между собой. Из этого факта легко получается следующая зависимость между дисперсиями:

$$C_x = n\bar{C}_y/3,$$

где  $\bar{C}_y$  – общая дисперсия теоретического комплекса. После преобразований общая дисперсия исходного комплекса  $C_y$  представляется в следующем виде (зависимость между дисперсиями для безграничных связанных рядов в общем случае получена Е.Е. Слуцким<sup>22</sup> [74]):

$$C_x = n\bar{C}_x/3 + \bar{C}_z,$$

где  $n\bar{C}_x/3$  и  $\bar{C}_z$  – это, соответственно, дисперсии, связанные с влиянием изучаемого фактора после выравнивания ряда частных средних и воздействием случайных и неучтенных факторов. Тогда сила влияния фактора после выравнивания определяется по следующей формуле:

$$R = \eta_x^2 = tC_x/C_y,$$

где  $t$  – коэффициент, прямо пропорциональный числу наблюдений в каждой градации комплекса, являющегося исходным для выравнивания ( $n$ ), и обратно пропорциональный количеству частных средних, по которым производится выравнивание (для данного частного случая,  $t = n/3$ ).

Выравнивание способом скользящей средней можно повторять многократно. Последним следует считать тот шаг  $h$ , после которого кривая частных средних по градиентам фактора становится достоверно одновершинной, что

<sup>22</sup> Слуцкий, Евгений Евгеньевич (1880-1948) – отечественный математик, статистик, экономист.

проверяется сравнением средних, например, по критерию Стьюдента. При количестве выравниваний  $h > 1$  значение  $t = 1$ , так как исходным для получения теоретического комплекса  $h$  является комплекс  $(h - 1)$ , для которого число наблюдений в каждой градации равно количеству частных средних, по которым проводится выравнивание.

Получение на практике равномерных комплексов возможно лишь в немногих случаях, когда градации фактора можно спланировать до проведения эксперимента (например, пространственные градиенты). Для количественных неравномерных комплексов (чаще всего это экоклина) значительно усложняется процесс нахождения коэффициента  $t$ , так как он является функцией от числа наблюдений. В качестве оценки величины  $n$  можно использовать среднее значение числа наблюдений в каждой градации фактора.

**Двухфакторный дисперсионный анализ с одним наблюдением в ячейке.** Этот вариант ANOVA рассмотрен в статье А.Г. Боголюбова<sup>23</sup> [6]. Автор продемонстрировал возможности дисперсионного анализа «пусть однократных, но обязательно должным образом организованных данных» [6, с. 84].

Пусть имеем два контролируемых фактора –  $N$  (число градаций  $I$ ) и  $R$  (число градаций  $J$ ). Варианты опыта для каждого сочетания факторов обозначим  $(i, j)$ ; при каждом сочетании условий эксперимента измерение проводится один раз.

Обозначим через  $y_{**}$  оценку генерального среднего:

$$y_{**} = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J y_{ij}.$$

Если факторы не влияют на  $y$  в каждом варианте  $(i, j)$  опыта, то все отклонения величин  $y_{ij}$  от генерального среднего случайны:

$$y_{ij} = y_{**} + e_{ij},$$

где  $e_{ij}$  – независимые нормально распределенные случайные величины, имеющие нулевое математическое ожидание и одинаковую дисперсию (ошибка воздействия неучтенных и случайных факторов). Получим средние величины при фиксированном уровне одного из факторов:

$$y_{i*} = \frac{1}{J} \sum_{j=1}^J y_{ij}, \quad y_{*j} = \frac{1}{I} \sum_{i=1}^I y_{ij}.$$

Главный эффект воздействия первого фактора ( $I$ ) – это величина отклонения среднего  $y_{i*}$  от генерального среднего ( $\alpha_i = y_{i*} - y_{**}$ ); аналогич-

но, для фактора  $J$  – ( $\beta_j = y_{*j} - y_{**}$ ). После этого проверяются следующие две нулевые гипотезы:

$$H_N: \text{все } \alpha_i = 0; \quad H_R: \text{все } \beta_j = 0.$$

Далее, процедура оценок дисперсий – традиционна. Полная сумма квадратов отклонений каждой из  $y_{ij}$  от генерального среднего представляется в виде суммы квадратов эффектов каждого фактора:

$$SS_{N \times R} = SS_N + SS_R + SS_e,$$

где

$$SS_N = \sum_{i=1}^I \sum_{j=1}^J \alpha_i^2,$$

$$SS_R = \sum_{i=1}^I \sum_{j=1}^J \beta_j^2,$$

$$SS_e = \sum_{i=1}^I \sum_{j=1}^J (y_{ij} - y_{i*} - y_{*j} + y_{**}).$$

Число степеней свободы для воздействующих факторов на единицу меньше числа градаций ( $\nu_N = I - 1$ ;  $\nu_R = J - 1$ ), а сумм квадратов ошибки –  $\nu_e = \nu_N \cdot \nu_R$ . Вычисляем статистики *распределения Фишера* ( $F_N = \nu_R SS_N / SS_e$  и

$F_R = \nu_N SS_R / SS_e$ ); далее, все просто: гипотезы  $H_N$  и  $H_R$  отвергаются на уровне значимости  $\alpha$  (верхний предел  $F$ -распределения), если  $F_N > F_{\alpha, \nu_N, \nu_e}$  и  $F_R > F_{\alpha, \nu_R, \nu_e}$ .

Также подробно А.Г. Боголюбов анализирует взаимодействие в двухфакторном анализе с одним наблюдением в ячейке.

Представляют интерес такие размышления автора (приведу достаточно обширные цитаты): «Практическое использование методов математической статистики редко укладывается в рамки условий, при соблюдении которых они были строго обоснованы. Например, хрестоматийное нарушение – нарушение условий нормальности распределений наблюдаемых случайных величин. Известно, что многие методы нечувствительны к нарушениям некоторых условий их обоснования. В частности, на выводы дисперсионного анализа слабо влияют отклонения распределений случайных величин от нормального. <...> Часто взаимодействия дисперсионного анализа трактуются как реально существующие взаимодействия факторов. При этом забывается, что исследуется изменчивость выбранной характеристики при варьировании факторов, а не факторы как таковые. Поэтому на основе дисперсионного анализа крайне нежелательны заключения о природе факторов или их взаимоотношениях. В наших исследованиях лучше вести речь о неаддитивных воздействиях факторов или неаддитивных эффектах, реже используя термин "взаимодействие". Одновременное действие факторов может оказать-

<sup>23</sup> Боголюбов, Александр Григорьевич (1954-2005) – математик, биофизик, эколог.

ся мультипликативным. В такой ситуации неудивительно отрицание гипотезы аддитивности эффектов. Нелинейность наблюдаемых процессов и синергетические свойства явлений тоже могут приводить к значимым неаддитивным эффектам. Но нельзя рассматривать значимые неаддитивные эффекты как окончательное доказательство, например, нелинейности процессов или синергетических свойств явлений. Как правило, значимые неаддитивные эффекты указывают на необходимость дополнительных исследований. Интерпретируя результаты дисперсионного анализа, полезно помнить о возможном косвенном воздействии контролируемых факторов на изучаемую характеристику» [Боголюбов, 1993, с. 89, 91].

Завершая этот раздел, добавлю, что методы (модели) современного дисперсионного анализа весьма разнообразны – это модели с постоянными, случайными, рандомизированными, смешанными и пр. факторами [24, 37, 49, 50].

### ИЗБРАННЫЕ ПРИМЕРЫ

Несколько примеров использования ANOVA в экологических исследованиях.

**Прямой градиентный анализ влияния пастбищной нагрузки на распределение видов растений (классический однофакторный алгоритм).** На Тургайском плато, на одном из участков Оренбургского заповедника в бессточной впадине оз. Айке (типичный участок зоны типчаково-ковыльных степей) была заложена трансекта длиной 1,2 км [20]. Выделено 4 стадии деградации:

- (1) слабый выпас; характерна довольно мощная дернина, рыхловато-комковатая структура, слабая уплотненность верхнего слоя почвы, высокая встречаемость видов;
- (2) умеренный выпас; подстилка отсутствует, проявляется скелетность почвы, заметно падает встречаемость видов;
- (3) стадия скотосоя; появляется солонцовый горизонт, структура почвы становится пылеватой, верхний слой сухой и плотный; разнообразие видов падает;
- (4) абсолютный скотосой; разрушение почвенного покрова, низкое разнообразие видов.

Проведенный градиентный анализ дал следующие результаты (табл. 1).

Выделенные индикаторные группы видов позволяют (при знании области экстраполяции) индентифицировать степень (стадию) выпаса. Показателем в этом плане рогач песчаный (*Ceratocarpus arenarius* L.), который формирует заросли именно в местах интенсивного выпаса сельско-

хозяйственных животных (выделено цветом в табл. 1).

Таблица 1

Результаты однофакторного ANOVA [20, с. 129]

| Виды                            | $\eta^2$ | Встречаемость видов по грациям фактора пастбищной дигрессии |      |      |      |
|---------------------------------|----------|---|------|------|------|
|                                 |          | (1)   | (2)  | (3)  | (4)  |
| <i>Stipa lessingiana</i>        | 0,99     | 95,0  | 1,0  |      |      |
| <i>Echinops ritro</i>           | 0,94     | 0,5   |      |      |      |
| <i>Phlomis tuberosa</i>         | 0,92     | 14,0  | 2,3  | 0,3  |      |
| <i>Serratula xeranthemoides</i> | 0,60     | 0,4   | 0,1  | 0,4  |      |
| <i>Eremogone micradenia</i>     | 1,00     |   | 0,5  |      |      |
| <i>Crinitaria villosa</i>       | 0,95     |   | 2,5  |      |      |
| <i>Artemisia austriaca</i>      | 0,72     | 2,5   | 19,3 | 23,5 |      |
| <i>Tanacetum achilleifolium</i> | 0,67     | 0,5   | 4,2  | 11,5 |      |
| <i>Ceratocarpus arenarius</i>   | 0,94     |   |      | 71,5 | 97,0 |

**Прямой градиентный анализ влияния влажности на распространение населения жуужелиц (классический однофакторный алгоритм).** Еще один пример классического использования ANOVA в экологических исследованиях: распределение жуужелиц по градиенту влажности в сосновых лесах Мордовского заповедника [82].

Таблица 2

Результаты однофакторного ANOVA [82, с. 201]

| Виды                                 | $\eta^2$ | Доля населения жуужелиц |      |      |
|--------------------------------------|----------|-------------------------|------|------|
|                                      |          | 1                       | 2    | 3    |
| <i>Pterostichus oblongopunctatus</i> | 0,32     | 0,80                    | 0,22 | 0,19 |
| <i>Calathus micropterus</i>          | 0,12     | 0,38                    | 0,28 | 0,05 |
| <i>Carabus hortensis</i>             | 0,10     |                         | 0,14 |      |
| <i>Carabus arcensis</i>              | 0,04     | 0,01                    | 0,12 | 0,12 |
| <i>Carabus glabratus</i>             | 0,04     |                         | 0,06 |      |
| <i>Pterostichus nigrita</i>          | 0,12     |                         |      | 0,17 |
| <i>Pterostichus niger</i>            | 0,02     | 0,04                    | 0,11 | 0,12 |
| <i>Agonum fuliginosum</i>            | 0,05     |                         |      | 0,07 |
| <i>Carabus granulatus</i>            | 0,31     |                         |      | 0,05 |

1 – сосняк ландышевый (мезофильные стадии), 2 – сосняк-черничник, 3 – сосняк сфагновый (гигрофитные местообитания).

Среди почвенной мезофауны в лесных экосистемах одно из ведущих мест по видовому составу и численному обилию занимают жуки семейства жуужелиц. С целью индикации среды и диагностики природных и антропогенных процессов были изучены комплексы жуужелиц

на 30 участках в сосновых лесах на территории Мордовского заповедника (отловлено около 50 тыс. экз. различных видов жуличиц, отработано 18,7 тыс. ловушко-суток). В табл. 2 представлены результаты прямого градиентного анализа распределения доли обилия видов жуличиц по фактору влажности.

Для 16 видов жуличиц определено достоверное (на 90-95% уровне значимости) влияние влажности на их распределение. Выделенные группы видов в целом согласуются с имеющимися в литературе данными и могут служить индикаторами степени увлажнения экотопов.

**Прямой градиентный анализ влияния фактора засоления на распределение видов растений (алгоритм с выравниванием).** В качестве примера процедуры выравнивания рассмотрим распределение проективного покрытия *Puccinellia tenuiflora* и *Plantago canescens* в зависимости от засоления (Намский район Якутии [10, с. 145-149]). Реальное распределение этих видов имеет двухвершинный характер с «провалом» в классах засоленности 2,5-3 и 2-2,5, соответственно. После процедуры однократного выравнивания были получены результаты, представленные в табл. 3.

Можно констатировать, что после выравнивания влияние фактора засоленности заметно снизилось (в 2-3 раза), хотя и осталось достоверным. Это делает более объективным наши представления о характере связи этих видов с данным фактором.

Таблица 3

Сравнение силы влияния фактора засоления на распределение двух видов солончаковых лугов Средней Лены до и после выравнивания

| Вид                           | А        |      | В        |      |
|-------------------------------|----------|------|----------|------|
|                               | $\eta^2$ | F    | $\eta^2$ | F    |
| <i>Puccinellia tenuiflora</i> | 0,40     | 46,7 | 0,17     | 14,3 |
| <i>Plantago canescens</i>     | 0,25     | 23,2 | 0,08     | 6,1  |

А – до выравнивания, В – после выравнивания; F – критерий Фишера (показатель достоверности силы влияния фактора).

Изложенный выше метод позволяет, минуя схему многофакторного анализа, извлечь более объективную информацию о влиянии одного фактора на изучаемый признак. Однако следует отметить, что данный алгоритм является приближенным и может использоваться только тогда, когда по каким-либо причинам невозможно построение многофакторных комплексов. Таким образом, при планировании того или иного эксперимента исследователь должен не

только иметь рабочую гипотезу, но и четко представлять тот математический метод, с помощью которого он будет подтверждать или опровергать ее. Если условия корректности эксперимента, напрямую или косвенно содержащиеся в любом математическом методе, соблюдены, то и использование данного метода можно считать обоснованным.

**Лишеоиндикация экологических условий в борах Приволжской возвышенности (классический 2-х факторный алгоритм).** Изучалось влияние экологических факторов на распределение наземных лишайников в сосняках-беломошниках Приволжской возвышенности (на территории Пензенской области [91]). Одной из задач исследования было выяснение влияния орографического фактора: 35 пробных площадок (50x50 см) было заложено на трансектах от гребня к основанию холма на склонах западной и восточной экспозиции. Четко просматривается 2-х факторная схема классического дисперсионного анализа, результаты которого представлены в табл. 4.

Представленные в табл. 4 результаты показывают, что среднее покрытие *Cladina rangiferina* на восточном склоне существенно отличается от такового на западном; заслуживает внимания также значительная совместная сила влияния обоих факторов (AB). На *Cladina arbuscula* и *Cladonia uncialis* данные факторы влияния не оказывают (покрытие *Cladonia uncialis* на восточном склоне чуть выше и различия менее достоверны).

Таблица 4

Результаты 2-х факторного ANOVA [91, с. 112]

| Виды                       | А     | В | AB   | Среднее покрытие по фактору А (%) |      |
|----------------------------|-------|---|------|-----------------------------------|------|
|                            |       |   |      | Е                                 | W    |
| <i>Cladina rangiferina</i> | 0,19  | - | 0,27 | 44,2                              | 25,0 |
| <i>Cladina arbuscula</i>   | -     | - | -    | 10,5                              | 10,8 |
| <i>Cladonia uncialis</i>   | 0,08* | - | -    | 6,3                               | 2,5  |

\* – уровень достоверности 80%; остальные значения достоверны на уровне 95%; А – экспозиция склона; В – удаление от гребня холма; Е – восточная экспозиция; W – западная экспозиция.

**Реакция свободноживущих инфузорий водохранилищ Средней и Нижней Волги на изменения некоторых факторов среды (классический 3-х факторный алгоритм).** По результатам комплексных гидробиологических

исследований в бассейне Волги (1988-1990 гг.) был сформирован 3-х факторный дисперсионный комплекс [21]: А – водохранилища (три градации [Куйбышевское, Саратовское, Волгоградское]; фактически, градиент «север – юг»), В – сезонность (3 градации [май, июль, октябрь]), С – морфометрия водохранилищ (3 градации [нижний бьеф, средняя часть, верхний бьеф]); кроме того, инфузории делились по биотопу на планктонные и бентосные. Результаты трехфакторного дисперсионного анализа для планктонного биотопа в 1990 г. (сила влияния факторов на параметры сообществ инфузорий и некоторые абиотические факторы) представлены в табл. 5.

**Таблица 5**  
**Результаты 3-х факторного ANOVA [21, с. 225]**

| Параметры         | А    | В    | С   | ВС   | ABC |
|-------------------|------|------|-----|------|-----|
| Численность       | 24,2 | 8,1  | 4,6 | 13,8 |     |
| Биомасса          |      | 7,4  |     |      |     |
| Продукция         |      | 7,5  |     |      |     |
| Температура воды  | 0,7  | 90,7 | 0,6 | 4,6  | 1,5 |
| Сапробность среды |      | 8,7  |     | 13,6 |     |

Воздействие факторов АВ и АС оказалось недостоверным на 95% уровне значимости.

Из результатов, представленных в табл. 5, очевидно, что самым «воздействующим» на инфузории фактором является сезонность (В); заметно и влияние «север – юг» (А) на численность инфузорий (Куйбышевское – 150, Саратовское – 130, Волгоградское – 400 тыс. экз./м<sup>3</sup>; [21, с. 220]). Всё это (использование дисперсионного анализа) соответствует современным тенденциям экологических исследований инфузорий – требуется «более пристальное изучение того, как взаимодействуют факторы в процессе определения динамики и численности» [49, р. 384].

Число примеров использования ANOVA в экологических исследованиях огромно. Естественно, больше всего работ [9], выполненных в рамках однофакторного дисперсионного анализа (из отечественных научных разработок назову еще [1, 44, 81, 90, 92]), меньше – двухфакторного (например, такие [17, 33, 77, 88 и др.]), и еще меньше – трехфакторного [3, 76 и др.).

Из сравнительно «свежих» работ по прямому градиентному анализу остановлюсь на монографии С.И. Мироновой [44]. Основные её результаты количественной обработки сосредоточены в главе «Ординационный анализ техногенной растительности», которая может быть

названа «Увертюрой<sup>24</sup> к однофакторному прямому градиентному анализу», континуум изменений растительности развернут в хроноклины, которые проанализированы в традиционных подходах школы уфимских фитоценологов (построение графиков распределения проективного покрытия видов во времени, определение средневзвешенной и силы влияния фактора времени на распределение видов). При этом анализ проведен для отвалов Западной и Южной Якутии по пяти группам видов («сквозного» распространения, тяготеющих к ранним, средним и более поздним стадиям сукцессии, имеющим двухвершинное распределение, т. е. находящихся под существенным влиянием не только фактора времени); кроме того, различался верхний (поверхность и склоны отвалов) и нижний уровни («пазухи» и берега искусственных водоемов). Теперь становится ясным, почему данный однофакторный анализ хроноклинов можно назвать лишь «увертюрой» – если бы Миронова довела обработку до конца (а фактически материал предполагает трехфакторную схему градиентного анализа: сукцессионное время, запад–юг, верх–низ), то можно было бы с полным правом назвать эту главу «Симфонией прямому градиентному анализу». Правда, в последнем разделе главы Миронова предпринимает попытку качественного сравнения хроноклинов разных вариантов и приходит к ряду интересных выводов (закономерное изменение фитосоциологического спектра по пяти классам сообществ, снижение гамма-разнообразия низких местообитаний за счет выпадения неустойчивых луговых видов, различная роль в ходе сукцессии злаковых и бобовых компонент техногенных экосистем, наконец, контрастность видового состава разных этапов сукцессии, оцененная по степени сходства).

### КУДА ПРИДЕМ?

Большинство разговоров о будущем – это разговоры о настоящем. При этом содержательно говорить о будущем способны далеко не все, особенно, если это касается специальных вопросов. Можно смело констатировать, что обработка экологической информации сегодня происходит в полном соответствии с *принципом Пилюлькина*<sup>25</sup> – дисперсионный анализ «прописывается», практически, во всех ситуациях.

<sup>24</sup> Раз пошли такие «музыкальные» аналогии, то работу [4], из которой почерпнуты многие из рассмотренных здесь примеров, я назвал бы «Гимном прямому градиентному анализу»...

<sup>25</sup> Доктор Пилюлькин из книги Н.Н. Носова "Приключения Незнайки и его друзей" лечил жителей

Методы математической статистики позволяют отсеять недостоверные гипотезы о структуре и динамике экосистем. Еще раз подчеркну, – математическая статистика позволяет только отвергать гипотезы; принятие гипотезы на основе некоторых статистических тестов еще не означает их справедливость и доказанность («гипотеза может быть только проверена, но никогда не может быть доказана!» [2, с. 38]). В этом можно видеть оценочный характер эмпирико-статистических моделей [65].

И, все-таки, завершая обсуждение дисперсионного анализа, я попытаюсь осторожно заглянуть немного вперед и умозрительно порассуждать о будущем прикладной статистики (прежде всего, статистики по-Фишеру).

**Классика.** Естественно, фундаментальные открытия Р. Фишера (такие как  $F$ -функция распределения, планирование экспериментов, рандомизация, концепция доверительного интервала, критерий значимости, сам термин «дисперсия», который, похоже, также был введен Фишером в 1916 г.) не исчезнут. Как отмечал в своей Фишеровской лекции<sup>26</sup> 1996 г. Б. Эфрон<sup>27</sup> [103, р. 112], «они (*фундаментальные открытия*. – *Г.Р.*) могут стать менее заметными, хотя сегодня мы используем эти идеи почти точно так, как их предложил Фишер, <...> могут скрыть влияние Фишера, но не уменьшат его. <...> На самом деле я предполагаю, что у старого Фишера будет очень хороший XXI-й век».

**Новые компьютерные возможности.** Широкое распространение сегодня высокопроизводительных компьютеров позволяет, *во-первых*, обрабатывать огромные массивы данных; *во-вторых*, использовать некоторые алгоритмы для определения более точных доверительных интервалов, которые были неосуществимы ранее (переход от линейных моделей к существенно нелинейным, нейронным сетям и пр.);

---

Цветочного города двумя препаратами – наружно использовал йод, внутрь – настойку...

<sup>26</sup> Фишеровская лекция (R.A. Fisher Lecture) является знаком высокого признания достижений в статистической науке, она читается на совместных заседаниях (Joint Statistical Meetings) в Северной Америке и впоследствии публикуется в статистическом журнале. Лектор получает мемориальную медаль и денежное вознаграждение в размере \$1000. Премия была учреждена в 1963 г. Североамериканским комитетом президентов статистических обществ (Committee of Presidents of Statistical Societies; COPSS) в знак признания заслуг Рональда Фишера.

<sup>27</sup> Эфрон, Брэдли (Bradley Efron, г. р. 1938) – американский статистик, предложил метод статистического бутстрэпа; член НАН США (1986).

*в-третьих*, Эфрон справедливо подчеркивает, что «теории Фишера хорошо подходят для жизни в компьютерный век. Казалось, что Фишер мыслит алгоритмически (to think in algorithmic terms)» [103, р. 112-113]. Действительно, оценки максимального правдоподобия, доверительные интервалы, таблицы ANOVA – все это выражено у Фишера алгоритмически и легко расширяется с помощью современных вычислений.

**К дискуссии о «мнимых повторностях (pseudoreplication)».** Десятая глава книги "Как лгать при помощи статистики" [86], написанной еще в 1954 г., называется «Как поставить статистика на место». Автор предлагает проверять статистические данные с помощью пяти простых вопросов:

- Кто это говорит?
- Откуда ему это известно?
- Чего не хватает?
- Не подменен ли объект исследования?
- Есть ли в этом смысл?

Жаль, я узнал об этой книге достаточно поздно.

Здесь я остановлюсь чуть подробнее (но не сильно), так как этой международной и достаточно бурной дискуссии мы с коллегами посвятили целый сборник статей [54].

В статье С. Хёлберта<sup>28</sup> [134, 135] о «мнимых повторностях», которая по свидетельству ряда исследователей и самого автора [30, 32, 133, 135] имела большое воспитательное и «санитарно-регла-ментирующее» значение, мнимая повторность определена «как проверка статистических гипотез при оценке эффекта влияния фактора по экспериментальным данным, когда группы воздействий не могут быть признаны повторными (хотя выборки были сформированы), или эти повторности не являются статистически независимыми» (Hurlbert, 1984, р. 187). Однако не все положения этой статьи встретили поддержку; более того, они вызвали и аргументированные возражения. Если резюмировать все «за» и «против» этой дискуссии (подробнее, см. [54]), то они сводятся к нескольким главным положениям.

- «Эксперимент включает пять последовательных стадий: гипотеза, планирование, реализация, статистический анализ и интерпретация. Гипотеза обладает первоочередной важностью, поскольку если она не

---

<sup>28</sup> Хёлберт, Стюарт (Stuart H. Hurlbert; г. р. 1940) – американский эколог, гидробиолог, лимнолог, статистик; эта статья признана научной классикой (индекс цитирования SCI более 2000); Американская статистическая ассоциация удостоила Хёлберта премией Дж. Снедекора за лучшую статью в области биометрии в 1984 г.

удовлетворяет некоторым критериям качества, то даже самый правильно проведенный эксперимент будет иметь не слишком большую ценность» [134, р. 188] – *с этим, достаточно очевидным, утверждением трудно спорить.*

- «Планируя эксперимент, Вы должны ясно и последовательно представлять место Вашей работы в логической структуре научного процесса и соответственно этому выбирать экспериментальные методы» [142, р. 36] – *таким образом, планирование эксперимента – это тактика; стратегией научного исследования является выбор «главной» установки: проводить исследование в дедуктивном или индуктивном духе.* В последнем случае, логика индукции требует «правильно организовать» анализируемую выборку (отбор объектов, повторности и пр.).
- «Мнимая повторность может быть определена в терминах дисперсионного анализа как проверка эффекта воздействия, основанная на оценке дисперсии, неадекватной рассматриваемой гипотезе» (Hurlbert, 1984, р. 190). Иными словами, логически корректное выявление причинной связи может искажаться воздействием посторонних факторов и, «поэтому в любом случае есть альтернатива объяснения, казалось бы, очевидного факта воздействия» [142, р. 27] – *также очевидное утверждение, т.к. математическая статистика ничего не доказывает, она только отвергает гипотезы.*
- «В сущности, проблема, затронутая С. Хёлбертом, чрезвычайно проста: без повторности результат воздействия неотличим от случайных влияний» [98, р. 394] – *естественно, для того, чтобы по выборке можно было получить достаточно точную (в вероятностном смысле) характеристику всей генеральной совокупности, необходимо, кроме планирования системы отбора образцов в выборку, планировать и само число этих образцов.*
- «Некоторые авторы привели доводы в пользу объективной необходимости и самодостаточности неповторяемых экологических исследований... И мы готовы признать, что некоторые исследования без повторностей могут быть корректно проанализированы с использованием таких статистических методов как анализ временных рядов, численный ресамплинг, ANOVA, или методов, основанных на байесовском подходе» [133, р. 261] – *этот справедливый вывод через 20 лет подтверждает и сам автор «мнимых повторностей» – «отсутствие повторности воздействий не только не со-*

ставляет мнимую повторность, но и не всегда мешает проведению корректных статистических тестов о результатах воздействия» [135, р. 594; см. выше, также, [6]].

Здесь же сошлюсь на аргументированный ответ [12, 28] последователям Хёлберта, которые «обрушились» [31], считая свое мнение «истиной в последней инстанции», на статью известного количественного фитосоколога В.С. Ипатова [27]. Полностью солидарен с Ипатовым и Боголюбовым в том, что для малых выборок «проверить нормальность распределения не представляется возможным» (28, с. 411). И готов подписаться под заключительными словами и В.И. Василевича (12, с. 414), – «оппоненты Ипатова очень односторонне подошли к проблеме применения статистических методов в биологии. Не следует приписывать оппоненту то, что он не писал и не думал. В дискуссии следует стараться понять позицию оппонента, а не обвинять его в идеологических отклонениях», – и В.С. Ипатова [28, с. 412], который справедливо считает, что все «это дурно пахнет». Могу предположить, что такой «беспардонный наезд» связан с необходимостью хоть чем-то отчитаться по полученному гранту между Финской Академией и РФФИ...

Так из-за чего же горел весь этот «сыр-бор»? Если несколько отстраниться и «отжать» обильно повторяющиеся рассуждения саморекламного характера<sup>29</sup> [30, 32, 134, 135], то в «сухом остатке» представлений Хёлберта будет лишь два совершенно бесспорных тезиса (или, как говорилось, «логика понимания того, что есть хороший эксперимент» [47, с. 4]):

- не всегда выводы, полученные при изучении частных выборок, корректно распространять на всю генеральную совокупность;
- оценка степени влияния фактора может оказаться ошибочной, если изучаемое воздействие должным образом не локализовано, а сопоставляемые данные взяты из недостаточно рандомизированных источников.

Завершая обсуждение своего участия в этой, в целом, совершенно неконструктивной дискуссии, соглашусь со следующей её оценкой [89, с. 91]: «обильное цитирование статьи (*имеется в*

<sup>29</sup> Чего, например, только стоит «строгий укор всей российской науке, не желающей читать взхлёб его труды. А эпиграф к другой собственной статье (2004) "чем разглагольствовать, читайте мои труды" – это же прямая цитата из профессора Выбегаллы...» [92]; уточню для молодежи: Амвросий Амбрузович Выбегалло – персонаж повестей "Понедельник начинается в субботу" и "Сказка о тройке" [75].

виду [134] – Г.Р.) в зарубежных источниках определяется скорее не её научно-методологической ценностью, либо грандиозностью полученных результатов, а чисто психологическими моментами:

- желанием примкнуть к развернутой критикующей компании и самому получить шанс уличить;
- попыткой отрицать правомочность клейма «псевдорепликация» на своих работах, либо работах своих коллег;
- просто как предупредительный сигнал рецензенту «Хёлберта читал, в своей работе учитывал и в этом направлении замечаний ко мне быть не должно».

Так завязывается 20-летняя дискуссия, которая лишь частично имеет вид научной».

Похоже, и сам инициатор дискуссии С. Хёлберт утратил к ней интерес: он пару раз после 2010 г. выступил с аналогичной критикой статей в психологии, а потом стал писать о глобалистике, утечке мозгов из Южной Америки, исламизации Сан-Диего, про «иммиграционные войны» и пр. (<http://www.bio.sdsu.edu/pub/stuart/Hurlbert-CV0517.pdf>).

**Неоклассика (?). Байесовский подход к статистике (A Bayesian World).** Действительно, за 100 лет с момента открытий Фишера в статистический арсенал все прочнее входят методы байесовской статистики (связана с именем Т. Байеса<sup>30</sup>; это статистическая теория, основанная на интерпретации вероятности [когда вероятность отражает *степень доверия* событию], которая может измениться при получении новой информации, в отличие от фиксированного значения, основанного на частотном подходе). Важность теоремы Байеса состоит в том, что она дает нам возможность очень точно измерять степень, в которой новые сведения должны менять наши представления о мире. Другими словами, по теореме (формуле) Байеса можно более точно пересчитывать вероятность, беря в учет как ранее известную информацию, так и данные новых наблюдений. Главная, видимо, особенность теоремы Байеса состоит в том, что для её практического применения обычно требуется огромное количество вычислений-пересчетов, а потому расцвет методов байесовых оценок пришелся как раз на революцию в компьютерных и сетевых инфотехнологиях.

После того, как столько сказано о формуле (теореме) Байеса, приведу её вид:

<sup>30</sup> Байес (Бейз), Томас (Thomas Bayes; 1702-1761) – британский математик, пресвитерианский священник, член Лондонского королевского общества (1742).

$$P(A/B) = \frac{P(B/A)P(A)}{P(B)},$$

где  $P(A/B)$  – вероятность того, что событие  $A$  произойдет при условии, что произошло событие  $B$ ;  $P(B/A)$  – вероятность того, что событие  $B$  произойдет при условии, что произошло событие  $A$ ;  $P(A)$  и  $P(B)$  – вероятности того, что произойдут события  $A$  и  $B$ , соответственно. Эта формула по своей значимости для науки и человечества в целом, по-видимому, не уступает уравнению Эйнштейна –  $E = mc^2$ .

«Байес не опубликовал свою теорему, она была напечатана спустя два года после его смерти, когда родственники Байеса попросили Ричарда Прайса<sup>31</sup> посмотреть записи на предмет чего-нибудь полезного для публикации. После того, как Прайс нашел и отредактировал записи, работа Байеса была прочитана в Британском Королевском обществе. Но почему-то идеи её не предавались публичной огласке – и французскому математику и физику Лапласу "пришлось" переоткрывать теорему Байеса снова, но уже в виде привычной нам формулы.

Легко заметить, что теорема Байеса не допускает вероятности чего-либо в 100%: она становится просто бессмысленной. Отличие веры от науки как раз в том, что всегда есть вероятность, что научная теория может ошибаться, с помощью теоремы Байеса и серии экспериментов вероятность эту можно уменьшить или увеличить. В вопросах веры нет места для ошибки – и наука там не работает» (<https://zen.yandex.ru/media/id/5ad580c8d7bf21db3be33855/>).

Итак, теорема Байеса дает нам критерий, позволяющий судить о том, адекватно ли мы используем новые знания. На этом основана концепция *идеального байесовского наблюдателя* – воображаемого существа, всегда использующего получаемые сведения наилучшим из возможных способов [45, 78, 137, 131].

«Бум» байесовской статистики приходится на последние десятилетия XX в. Уже в 1974 г. Д. Линдли<sup>32</sup> предсказывал, что XXI-й век будет

<sup>31</sup> Прайс, Ричард (Richard Price; 1723-1791) – британский философ, проповедник, публицист; член Лондонского королевского общества (1765).

<sup>32</sup> Линдли, Деннис (Dennis Victor Lindley; 1923-2013) – британский математик, статистик. Он ищет более глубокое обоснование идей Неймана, Пирсона, Вальда, Фишера, находит его в теории Байеса и становится «великим миссионером для Байесовского Евангелия – great missionary for the Bayesian gospel» ([https://en.wikipedia.org/wiki/Dennis\\_Lindley](https://en.wikipedia.org/wiki/Dennis_Lindley)). Когда Линдли приехал в Университетский колледж Лондона и возглавил кафедру статистики, «это было так, как будто Свидетель Иеговы был избран Папой

байесовским, а Г. Джеффрис<sup>33</sup> писал, что теорема Байеса «для теории вероятности, то же, что теорема Пифагора для геометрии» [136, р. 31]. Координирующим центром продвижения байесовской статистики стал симпозиум, проводимый в Валенсии раз в четыре года по инициативе Х. Бернардо<sup>34</sup>.

В формуле Байеса два ключевых элемента:  $P(A/B)$  и  $P(B/A)$ . Величина  $P(A/B)$  говорит нам, насколько мы должны изменить наше представление об окружающем мире ( $A$ ) после получения новых сведений ( $B$ ). Величина  $P(B/A)$  говорит нам, каких сведений ( $B$ ) мы должны ожидать, исходя из нашего убеждения ( $A$ ). Это позволяет нам делать предсказания и отслеживать ошибки в них. Руководствуясь своими представлениями об окружающем мире, мы можем предсказывать характер событий, которые будут отслежены –  $P(B/A)$ . Что же происходит, когда такое предсказание оказывается ошибочным? Отслеживать ошибки в подобных предсказаниях особенно важно, потому что мы можем использовать их для уточнения и улучшения своих представлений об окружающем мире, о  $P(A/B)$ . После внесения уточнения мы получаем новое представление о мире, и может снова повторить ту же процедуру, сделав новое предсказание о характере наблюдаемых событий. С каждым повтором этого цикла ошибка в предсказаниях уменьшается. Фактически, байесовская процедура является итеративной.

Интересный факт: «Термин "байесовский" (Bayesian) впервые применил Рональд Фишер. Но в его лексиконе этот термин был ругательным. Он считал, что теория Байеса "должна быть полностью отвергнута"» (<https://dyakonov.org/2018/07/30/байесовский-подход/>).

Мир прикладной статистики, похоже, нуждается в эффективном компромиссе между байесовскими и частотными идеями, но пока, явной замены для фишеровской статистики нет.

**Другие дороги.** Естественно, кроме перечисленных направлений коллеги отмечают и другие возможности будущего развития прикладной статистики [50, 64, 103, 136, 145, 147]. Так, например, выделены такие пять «точек роста» (Орлов, 2016): **непараметрика, робастность,**

**бутстреп, статистика интервальных данных, статистика нечисловых данных.**

Л. Сэвидж<sup>35</sup> писал, что его наставники по статистике считали, «что статистические методы Фишера для научных работников являются серьезным введением исследователя в статистику. <...> Они и некоторые другие, хотя, конечно, не все, давали один и тот же совет: «чтобы стать статистиком, заниматься статистикой и размышлять требуется фишеровское терпение, уважение и скептицизм» [147, р. 441-442].

## ЗАКЛЮЧЕНИЕ

В силу того, что основной целью данной статьи было рассмотрение проблем дисперсионного анализа, «за бортом» моего внимания остались генетико-математические аспекты творчества Рональда Эйлмера Фишера и, конечно, его фундаментальная теорема естественного отбора, опубликованная в 1930 г. в монографии "The Genetical Theory of Natural Selection" [85]. Именно соединение математических методов менделевской генетики с дарвиновской теорией естественного отбора и следует признать основным достижением Фишера в биологии; эта концепция легла в основу популяционной генетики и современной синтетической теории эволюции. А фундаментальная теорема естественного отбора продолжала и продолжает привлекать к себе внимание «математических экологов» [2, 5, 7, 56, 57, 70, 71, 96, 99, 125, 138, 144 и др.].

Тривиально, но следует избегать суждений о великих умах по сегодняшним меркам даже относительно недавнего прошлого. И все-таки, большинство исследователей сходятся в том, что Фишер был великим естествоиспытателем:

- «Сэр Рональд Фишер был настоящим гением. Он был фактическим основателем не только популяционной генетики, но и, во многом, современной статистики и ввел математическое определение информации задолго до Клода Шеннона» [34, с. 23).
- «Гений, который практически в одиночку заложил основы современной статистической науки» [132, р. 157].
- «Даже сегодня я иногда встречаю генетиков, которые спрашивают меня, правда ли, что великий генетик Р.А. Фишер также был

– it was as though a Jehovah's Witness had been elected Pope» [146, р. 665].

<sup>33</sup> Джеффрис, Гарольд (Sir Harold Jeffreys; 1891-1989) – британский математик, статистик, геофизик, астроном; член Лондонского королевского общества (1925), иностранный член Национальной академии наук США (1945), президент Королевского астрономического общества (1955-1957).

<sup>34</sup> Бернардо, Хосе (José-Miguel Bernardo; г. р. 1950) – испанский математик, статистик.

<sup>35</sup> Сэвидж, Леонард (1917-1971) – американский математик, статистик; лауреат Нобелевской премии по экономике М. Фридман сказал, что Сэвидж был «одним из немногих людей, которых я встречал и которого я без колебаний назвал бы гением» [130, р. 146].

важным (знаковым) статистиком» [147, р. 445].

- «Р.А. Фишер был величайшим статистиком своего времени, возможно, всех времен. Он был одним из великого трио, в которое входили он сам, Сьюэлл Райт и Дж.Б.С. Холдейн<sup>36</sup>, заложившие математические основы эволюционной теории. Его книга "Генетическая теория естественного отбора" [115] была названа самой глубокой книгой об эволюции со времен Дарвина» [129, р. 415].
- В 20-м столетии чл.-корр. РАН И.А. Захаров-Гезехус выделил 33 основных события в истории генетики; одно из них: «1930-1932. Математическая теория эволюции популяций дана в работах Р.А. Фишера, С. Райта, Дж. Холдейна» [23, с. 69].
- «Хотя рассуждения Фишера и не были всегда убедительными, он является одним из крупнейших (если не крупнейшим) статистиков, когда-либо живших» [72, с. 111].
- «Фишер был гением первого ранга, который твердо претендует на звание самого важного прикладного математика 20-го века. Его работа обладает уникальным качеством смелого математического синтеза в сочетании с максимальной практичностью... Это печать великого мыслителя, а статистика и наука в целом находятся у него в долгу» [103, р. 113].

Наконец, Л. Закс<sup>37</sup> [22, с. 514], завершая теоретическую часть своего учебника, называет пять важнейших дат в истории теории вероятностей и математической статистики. В это число попадает и первая половина XX века с работами К. Пирсона, Р. Фишера, Э. Пирсона, Е. Неймана и А.Н. Колмогорова.

А завершить эту статью «по поводу круглой даты» хочу еще одной, фактически, цитатой: «All models are wrong; the practical question is how wrong do they have to be to not be useful – Все модели ошибочны; практический вопрос – насколько неправильными они должны быть, чтобы стать бесполезными» и «Essentially, all models are wrong, but some are useful – В сущно-

сти, все модели неправильны, но некоторые полезны» [97<sup>38</sup>, Drape, р. 74, 424].

## СПИСОК ЛИТЕРАТУРЫ

1. **Абакарова Б.А., Асадулаев З.М.** Внутривидовая изменчивость признаков *Rosa oxydon* Boiss. на Гунибском плато // Юг России: экология, развитие. 2008. № 4. С. 48-52.
  2. **Абросов Н.С., Боголюбов А.Г.** Экологические и генетические закономерности сосуществования и коэволюции видов. Новосибирск: Наука, 1988. 333 с.
  3. **Бессчетнова Н.Н.** Сравнительная оценка клонов плюсовых деревьев сосны обыкновенной (*Pinus sylvestris* L.) по параметрам хвои в трехфакторном дисперсионном анализе // Лесной вестн. 2012. № 6. С. 199-204.
  4. Биоиндикация: теория, методы, приложения. Тольятти: ИЭВБ РАН, 1994. 266 с.
  5. **Боголюбов А.Г.** О принципе конкурентного исключения и механизмах поддержания видовой разнообразия сообществ // Биол. науки. 1989. № 11. С. 5-18.
  6. **Боголюбов А.Г.** Как поступать с однократными наблюдениями, или двухфакторный дисперсионный анализ с одним наблюдением в ячейке // Ботан. журн. 1993. Т. 78, № 9. С. 84-93.
  7. **Боголюбов А.Г.** Математические модели эколого-генетических процессов конкуренции видов: Автореф. дис. ... докт. физ.-мат. наук. СПб.: СПб. техн. ун-т, 1995. 34 с.
  8. **Боровиков В.П.** Популярное введение в современный анализ данных в системе STATISTICA. М.: Горячая линия - Телеком, 2013. 288 с.
  9. **Будилова Е.В., Дрогалина Ж.А., Терехин А.Т.** Основные направления современной экологии и ее математический аппарат: анализ публикаций // Журн. общ. биол. 1995. Т. 56, № 2. С. 179-189.
  10. **Бурцева Е.И.** Опыт количественного анализа растительности солончаковых лугов в пойме Средней Лены // Экология и ценология лугов Центральной Якутии. Якутск: Изд-во Якутс. госун-та, 1978. С. 128-164.
  11. **Бююль А., Цёфель П.** SPSS: Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. СПб.: «ДиаСофтЮП», 2005. 608 с.
  12. **Василевич В.И.** Некоторые комментарии к дискуссии об использовании статистических методов в биологии // Ботан. журн. 2012. Т. 97, № 3. С. 109-110.
  13. **Воробьев В.С.** К 100-летию со дня смерти Фрэнсиса Гальтона – основателя психогенетики // Теоретическая и экспериментальная психология. 2011. Т. 4, № 3. С. 85-100.
- <sup>38</sup> Бокс, Джордж (George Edward Pelham Box; 1919-2013) – британский, американский статистик; член Лондонского королевского общества (1985); в контексте данной статьи интересен тот факт, что он был женат на Джоан (Joan Fisher) – дочери Р. Фишера и авторе его биографии [97].

<sup>36</sup> Райт, Сьюэлл (Sewall Green Wright; 1889-1988) – американский генетик, эволюционист, статистик. Холдейн, Джон (John Burdon Sanderson Haldane; 1892-1964) – британский генетик, эволюционист, физиолог, биохимик, биометрист, популяризатор и философ науки; член Лондонского королевского общества (1932), иностранный, почетный член АН СССР (1942).

<sup>37</sup> Закс, Лотар (Lothar Sachs; 1929-2019) – немецкий статистик.

14. **Гинзбург Э.Х.** Сравнение оценок показателя силы влияния // Генетика, 1969. Т. 5, № 4. С. 150-160.
15. **Гинзбург Э.Х.** Оценка показателя силы влияния и планирование дисперсионного комплекса // Генетика. 1973. Т. 9, № 3. С. 156-162.
16. **Глэд Д.** Будущая эволюция человека – Евгеника XXI века. М.: «Издательство Захарова», 2005. 176 с.
17. **Гонгальский К.Б.** Лесные пожары и почвенная фауна. М.: Тов-во науч. изд. КМК, 2014. 169 с.
18. **Грейг-Смит П.** Количественная экология растений. М.: Мир, 1967. 359 с.
19. **Дьяконов А.** Байесовский подход // Анализ малых данных. КвазиНаучный блог Александра Дьяконова. 30.07.2018. [https://dyakonov.org/2018/07/30/байесовский-подход].
20. **Елизаров А.В.** Прямой градиентный анализ пастбищной дигрессии типчаково-ковыльной степи на Тургайском плато // Биоиндикация: теория, методы, приложения. Тольятти: ИЭВБ РАН, 1994. С. 126-130.
21. **Жариков В.В., Ротарь Ю.М.** К оценке реакции свободноживущих инфузорий водохранилищ Средней и Нижней Волги на изменения некоторых факторов среды // Биоиндикация: теория, методы, приложения. Тольятти: ИЭВБ РАН, 1994. С. 205-225.
22. **Закс Л.** Статистическое оценивание. М.: Статистика, 1976. 598 с.
23. **Захаров И.А.** Генетика в XX веке: Очерки по истории. М.: Наука, 2003. 78 с.
24. **Ивантер Э.В., Коросов А.В.** Элементарная биометрия. Петрозаводск: Изд-во ПетрГУ, 2010. 104 с.
25. **Иванюшкин А.Я.** К.Э. Циолковский и русская евгеника // Здравоохранение Российской Федерации. 2014. Т. 58, № 1. С. 48-52.
26. **Инге-Вечтомов С.Г.** Генетика с основами селекции: Учеб. для биол. спец. ун-тов. М.: Высш. шк., 1989. 591 с.
27. **Ипатов В.С.** Осторожно – биометрика. (Об использовании оценок «достоверности» при исследовании количественных закономерностей) // Ботан. журн. 2010. Т. 95, № 10. С. 1494-1498.
28. **Ипатов В.С.** По поводу статьи М.В. Козлова и Е.Л. Воробейчика «Осторожно – некорректные подходы к анализу данных или об использовании непараметрических методов в ботанических исследованиях» // Ботан. журн. 2012. Т. 97, № 3. С. 411-412.
29. **Канаев И.И.** Фрэнсис Гальтон, 1822-1911. Л.: Наука, 1972. 135 с.
30. **Козлов М.В.** Мнимые повторности (pseudoreplication) в экологических исследованиях: проблема, не замеченная российскими учеными // Журн. общ. биол. 2003. Т. 64, № 4. С. 292-307.
31. **Козлов М.В., Воробейчик Е.Л.** Осторожно – некорректные подходы к анализу данных, или об использовании непараметрических методов в ботанических исследованиях // Ботан. журн. 2012. Т. 97, № 3. С. 402-410.
32. **Козлов М.В., Хёлберт С.Х.** Мнимые повторности, бесплодные дискуссии и интернациональная сущность науки: ответ Д.В. Татарникову // Журн. общ. биол. 2006. Т. 67, № 2. С. 145-152.
33. **Крюкова А.В., Абрамова Л.М.** Влияние экологических факторов на изменчивость морфометрических параметров редкого вида *Iris punila* L. // Изв. Саратов. ун-та. Нов. серия. Сер.: Химия. Биология. Экология. 2018. Т. 18, № 2. С. 232-236.
34. **Кунин Е.В.** Логика случая. О природе и происхождении биологической эволюции / Автор. пер. с англ. = The Logics of Chance. The Nature and Origin of Biological Evolution. М: ЗАО "Издательство Центрполиграф", 2014. 527 с. (Koonin E.V. The Logic of Chance: The Nature and Origin of Biological Evolution. Upper Saddle River (NJ): FT Press, 2011. 516 p.).
35. **Лакин Г.Ф.** Биометрия / 4-е изд. М.: Высш. шк., 1990. 352 с.
36. **Лукомский Я.И.** Теория корреляции и её применение к анализу производства. М.: Госстатиздат, 1958. 387 с.
37. **Любищев А.А.** Дисперсионный анализ в биологии. М.: Изд-во Моск. ун-та, 1986. 200 с.
38. **Миллс Ф.** Статистические методы. М.: Госстатиздат, 1958. 799 с. (Mills F.C. Statistical Methods / 3<sup>rd</sup> ed. N. Y.: Henry Holt & Company, 1955. 842 p.).
39. **Мильчаков К.С.** Дисперсионный анализ ANalysis Of VAriance (ANOVA) просто о сложном // Блог «Литобзор». URL. [https://lit-review.ru/biostatistika/dispersionnyj-analiz-anova/].
40. **Миркин Б.М.** Статистический анализ экоклиннов // Ботан. журн. 1971. Т. 56, № 12. С. 1772-1788.
41. **Миркин Б.М., Наумова Л.Г.** Наука о растительности (история и современное состояние основных концепций). Уфа: Гилем, 1998. 413 с.
42. **Миркин Б.М., Розенберг Г.С.** Количественные методы классификации, ординации и геоботанической индикации // Итоги науки и техники. Ботаника. М.: ВИНТИ, 1979. Т. 3. С. 71-137.
43. **Миркин Б.М., Розенберг Г.С., Наумова Л.Г.** Словарь понятий и терминов современной фитоценологии. М.: Наука, 1989. 223 с.
44. **Миронова С.И.** Техногенные сукцессионные системы растительности Якутии (на примере Западной и Южной Якутии). Новосибирск: Наука, 2000. 151 с.
45. **Моррис У.Т.** Наука об управлении: Байесовский подход. М.: Мир, 1971. 304 с.
46. **Мурзина Е.** Рональд Эйлмер Фишер / Ronald Aylmer Fisher. Биография // Блог «Люди. Peoples.ru». 2010. [https://www.peoples.ru/science/mathematics/ronald\_aylmer\_fisher/index.html].
47. **Налимов В.В., Голикова Т.И.** Логические основания планирования эксперимента. М.: Металлургия, 1981. 152 с.
48. **Ниворожкина Л.И., Арженовский С.В., Рудяга А.А. и др.** Статистические методы анализа данных: Учебник. М.: РИОР; ИНФРА-М, 2016. 333 с.
49. **Орлов А.И.** Математика случая. Вероятность и статистика – основные факты: Учебное пособие. М.: МЗ-Пресс, 2004. 110 с.

50. Орлов А.И. Состояние и перспективы развития прикладной и теоретической статистики // Политематический сетевой электронный науч. журн. КубГАУ. 2016. № 115 (01). С. 1-25. (<https://cyberleninka.ru/article/n/sostoyanie-i-perspektivy-razvitiya-prikladnoy-i-teoreticheskoy-statistiki>).
51. Плохинский Н.А. Биометрия / 2-е изд. М.: Изд-во Моск. ун-та, 1970. 367 с.
52. Плохинский Н.А. Математическое оснащение биологов // Методы современной биометрии. М.: МГУ, 1978. С. 27-35.
53. Плохинский Н.А. Алгоритмы биометрии / 2-е изд. М.: Изд-во Моск. ун-та, 1980, 150 с.
54. Проблемы экологического эксперимента (Планирование и анализ наблюдений). Тольятти: Самар. НЦ РАН; Кассандра, 2008. 274 с.
55. Пушкин А.С. Путешествие в Арзум во время похода 1829 года // ПСС в 10-и т., изд. 2. Т. 6. М.: Изд-во АН СССР, 1957. С. 637-701.
56. Пых Ю.А. Равновесие и устойчивость в моделях популяционной динамики. М.: Наука, 1983. 181 с.
57. Ратнер В.А. Обобщенная приспособленность и фундаментальная теорема естественного отбора Фишера // Исследования по теоретической генетике. Новосибирск: Наука, 1972. С. 10-26.
58. Розенберг Г.С. Десять программ обработки экспериментальных данных на ЭВМ "Наири-К". Препринт Ин-та биол. БФАН СССР. Уфа, 1976а. 34 с.
59. Розенберг Г.С. Об одном несложном методе уточнения факториальной дисперсии в однофакторном дисперсионном комплексе // Аспекты оптимизации количественных исследований растительности. Уфа: БФАН СССР, 1976б. С. 92-103.
60. Розенберг Г.С. Модели в фитоценологии. М.: Наука, 1984. 240 с.
61. Розенберг Г.С. Анализ структуры и динамики сложных систем на ЭВМ (с примерами из фитоценологии). Уфа: БФАН СССР, 1985. 82 с.
62. Розенберг Г.С. О процедурах выравнивания частных средних в дисперсионном анализе // Биоиндикация: теория, методы, приложения. Тольятти: ИЭВБ РАН, 1994. С. 80-85.
63. Розенберг Г.С. Количественные методы фитоиндикации // Экологический мониторинг. Методы биологического и физико-химического мониторинга. Часть III. Н. Новгород: ННГУ, 1998. С. 5-27.
64. Розенберг Г.С. Статистические методы в фитоценологии на рубеже тысячелетий (к 50-летию выхода монографии П. Грейг-Смита) // Актуальные проблемы геоботаники: III Всерос. шк.-конф. Лекции. Петрозаводск: КарелНЦ РАН, 2007. С. 72-116.
65. Розенберг Г.С. Введение в теоретическую экологию / В 2-х т.; Изд. 2-е, исправленное и дополненное. Тольятти: Кассандра, 2013. Т. 1. 565 с. Т. 2. 445 с.
66. Розенберг Г.С. 20-20 (экологические юбилей года). Тольятти: Анна, 2019. 78 с.
67. Розенберг Г.С., Долотовский И.М. Еще раз о показателях силы влияния // Биол. науки. 1988. № 9. С. 105-110.
68. Розенберг Г.С., Шитиков В.К., Мозговой Д.П. Экологическая информатика: Учебное пособие. Самара: Изд-во Самар. ун-та, 1993. 151 с.
69. Рохо А. Возможно да, возможно нет. Фишер. Статистический вывод. М.: Де Агостини, 2015. 176 с. (Сер. Наука. Величайшие теории. Вып. 47).
70. Свирежев Ю.М. Возможные пути обобщения фундаментальной теоремы естественного отбора Р. Фишера // Журн. общ. биол. 1974. Т. 35, № 4. С. 590-599.
71. Свирежев Ю.М., Пасеков В.П. Основы математической генетики. М.: Наука, 1982. 511 с.
72. Секей Г. Парадоксы в теории вероятностей и математической статистике. М.: Мир, 1990. 240 с.
73. Симушкин С.В. (Составитель). Дисперсионный анализ. Методические разработки по специальному курсу. Ч. I. Оценки метода наименьших квадратов. Критерий Фишера. Однофакторный дисперсионный анализ. Казань: Казан. ун-т, 1998. 91 с.
74. Слуцкий Е.Е. О сложении случайных величин как причине циклических процессов // Вопросы конъюнктуры. 1927. Т. 3, № 1. (см.: Слуцкий Е.Е. Избранные труды. Теория вероятности и математическая статистика. М.; Л.: АН СССР, 1960. С. 99-132).
75. Стругацкий А.Н., Стругацкий Б.Н. Понедельник начинается в субботу; Сказка о Тройке: Фантаст. повести. М.: Книжный сад; СПб.: ИНТЕРОКО, 1993. 446 с. (Сер.: St. Petersburg).
76. Тимофеев А.В. К применению дисперсионного трехфакторного анализа в ботанико-географических исследованиях // Вестн. Санкт-Петербургского ун-та. Сер. 7. Науки о Земле. 2003. Вып. 1 (№ 7). С. 137-141.
77. Трушкова М.А., Тютикова А.А. Оценка экологической структуры сообществ мелких млекопитающих Среднего Поволжья методом дисперсионного анализа // Современные проблемы науки и образования. 2016. № 6. (<http://www.science-education.ru/ru/article/view?id=25885>).
78. Тулупьев А.Л., Николенко С.И., Сироткин А.В. Байесовские сети: логико-вероятностный подход. СПб: Наука, 2006. 607 с.
79. Уиттекер Р. Сообщества и экосистемы. М.: Прогресс, 1980. 328 с.
80. Урбах В.Ю. Биометрические методы: (Статистическая обработка опытных данных в биологии, сельском хозяйстве и медицине). М.: Наука, 1964. 415 с.
81. Федосеева Е.В., Терехова В.А., Цесаренко О.В., Гладкова М.М. Обработка результатов токсикологических исследований в статистической программе R // Принципы экологии. 2015. № 3. С. 12-26.
82. Феокистов В.Ф. Жужелицы как индикаторы экологических условий в лесных фитоценозах // Биоиндикация: теория, методы, приложения. Тольятти: ИЭВБ РАН, 1994. С. 191-201.
83. Фишер Р. Статистические методы для исследователей. М.: Госстатиздат, 1958. 267 с.
84. Фишер Р.А. Математика дамы, дегустирующей чай // Современные проблемы математики. Сб. статей. М.: Знание, 1981. С. 46-58.

85. **Фишер Р.** Генетическая теория естественного отбора. М.: Институт компьютерных исследований, Регулярная и хаотическая динамика, 2011. 294 с.
86. **Хафф Д.** Как лгать при помощи статистики = How to Lie with Statistics. М.: Альпина Паблишер, 2015. 163 с.
87. **Шеффе Г.** Дисперсионный анализ. М.: Физматгиз, 1963. 625 с. (Scheffe H. The Analysis of Variance. N. Y.: Wiley; L.: Chapman and Hall, 1959. 477 p.).
88. **Шитиков В.К., Розенберг Г.С., Зинченко Т.Д.** Количественная гидроэкология: методы, критерии, решения: в 2-х кн. М.: Наука, 2005. Кн. 1. 281 с.; Кн. 2. 337 с.
89. **Шитиков В.К., Розенберг Г.С., Крамаренко С.С., Якимов В.Н.** Современные подходы к статистическому анализу экспериментальных данных // Проблемы экологического эксперимента (Планирование и анализ наблюдений). Тольятти: Самар. НЦ РАН; Кассандра, 2008. С. 212-250.
90. **Шпаков А.Э., Волчков Ю.А.** Генотипическая изменчивость сортовых популяций табака по экологической пластичности // Юг России: экология, развитие. 2010. № 2. С. 31-35.
91. **Шустов М.В.** Лихеноиндикация экологических условий в борах Приволжской возвышенности // Биоиндикация: теория, методы, приложения. Тольятти: ИЭВБ РАН, 1994. С. 107-112.
92. Экосистема эстуария реки Невы: биологическое разнообразие им экологические проблемы. М.: Тов-во науч. изд. КМК, 2008. 477 с.
93. **Юдин Т.И.** Евгеника. Учение об улучшении природных свойств человека / 2-е перераб. и доп. изд. М.: Изд. М. и С. Сабашниковых, 1928. 288 с.
94. **Aldrich J.** A Guide to R. A. Fisher. Biographical Sketch. URL. <http://www.economics.soton.ac.uk/staff/aldrich/fisherguide/rafframe.htm>.
95. **Bennett J.H.** (Ed.). Natural Selection, Heredity, and Eugenics. Oxford: Oxford Univ. Press, 1983. 316 p.
96. Response Surfaces. N. Y.: John Wiley & Sons, 1987. 669 p.
97. **Box J.F.** R. A. Fisher: The Life of a Scientist. N. Y.: Wiley, 1978. 526 p.
98. **Cottenie K., De Meester L.** Comment to Oksanen (2001): reconciling Oksanen (2001) and Hurlbert (1984) // Oikos. 2003. V. 100. P. 394-396.
99. **Crow J.F.** Perspective: Here's to Fisher, additive genetic variance, and the fundamental theorem of natural selection // Evolution. 2002. V. 56, No. 7. P. 1313-1316.
100. **Dawkins R.** River Out of Eden: A Darwinian View of Life. N. Y.: Basic Books, 1995. 172 p.
101. **Doll R.** Fisher and Bradford Hill: Their personal impact // Internat. J. Epidemiology. 2003. V. 32, No. 6. P. 929-931. (Discussion P. 931-848).
102. **Doll R., Hill A.B.** Smoking and carcinoma of the lung // British Med. J. 1950. V. 2 (4682). P. 739-748.
103. **Efron B.** R. A. Fisher in the 21<sup>st</sup> century. Invited paper presented at the 1996 R. A. Fisher Lecture // Statist. Sci. 1998. V. 13, No. 2. P. 95-122.
104. **Fisher R.A.** Mendelism and biometry. Unpublished manuscript. 1911. (цит. по: [Bennett, 1983, p. 51-58]).
105. **Fisher R.A.** On an absolute criterion for fitting frequency curves // Messenger Math. 1912. V. 41. P. 155-160.
106. **Fisher R.A.** Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population // Biometrika. 1915. V. 10. P. 507-521.
107. **Fisher R.A.** The correlation between relatives on the supposition of Mendelian inheritance // Trans. R. Soc. Edinburgh. 1918. V. 52. P. 399-433.
108. **Fisher R.A.** Studies in Crop Variation. I. An examination of the yield of dressed grain from Broadbalk // J. Agricultural Sci. 1921. V. 11, No. 2. P. 107-135.
109. **Fisher R.A.** On the interpretation of  $\chi^2$  from contingency tables, and the calculation of P // J. Royal Stat. Society. 1922a. V. 85, No. 1. P. 87-94.
110. **Fisher R.A.** On the mathematical foundations of theoretical statistics // Philos. Trans. R. Soc. London A. 1922b. V. 222A. P. 309-368.
111. **Fisher R.A.** On the dominance ratio // Proceedings R. Soc. Edinburgh. 1922c. V. 42. P. 321-341.
112. **Fisher R.A.** On a distribution yielding the error functions of several well-known statistics // Proceedings of the International Congress of Mathematics, Toronto. 1924a. V. 2. P. 805-813.
113. **Fisher R.A.** Studies in crop variation. III. The influence of rainfall on the yield of wheat at Rothamsted // Philosophical Trans. R. Soc. B. 1924b. V. 213 (402-410). P. 89-142.
114. **Fisher R.A.** Applications of "Student's" distribution // Metron. 1925. V. 5. P. 90-104.
115. **Fisher R.A.** The Design of Experiments. Edinburgh: Oliver & Boyd. 1935. 252 p.
116. **Fisher R.A.** What sort of man is Lysenko? // Listener. 1948. V. 40. P. 874-875.
117. **Fisher R.A.** Statistical methods in genetics // Heredity. 1952. V. 6. P. 1-12.
118. **Fisher R.A.** Statistical methods and scientific induction // J. Royal Stat. Soc., Ser. B. 1955. V. 17. P. 69-78.
119. **Fisher R.A.** Statistical Methods and Scientific Inference. L.; Edinburgh: Oliver & Boyd. 1956. 175 p.
120. **Fisher R.A.** Dangers of cigarette smoking // British Med. J. 1957. V. 2, No. 5039. P. 297-298.
121. **Fisher R.A.** Lung cancer and cigarettes? // Nature. 1958a. V. 182, No. 4628. P. 108.
122. **Fisher R.A.** Cancer and smoking. Nature. 1958b. V. 182, No. 4635. P. 596.
123. **Fisher R.A.** Cigarettes, cancer, and statistics // Centennial Rev. 1959. V. 2. P. 151-166.
124. **Fisher R.A., Ford E.B.** The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* L. // Heredity. 1947. V. 1, No. 2. P. 143-174.
125. **Fisher R.A., Lyon M.F., Owen A.R.G.** The Sex chromosome in the house mouse // Heredity. 1947. V. 1, No. 3. P. 355-365.
126. **Fisher R.A., MacKenzie W.A.** Studies in Crop Variation. II. The manorial response of different potato

- varieties // *J. Agricultural Sci.* 1923. V. 13, No. 3. P. 311-320.
127. **Fisher R.A., Snell G.D.** A twelfth linkage group of the house mouse // *Heredity*. 1948. V. 2, No. 2. P. 271-273.
128. **Fisher R.A., Yates F.** *Statistical Tables for Biological, Agricultural and Medical Research*. L.: Oliver & Boyd, 1938. 146 p.
129. **Ford E.B.** R. A. Fisher: an appreciation // *Genetics*. 2005. V. 171, No. 2. P. 415-417.
130. **Friedman M., Friedman R.** *Two Lucky People: Memoirs*. Chicago: Univ. Chicago Press, 1998. 660 p.
131. **Gelman A., Carlin J.B., Stern H.S., Dunson D.B., Vehtari A., Rubin D.B.** *Bayesian Data Analysis / 3<sup>rd</sup> Ed.* L.: Chapman and Hall/CRC, 2013. 668 p.
132. **Hald A.** *A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713-1935*. N. Y.: Springer Science & Business Media, 2008. 225 p.
133. **Heffner R.A., Butler M.J. IV, Reilly C.K.** Pseudoreplication revisited // *Ecology*. 1996. V. 77, № 8. P. 2558-2562.
134. **Hurlbert S.H.** Pseudoreplication and the design of ecological field experiments // *Ecol. Monographs*. 1984. V. 54. P. 187-211. (рус. перевод: Хёрлберт С. Мнимые повторности и планирование экологических полевых экспериментов // *Планирование экологического эксперимента*. Тольятти: СамНЦ РАН; Кассандра, 2008. С. 9-45).
135. **Hurlbert S.H.** On misinterpretations of pseudoreplication and related issues: A reply to Oksanen // *Oikos*. 2004. V. 104. P. 591-597.
136. **Jeffreys H.** *Scientific Inference / 3<sup>rd</sup> ed.* Cambridge: Univ. Press, 1973. 244 p.
137. **Lee P.M.** *Bayesian Statistics: An Introduction*. N. Y.: John Wiley & Sons, 2012. 488 p.
138. **Li C.C.** 1967. Fundamental theorem of natural selection // *Nature*. 1967. V. 214. P. 505-506.
139. **Ludbrook J.** R. A. Fisher's life and death in Australia, 1959-1962 // *Amer. Statist.* 2005. V. 59 (May). P. 164-165.
140. **Lysenko T.D.** *The Science of Biology Today*. N. Y.: Internat. Publ., 1948. 62 p.
141. **Magnello M.E.** The non-correlation of biometrics and eugenics: Rival forms of laboratory work in Karl Pearson's Career at University College London, (In two parts) // *History of Science*. 1999. V. 37. P. 79-106, P. 125-150.
142. **Oksanen L.** Logic of experiments in ecology: is pseudoreplication a pseudoissue? // *Oikos*. 2001. V. 94. P. 27-38.
143. **Plutynski A.** What was Fisher's fundamental theorem of natural selection and what was it for? // *Stud. Hist. Phil. Biol. & Biomed. Sci.* 2006. V. 37. P. 59-82.
144. **Price G.R.** Fisher's fundamental theorem made clear // *Annals of Human Genetics*. 1972. V. 36. P. 129-140.
145. **Rao C.R.** R.A. Fisher: The founder of modern statistics // *Stat. Sci.* 1992. V. 7, No. 1. P. 34-48.
146. **Rivett P.** Review of aspects of uncertainty. A Tribute to D.V. Lindley // *J. Oper. Res. Soc.* 1995. V. 46, No. 5. P. 663-665.
147. **Savage L.J.** On rereading R. A. Fisher // *Ann. Statistics*. 1976. V. 4, No. 3. P. 441-500.
148. **Snedecor G.W.** *Calculation and Interpretation of Analysis of Variance and Covariance* // Iowa: Iowa State College Division of Industrial Science, 1934. 96 p.
149. **Threlkeld S.T.** Daphnia population fluctuations: Patterns and mechanisms // *Daphnia: Memorie Dell'istituto Italiano Di Idrobiologia: Dott. Marco De Marchi / Ed. by R.H. Peters, R. de Bernardi. Roma; Pallanza (Italy): Consiglio Nazionale*, 1987. V. 45. P. 367-388.
150. **Wald A.** Sequential tests of statistical hypotheses // *Ann. Math. Statist.* 1945. V. 16, No. 2. P. 117-186.
151. **Wald A.** *Sequential Analysis*. N. Y.: John Wiley and Sons, 1947. 212 p. (Вальд А. Последовательный анализ. М: Физматлит, 1960. 328 с.).
152. **Whittaker R.H.** Direct gradient analysis: techniques // *Handbook of Vegetation Science. Part 5. Ordination and Classification of Vegetation*. Hague: Dr. W. Junk B.V. Publ., 1973. P. 7-30.