

DOI 10.35264/1996-2274-2020-2-65-79

ОБРАБОТКА ЭКСПЕРТНЫХ ОЦЕНОК ПРИ ПРОВЕДЕНИИ ЭКСПЕРТИЗЫ НАУЧНЫХ ПРОЕКТОВ С ПРИВЛЕЧЕНИЕМ ЭКСПЕРТОВ ФЕДЕРАЛЬНОГО РЕЕСТРА

Р.Р. Илющенко, нач. отдела ФГБНУ НИИ РИНКЦЭ, renaldi@extech.ru

Т.В. Осипова, вед. аналитик ФГБНУ НИИ РИНКЦЭ, osipovatv@extech.ru

Рецензент: Т.А. Яркова

Статья посвящена вопросам повышения точности экспертных оценок при организации экспертизы научно-технических и инновационных проектов в информационной системе Федерального реестра экспертов. Рассмотрены основные методы обработки экспертных оценок. Показано, что для экспертно-аналитических работ, выполняемых информационной системой Федерального реестра экспертов, оптимальным методом обработки экспертных оценок является их усреднение в рамках экспертных групп. Проведен анализ оценочных шкал типовых экспертных анкет, определены статистические характеристики экспертных оценок. Обоснован выбор лучшей групповой оценки для типовых схем экспертизы, в качестве которой предложено использовать среднее арифметическое частных оценок. Рассмотрены критерии принятия решения и пороговые значения экспертных оценок. Предложен способ повышения достоверности и точности экспертных оценок путем ранжирования оцениваемых показателей по степени их значимости с последующим формированием частной итоговой оценки проекта в виде взвешенной суммы.

Ключевые слова: экспертиза, экспертно-аналитическое исследование, экспертная группа, экспертная оценка, частная оценка, групповая оценка, средняя оценка, достоверность оценки, точность оценки, метод обработки оценок, экспертная анкета, оцениваемый показатель, критерий оценки, оценочная шкала, весовой коэффициент, среднее арифметическое, медиана, ранжировка.

PROCESSING OF EXPERT ESTIMATES WHEN CONDUCTING AN EXAMINATION OF SCIENTIFIC PROJECTS WITH THE INVOLVEMENT OF EXPERTS OF THE FEDERAL ROSTER

R.R. Ilyushchenko, Head of Department, SRI FRCEC, renaldi@extech.ru

T.V. Osipova, Leading Analyst, SRI FRCEC, osipovatv@extech.ru

The article is devoted to the issues of increasing the accuracy of expert assessments when organizing the examination of scientific, technological and innovative projects in the information system of the Federal Roster of Experts. The main methods of processing expert assessments are considered. It is shown that for expert and analytical work performed by the information system of the Federal Register of Experts, the optimal method for processing expert assessments is their averaging within expert groups. The analysis of the rating scales of typical expert questionnaires is carried out, the statistical characteristics of expert assessments are determined. The choice of the best group assessment for typical examination schemes is substantiated, as it is proposed to use the arithmetic mean of private assessments. The criteria for making a decision and the threshold values of expert assessments are considered. A method is proposed for increasing the reliability and accuracy of expert assessments by ranking the estimated indicators according to their degree of

significance with the subsequent formation of a private final assessment of the project in the form of a weighted sum.

Keywords: expert examination, expert-analytical research, expert group, expert assessment, private assessment, group assessment, average assessment, assessment reliability, assessment accuracy, assessment method, expert questionnaire, assessed indicator, assessment criterion, assessment scale, weight coefficient, arithmetic mean, median, ranking.

Введение

В практике экспертно-аналитических исследований (ЭАИ), проводимых с привлечением экспертов Федерального реестра экспертов научно-технической сферы (Реестр), экспертам приходится давать количественные и качественные оценки различным объектам экспертизы.

Под объектами экспертно-аналитических исследований (экспертизы) в области науки, технологий и техники (ОЭ) обычно понимают предложения или проекты по развитию научной, научно-технической и инновационной сфер, направления и тематику исследований и разработок, а также оценки состояния научно-технологического комплекса и результаты научно-технической деятельности. Конкретные объекты ЭАИ определяются их заказчиками по согласованию с экспертной организацией и/или организацией – оператором Реестра.

Совокупность критериев, показателей и оценочных шкал, по которым осуществляется экспертиза ОЭ, разрабатывается применительно к каждому виду ЭАИ и утверждается заказчиком экспертизы, в качестве которого, как правило, выступает Минобрнауки России.

Число критериев и показателей в анкете может варьироваться от нескольких единиц до нескольких десятков, причем в анкете они обычно объединяются в разделы. Сумма баллов по всем критериям (показателям) образует общую оценку ОЭ. Кроме того, в анкете могут присутствовать вопросы, требующие нечисловой оценки («да/нет», «соответствует/не соответствует», «имеется/отсутствует» и т. п.). Итоговое мнение об ОЭ эксперт также выражает в нечисловой форме («поддержка целесообразна/поддержка нецелесообразна»).

ОЭ обычно поступают на экспертизу пакетами, содержащими от нескольких десятков до нескольких сотен проектов различной тематики. Экспертиза выполняется экспертными группами (ЭГ), в которых эксперты работают параллельно и независимо друг от друга. В информационной системе Реестра (ИС ФРЭ) ЭГ формируются чаще всего в составе трех экспертов, при этом в рамках ЭАИ, в зависимости от числа подлежащих экспертизе проектов, может быть сформировано от нескольких десятков до нескольких сотен ЭГ.

Эксперт может одновременно входить в состав нескольких ЭГ, а одна и та же группа может выполнять экспертизу нескольких однотипных проектов. Состав ЭГ эксперту неизвестен, тем более исключено прямое общение между экспертами внутри группы. Каждый эксперт, работая в своем личном кабинете в ИС ФРЭ, имеет доступ только к экспертной анкете и материалам проектов, предложенных ему для оценки.

Данная форма организации ЭАИ позволяет избежать таких негативных явлений, как конформизм (стремление «подогнать» свою оценку под мнение большинства), давление авторитетом и возможность сговора. Несмотря на то что схема с тремя независимыми экспертами не является наилучшей с точки зрения достоверности итоговой оценки [1], она, тем не менее, применяется наиболее часто, поскольку предпочтительна по другим соображениям, в первую очередь временным и финансовым.

Формируемые из пулов Реестра ЭГ чаще всего оказываются неоднородными, т. е. составленными из экспертов разного уровня компетентности по отношению к оцениваемым проектам. Это объясняется рядом причин.

Во-первых, аккредитованные в Реестре эксперты представляют собой лишь часть научного сообщества, и вероятность того, что в Реестре найдется эксперт, вектор компетентно-

сти которого будет точно соответствовать научно-техническому профилю конкретного ОЭ, представляется крайне малой.

Во-вторых, существенную роль играет такой субъективный фактор, как отсутствие у эксперта возможности выполнить предложенную экспертизу в установленные сроки (занятость по работе, командировка, болезнь и т. п.), вследствие чего экспертное задание приходится предлагать менее компетентным экспертам из сформированного рейтингового списка.

Применяемые в настоящее время процедуры формирования экспертных пулов и групп [2] предполагают ранжирование экспертов в порядке убывания их компетентности, при этом проведение экспертизы предлагается в первую очередь экспертам, компетентность которых максимально близка к научно-техническому профилю ОЭ.

Наиболее неоднородными являются группы в тех ЭАИ, при организации которых используются автоматизированные процедуры подбора экспертов, как, например, в экспертизе конкурсных заявок на получение грантов Президента России молодым российским ученым – кандидатам наук, докторам наук и ведущим научным школам [3]. Следствием неоднородности ЭГ является расхождение экспертных оценок, часто весьма значительное. Стоит, однако, отметить, что данное явление наблюдается также и в сравнительно однородных группах, однако такие случаи редки, и в большинстве своем однородные ЭГ демонстрируют удовлетворительную согласованность оценок.

Если в результате экспертизы ОЭ от группы экспертов получены разные оценки, то возникает задача формирования итоговой оценки, которая может считаться мнением группы и должна быть представлена в сводном экспертном заключении. Выбор релевантного метода обработки частных экспертных оценок, обеспечивающего получение наилучшей (с точки зрения достоверности и точности) групповой оценки, может быть сделан на основе анализа особенностей ЭАИ, выполняемых в ИС ФРЭ, характеристик экспертных анкет и используемых в них измерительных шкал.

Методы обработки экспертных оценок

На сегодняшний день в научной литературе описаны различные методы обработки экспертных оценок, ни один из которых, к сожалению, нельзя признать универсальным. Основные широко используемые в настоящее время методы – это *проверка согласованности мнений* экспертов (или *классификация* экспертов, если нет согласованности) и *усреднение мнений* экспертов внутри согласованной группы [4]. В тех видах экспертизы, где ответы экспертов связаны с объектами нечисловой природы (градации качественных признаков, ранжировки, разбиения, результаты парных сравнений, нечеткие предпочтения и т. д.), для анализа применимы *методы нечисловой статистики* [5, 6].

Проверка согласованности мнений экспертов внутри группы

Данный метод применяется в целях выделения в экспертной группе одной или нескольких подгрупп, оценки которых совпадают или близки друг к другу. Если число экспертов, одинаково оценивших ОЭ, составляет большинство ЭГ, то их оценка принимается за мнение группы, а остальные (девиантные) оценки попросту отбрасываются. Данный подход подвергается справедливой критике рядом авторов, считающих, что в отношении экспертной оценки ОЭ (особенно научно-технических) нельзя безапелляционно утверждать, что «большинство всегда право». Поскольку неизвестно «истинное» значение оценки, которой заслуживает ОЭ, частная оценка, значительно отстоящая по шкале от других, может оказаться самой точной. На практике указанная проблема обычно решается путем проведения второго тура экспертизы с переводом ее в формат «круглого стола», особенно когда речь идет об оценке ОЭ, реализация которых связана с большими рисками или затратами. Экспертиза при этом может превращаться в научную дискуссию со всеми вытекающими последствиями, однако в итоге часто удается достигнуть консенсуса или же максимально сблизить оценки.

Классификация экспертов

Классификация экспертов означает их ранжирование по «весу», для чего каждому эксперту, входящему в ЭГ, присваивается определенный ранг, количественно выражаемый весовым коэффициентом. В случае расхождения частных оценок применение метода классификации позволяет сформировать групповую оценку двумя способами:

- принятием в качестве групповой оценки эксперта, имеющего наивысший ранг;
- вычислением *средневзвешенной* оценки, получаемой путем корректировки частных оценок с учетом весовых коэффициентов экспертов и последующим усреднением этих оценок.

Главной трудностью при выборе данного метода является отсутствие научно обоснованных, лишенных субъективизма методик назначения экспертам весовых коэффициентов.

Простой механический подсчет количества научных публикаций, патентов, выполненных ранее научно-исследовательских работ, индексов цитируемости позволяет, в принципе, ранжировать экспертов по этим показателям, однако без их качественного анализа полученные рейтинги могут оказаться ошибочными. Анализ научных достижений экспертов помимо значительных временных затрат предъявляет высокие требования к профессиональной подготовке администраторов Реестра, предполагающей наличие глубоких знаний в различных областях науки, что на практике недостижимо.

Ранжирование экспертов на основе формальных признаков (ученая степень, звание, должность, награды и премии и т. п.) таит в себе определенный произвол и также чревато ошибками [7]. Другая проблема заключается в том, что сама идея корректировки экспертных оценок в сторону их завышения или занижения путем умножения на весовые коэффициенты экспертов представляется спорной. Очевидно, что задача поиска эффективных алгоритмов ранжирования экспертов на основе всей совокупности данных об эксперте, имеющих в распоряжении администраторов Реестра, актуальна и требует решения.

Усреднение мнений экспертов внутри группы

Усреднение является наиболее широко применяемым методом обработки частных оценок, при этом в качестве средней (групповой) оценки в зависимости от типа оценочной шкалы могут быть выбраны различные средние (мода, медиана, среднее арифметическое, среднее геометрическое, среднее гармоническое, среднее степенное или среднее по Колмогорову) частных оценок. Свойства средних оценок подробно описаны в соответствующей литературе и в настоящей статье не рассматриваются. Достаточно сказать, что все указанные средние оценки лежат между минимумом и максимумом из чисел имеющегося ряда. Для целей настоящего исследования наибольший интерес представляют две средние оценки: медиана и среднее арифметическое.

Медиана – такое число в числовом ряду, что половина из элементов ряда больше него, а другая половина – меньше. Например, для схемы экспертизы с тремя экспертами медианой является оценка, имеющая среднее значение из трех. Медиана – наилучшая (и единственная) средняя оценка в номинальных и порядковых шкалах качественных признаков. Для экспертных оценок, имеющих нечисловую природу, удобной средней оценкой является медиана Кемени – Снелла [8], часто называемая просто медианой Кемени. Медиана Кемени предложена в качестве средней оценки группы в тех случаях, когда частные оценки представляет собой не числа, а ранжировки, т. е. некоторые последовательности, упорядоченные экспертами по их предпочтениям. С математической точки зрения вычисление медианы Кемени является задачей целочисленного программирования и сводится к отысканию среди полученных от экспертов упорядочений такого, для которого сумма расстояний до оценок других экспертов будет минимальной. Найденное таким образом упорядочение и называют медианой Кемени.

Вычисление *среднего арифметического* частных оценок является самым простым способом получения средней оценки в том предположении, что частные оценки представляют собой случайные числа. Однако из теории [9] известно, что данная оценка является наилучшей только для случайной величины, имеющей нормальное (гауссово) распределение.

В тех же случаях, когда распределение случайной величины неизвестно, нельзя утверждать, что данная оценка будет наилучшей.

Экспертизы, выполняемые в ИС ФРЭ, имеют характерные особенности, к основным из которых относятся:

- малочисленность экспертных групп (максимум три человека), что обусловлено, главным образом, бюджетными ограничениями;
- проведение экспертизы в один этап;
- использование анкет, предполагающих «оцифровку» ответов эксперта и представление их в виде балльных оценок по показателям проекта и по проекту в целом;
- применение метода ранжирования экспертов только на этапе формирования ЭГ (присвоенные экспертам весовые коэффициенты не учитываются при обработке экспертных оценок; считается, что в рамках одной ЭГ все эксперты имеют одинаковый ранг).

Указанная специфика ЭАИ накладывает ограничения на выбор метода обработки экспертных оценок: малочисленность ЭГ не позволяет выделять в них согласованные подгруппы, а классификация экспертов также не находит применения по причинам, указанным выше.

Таким образом, усреднение мнений экспертов внутри ЭГ остается пока единственным методом обработки экспертных оценок, реализуемым на текущем этапе развития системы экспертиз Реестра.

Логика оценок, экспертные анкеты и оценочные шкалы

При выставлении оценок эксперт анализирует материалы ОЭ, включающие его формализованное описание и различные приложения справочного характера. В зависимости от уровня компетентности и под влиянием различных факторов случайного характера оценки эксперта могут иметь то или иное отклонение от «истинных» значений оцениваемых показателей. Разница между «истинным» значением показателя и его оценкой образует ошибку экспертизы.

Возникает вопрос: что считать «истинной» оценкой проекта?

С точки зрения традиционной логики, экспертная оценка является суждением, отражающим знания эксперта об оцениваемом ОЭ, и, как любое суждение, может быть истинной или ложной. В современной логике вместо термина «суждение» обычно используют термин «высказывание», обозначающий грамматически правильное предложение, взятое вместе с выражаемым им смыслом. В нашем случае мы имеем дело с особым типом высказываний, а именно – с *оценочными высказываниями*. Понятийно оценочные высказывания противопоставляются дескриптивным (описательным) высказываниям. Оценка является выражением ценностного отношения утверждения к объекту, противоположного описательному, или истинностному, отношению [10]. Логическая структура и логические связи оценочных высказываний изучаются *логикой оценок*, являющейся ветвью модальной логики и слагающейся из логики абсолютных оценок и логики сравнительных оценок (логики предпочтений). Логика абсолютных оценок оперирует такими аксиологическими модальностями, как «хорошо», «плохо», «безразлично», в то время как логика предпочтений использует понятия «лучше», «хуже», «равноценно». Логика абсолютных оценок и логика предпочтений не сводимы друг к другу и являются самостоятельными разделами логики оценок. «Лучше» не определимо через «хорошо»: лучше один другого могут быть и два «хороших», и два «плохих» ОЭ.

Принципиально важным является то, что оценочное высказывание не является ни истинным, ни ложным. Истина характеризует отношения между описательным высказыванием и действительностью; оценки же не являются описаниями. Они могут характеризоваться как целесообразные, эффективные, разумные, обоснованные и т. п., но не как «истинные» или «ложные». Принимаемый логикой оценок *принцип Юма* отвергает выводимость оценок из описаний и описаний из оценок. Из этого следует вполне практический вывод: при обработке экспертных оценок ни одно экспертное мнение нельзя отбрасывать как заведомо неправильное, так же как и считать чье-то мнение единственно верным. Все оценки имеют право на существование и в этом смысле равноценны.

В соответствии с принципом *аксиологической полноты* поступающий на экспертизу ОЭ может быть «хорошим», «плохим» или «безразличным». Однако этот принцип справедлив только в случае предположения, что множество ОЭ, о ценности которых имеется определенное представление, совпадает с множеством всех ОЭ.

Очевидно, что, например, для заявок на выполнение научно-технических и инновационных проектов такое предположение не всегда оправданно, поскольку данные проекты ориентированы, как правило, на получение результатов, не существующих на момент проведения экспертизы. Не будет преувеличением сказать, что материалы таких заявок, дескриптивные по форме, по сути являются оценочными, поскольку отражают взгляды их авторов на результаты, планируемые к получению в будущем. В случае экспертизы научно-технических отчетов дело обстоит проще: здесь принцип *аксиологической полноты* соблюдается, и экспертная оценка может быть эффективной.

Другой принцип – принцип *аксиологической непротиворечивости* – требует внутренней непротиворечивости самой системы принятых оценок. Другими словами, противоречащие друг другу состояния не могут быть вместе «хорошими» или «плохими». К сожалению, реальные множества оценок, используемые в экспертных анкетах, нередко непоследовательны, что ставит в тупик экспертов и повышает вероятность экспертной ошибки. Например, в ряде экспертных анкет большая научная новизна результатов предусматривает и более высокие баллы, и в то же время высокие баллы выставляются за отсутствие рисков неполучения запланированных результатов. Здесь мы сталкиваемся с очевидным противоречием: известно, что проекты, в которых отсутствуют риски, как правило, не содержат новизны, и наоборот. Однако формально и те и другие проекты могут набрать одинаковое количество баллов, что вызовет трудности при определении лучшего из проектов.

Подлежащий экспертизе ОЭ, имеющий вид формализованного описания, с логической точки зрения есть набор дескриптивных высказываний, которые могут соответствовать или не соответствовать действительности, т. е. быть истинными или ложными. Задача эксперта заключается в проверке на основе своих компетенций истинности этих высказываний и выражении к ним своего ценностного отношения в виде оценочных высказываний по основаниям, устанавливаемым экспертной анкетой. В сложившейся экспертной практике возможные оценочные высказывания (ответы эксперта) «оцифровываются», т. е. каждому из них ставится в соответствие определенное количество баллов. Такая система оценки удобна при проведении конкурсов, когда из множества проектов необходимо выбрать лучшие для их последующей реализации.

Среди множества оценок, которые могут быть поставлены ОЭ по заданной измерительной шкале, существует оценка, наиболее точно характеризующая ОЭ. Эту гипотетическую оценку, значение которой неизвестно, мы условно и называем «истинной». Если бы имелась возможность привлечь к экспертизе проекта всех существующих специалистов в тематической области ОЭ, то полученное множество частных оценок представляло бы собой генеральную совокупность, по которой можно было бы определить теоретическое распределение оценки и его параметры. Тогда за «истинное» значение оценки можно было бы принять какую-либо характеристику *центра распределения* – моду, медиану или математическое ожидание [11]. Однако на практике оценку ОЭ приходится формировать по ограниченной выборке из генеральной совокупности, чаще всего по выборке, состоящей из трех элементов. При таком малом объеме выборки провести корректный статистический анализ в целях определения теоретического распределения оценки не представляется возможным.

Например, проверка по стандартным методикам ГОСТ Р 8.736–2011 гипотезы о нормальности распределения требует наличия выборки объемом как минимум 15 элементов [12].

Распределение итоговых экспертных оценок можно определить косвенным путем на основе анализа оценочных шкал, используемых в экспертных анкетах. В табл. 1 представлены оценочные шкалы некоторых типовых анкет, по которым с использованием ИС ФРЭ проводились ЭАИ в период 2014–2020 гг.

Таблица 1

Оценочные шкалы типовых экспертных анкет

Обозначение ЭАИ	Кол-во проектов	Объект экспертизы	Оцениваемые показатели	
			Кол-во	Шкала (баллы)
ЭАИ-1	99	Научно-исследовательские проекты, предлагаемые ведущими университетами Российской Федерации – участниками программы повышения конкурентоспособности вузов среди ведущих мировых научно-образовательных центров	1 14	0; 5 ¹ 0–5 ²
ЭАИ-2	412	Конкурсные заявки на реализацию проектов в рамках Федеральной целевой программы (ФЦП) «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы», мероприятие 1.3	2 1 2	0–3 0–5 0–7
ЭАИ-3	122	Конкурсные заявки на реализацию проектов в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы», мероприятие 1.4	2 3 2	0–2 0–3 0–5
ЭАИ-4	556	Проекты, отобранные для участия в конкурсе на финансирование в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы», мероприятия 1.2–1.4	9 2 1 1 1	0–1 0–2 0–3 0–4 0–5
ЭАИ-5	31	Проекты, отобранные для участия в конкурсе на финансирование в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы» по мероприятиям 2.1, 2.2 (критерии 1, 2)	1 1 1 1 1 1 4 1	0–1 0–2 0–3 0; 2–3 0; 1; 2; 4 0; 2; 4 0; 3; 5 0; 3; 5; 8
ЭАИ-6	60	Проекты, заявленные на участие в конкурсном отборе на предоставление грантов на государственную поддержку центров национальной технологической инициативы на базе образовательных организаций высшего образования и научных организаций	1 7 1 9 4 7 1	0–1 0–2 0,5–2,0 0–3 0–4 0–5 1–5
ЭАИ-7	31	Проекты, отобранные для участия в конкурсе на финансирование в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы» по мероприятиям 2.1, 2.2 (критерий 3)	1 1 3 1 1 1 1 1 1	0–1 0–2 0; 1; 3 0; 1; 3; 5 0; 1; 2; 4; 5; 7 0; 2; 4; 5; 7 0; 2; 4 (+2) (+2) (+2) ³ 0; 1; 2; 4; 6; 8 0; 4 (+2) (+2) (+2)

¹ По этой шкале показатель может получить оценку 0 или 5 баллов.

² По этой шкале показатель может получить оценку от 0 до 5 баллов.

³ Элемент шкалы (+2) означает, что если поставлена оценка 4 балла, то эксперт может трижды добавить к ней дополнительно по 2 балла при соответствии проекта определенным условиям.

Продолжение таблицы 1

Обозначение ЭАИ	Кол-во проектов	Объект экспертизы	Оцениваемые показатели	
			Кол-во	Шкала (баллы)
ЭАИ-8	1730	Отчеты о результативности научной деятельности за 5 лет организаций, подведомственных Минобрнауки России	1	0; 5; 10
			2	0; 5; 15
			1	0; 5; 15; 20; 30
			1	0; 5; 15; 30
			1	0; 5; 20; 30 (+10)
			1	0; 10; 15
			1	0; 10; 15 (+5)
			1	0; 10; 15 (+10)
			2	0; 10; 20 (+10)
ЭАИ-9	1661	Проекты, заявленные на конкурс на получение грантов Президента Российской Федерации для государственной поддержки молодых российских ученых	1	0; 1 (+1) (+1) (+2)
			1	0; 1; 5 (+5)
			1	0; 1; 5; 10; 20 (+5) (+5)
			1	0; 2 (+3)
			1	0; 2; 5 (+3) (+7)
			1	0; 5 (+5) (+5)
			1	0; 10
ЭАИ-10	132	Конкурсные заявки на соискание стипендии Президента Российской Федерации для молодых ученых и аспирантов, осуществляющих перспективные научные исследования и разработки по приоритетным направлениям модернизации российской экономики	1	0; 0,7; 1,0
			1	0,5; 1,0
			2	1–5
			1	1–7
			1	1–8
			1	1–10
			1	1–15
			1	1–20
ЭАИ-11	913	Отчетные материалов получателей грантов Президента Российской Федерации для государственной поддержки молодых российских ученых	1	0; 1; 5; 20 (+5) (+5)
			1	0; 5
			1	0; 5 (+5) (+5)
			1	0; 5 (+10)
			1	0; 5; 15
			1	0; 6; (+1) (+1) (+2)
ЭАИ-12	110	Конкурсные заявки в рамках реализации мероприятия «Обучение граждан по программам непрерывного образования в образовательных организациях, реализующих дополнительные образовательные программы и программы профессионального обучения» Федерального проекта «Новые возможности для каждого» Национального проекта «Образование»	1	0; 10
			6	0; 20
			5	0; 30
			4	0; 50
			2	0; 100
			2	0; 200
			ЭАИ-13	75
5	0; 7			
10	0; 10			
2	0; 17			
2	0; 18			
2	0; 20			
1	0; 30			
1	0; 35			

Окончание таблицы 1

Обозначение ЭАИ	Кол-во проектов	Объект экспертизы	Оцениваемые показатели	
			Кол-во	Шкала (баллы)
ЭАИ-14	141	Конкурсные заявки на право получения грантов Президента Российской Федерации ведущим научным школам в 2020 г.	1	0; 1; 5; 10; 20 (+5) (+5)
			1	0; 1; 5 (+5) (+5)
			1	0; 2 (+3) (+5)
			1	0; 2; 5 (+3) (+7)
			1	0; 5
			1	0; 5 (+5) (+5)
			1	0; 10
ЭАИ-15	14	Представления на соискание Премии союзного государства в области науки и техники российским и белорусским ученым и специалистам	3	0; 1; 3
			1	0; 2; 3
			1	0; 2; 3; 4; 5
			4	0; 3; 5
			1	0; 5
ЭАИ-16	269	Конкурсные заявки на получение персональных стипендий им. Ж.И. Алферова для молодых ученых в области физики и нанотехнологий	1	0; 5 (+2) (+3)
			1	0; 5 (+5)
			2	0; 5 (+10)
			1	0; 5; 10
			1	0; 5; 15 (+10) (+5) (+5)
			1	(+5)

Как видно из табл. 1, экспертные анкеты содержат от 5 до 30 и более показателей, а оценочные шкалы содержат от 2 до 7 значений. Для облегчения работы экспертов анкеты составляются таким образом, что, оценивая тот или иной показатель, эксперт должен выбрать из списка один или несколько возможных вариантов ответа, каждому из которых поставлено в соответствие определенное количество баллов (рис. 1).

1. ОПЫТ И ДОСТИЖЕНИЯ СОИСКАТЕЛЯ В НАУЧНОМ НАПРАВЛЕНИИ ПРОЕКТА
(за 3 года, предшествующих году проведения конкурса)

1.1. Опыт участия соискателя в научных исследованиях

Наличие завершенных научных и научно-исследовательских работ с участием соискателя

0 - Отсутствуют научные работы, в которых соискатель является исполнителем, либо оценка затруднена

1 - Представлены сведения менее, чем о пяти НИР или НИОКТР, в которых соискатель является исполнителем

5 - Представлены сведения о пяти и более НИР или НИОКТР, в которых соискатель является исполнителем

+5 - Представлены сведения о НИР и НИОКТР, в которых соискатель является научным руководителем или ответственным исполнителем

Рис. 1. Фрагмент анкеты с «оцифрованными» вариантами ответа

Шкалы отличаются многообразием даже в пределах одной экспертной анкеты, при этом устанавливаемые в них значения (баллы) свидетельствуют о достаточно произвольном подходе к «оцифровке» ответов на вопросы анкеты. По своему типу используемые в анкетах

шкалы являются порядковыми, а оценки по показателям представляют собой дискретные величины. Поскольку невозможно заранее указать, какую оценку поставит тот или иной эксперт по тому или иному показателю, можно считать данные оценки дискретными случайными величинами и применять для их анализа методы теории вероятностей и математической статистики.

В примере, приведенном на рис. 1, оцениваемым показателем является опыт участия соискателя в научных исследованиях, а критерием оценки – наличие научных и научно-исследовательских работ с участием соискателя и их количество за последние три года. Видно, что использованная для оценки данного показателя измерительная шкала является порядковой, значения в ней расположены по возрастанию, а единица измерения – балл. Производить какие-либо арифметические действия с элементами шкалы, например вычислять соотношения или разности, с логической точки зрения не имеет смысла. Например, соискатель, участвовавший в качестве исполнителя в четырех научно-исследовательских работах (НИР), получит по данной анкете оценку в 1 балл, а исполнитель пяти НИР – уже 5 баллов. Но это не означает, что последний в 5 раз лучше первого. Разница в оценках этих соискателей составляет 4 балла, но точно такая же разница была бы в случае, если бы второй соискатель принял участие в выполнении не 5, а, например, 15 НИР. По логике анкеты также выходит, что заявитель, выполнивший в качестве научного руководителя одну НИР, и научный руководитель четырех НИР равноценны (оба получают по 6 баллов), что, очевидно, нельзя признать правильным. Подобные нестыковки, к сожалению, встречаются в экспертных анкетах довольно часто, и рассмотренный пример показывает, что абстрактные количественные оценки не всегда точно отражают измеряемые характеристики объекта экспертизы.

В 2013 г. ФГБНУ НИИ РИНКЦЭ была предпринята попытка использовать при проведении экспертизы научных проектов нечисловые показатели. Для этой цели была разработана экспертная анкета (рис. 2), в которой для оценки проекта использовались биполярные шкалы качественных признаков: «соответствует/не соответствует»; «да/нет»; «скорее да, чем нет/скорее нет, чем да»; «имеется/отсутствует» и т. д. Анкета прошла апробацию в двух тестовых ЭАИ (5 и 20 проектов соответственно) и получила положительные отзывы экспертов, участвовавших в ее тестировании. Новшеством являлось также то, что после заполнения анкеты экспертное заключение формировалось автоматически в виде связного текста, что облегчало работу экспертов и сокращало общее время экспертизы. Однако идея не получила дальнейшего развития. Связано это было с тем, что на этапе принятия решения проекты, рекомендованные экспертами к поддержке, приходилось вновь рассматривать и каким-то образом рейтинговать в целях выбора из них лучших. В отсутствие количественных оценок сделать это было крайне сложно. Поэтому начиная с 2014 г. при проведении ЭАИ стала вновь применяться и применяется по настоящее время балльная система оценок.

Многообразие используемых в анкетах оценочных шкал определяется, главным образом, спецификой конкретных ЭАИ, и универсального решения до настоящего времени здесь нет. Однако задача разработки для типовых видов экспертизы стандартизованных анкет, использующих оценочные шкалы, основанные на интуитивно понятной метрике, давно назрела.

Статистические характеристики экспертных оценок

В предыдущем параграфе было показано, что оценки по каждому показателю экспертной анкеты можно рассматривать как дискретные случайные величины с неизвестными в общем случае законами распределения. Выяснение этих законов и их параметров невозможно ввиду малых объемов выборок. Однако поскольку сумма оценок по всем показателям образует общую оценку ОЭ в баллах, то согласно тереме Ляпунова эта общая оценка также есть случайное число, которое, приняв в качестве допущения независимость оценок по показателям, можно уверенно считать распределенным нормально [13]. Данный вывод позволяет однозначно выбрать наилучшую среднюю оценку ОЭ, которую можно считать мнением ЭГ.

Из теории обработки наблюдений [14] известно, что наилучшей точечной оценкой наиболее вероятного значения наблюдаемой величины x , полученной по нормально распределенной выборке $x_1, x_2, \dots, x_i, \dots, x_n$, является выборочное среднее $x_{\text{ср}}$, определяемое по формуле:

$$x_{\text{ср}} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (1)$$

Следовательно, наилучшей групповой оценкой ОЭ будет среднее арифметическое частных оценок. Необходимо подчеркнуть, что вычислять подобным образом средние оценки по отдельным показателям нельзя, так как в порядковой шкале в качестве средних можно использовать только члены вариационного ряда, в частности медиану, но нельзя использовать среднее геометрическое, среднее арифметическое и т. д. В количественных шкалах из всех обобщенных средних по А.Н. Колмогорову можно использовать: в шкале интервалов – только среднее арифметическое, а в шкале отношений – только степенные средние [6].

Оценка экономических показателей предлагаемой НИР	
Соответствует ли заявленная стоимость НИР объему и сложности предполагаемых работ?	
Заявленная стоимость НИР: 4,3 млн.руб.	
<input type="radio"/> да <input type="radio"/> нет <input checked="" type="radio"/> оценка затруднена	
Соответствует ли заявленная продолжительность НИР объему и сложности предполагаемых работ?	
Заявленная продолжительность НИР: 4 года	
<input type="radio"/> да <input checked="" type="radio"/> нет приведите Вашу оценку продолжительности НИР (лет)	
<input type="text" value="2-3 года"/>	
<input type="radio"/> оценка затруднена	
Как Вы оцениваете перспективы коммерциализации результатов НИР?	
<input type="radio"/> коммерциализация возможна в краткосрочной перспективе (2-3 года) <input checked="" type="radio"/> коммерциализация возможна в среднесрочной перспективе (4-6 лет)	

Рис. 2. Фрагмент экспертной анкеты, использующей нечисловые оценочные шкалы

Вычисленная по формуле (1) средняя оценка ОЭ ни в коей мере не может считаться его «истинной» оценкой или даже максимально к ней близкой. Все элементы выборки (частные оценки) могут оказаться далеко отстоящими от гипотетического «истинного» значения и находиться как в зоне недооценки, так и в зоне переоценки. По полученной выборке можно лишь утверждать, что, по мнению данной ЭГ, наиболее вероятным значением оценки ОЭ

является x_{cp} . Также можно утверждать, что при увеличении объема выборки, т. е. при увеличении численности ЭГ, средняя оценка будет приближаться к своему наиболее вероятному значению – математическому ожиданию.

Кроме среднего значения по выборке можно также определить среднеквадратическое отклонение экспертной оценки, характеризующее расхождение мнений экспертов внутри ЭГ. Несмещенная оценка среднеквадратического отклонения величины x_{cp} определяется по формуле Бесселя:

$$s = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - x_{cp})^2} . \quad (2)$$

Рассчитанные по формулам (1)–(2) статистические характеристики экспертных оценок, полученных группами из трех экспертов в некоторых видах ЭАИ, представлены в табл. 2.

Таблица 2

Статистические характеристики экспертных оценок

Обозначение			Экспертные оценки, балл			Статистические характеристики оценок	
ЭАИ	Группа	Проект	Эксперт 1	Эксперт 2	Эксперт 3	x_{cp}	s
ЭАИ-1	ЭГ-1	П1.1	69	50	66	61,7	5,9
		П1.2	51	70	59	60,0	5,5
	ЭГ-2	П2.1	80	70	67	72,3	3,9
		П2.2	77	80	78	78,3	0,9
	ЭГ-3	П3.1	52	55	75	60,7	7,2
		П3.2	56	47	78	60,3	9,2
ЭАИ-6	ЭГ-4	П4.1	36	89	52	59,0	15,7
		П4.2	62,5	83,5	81,5	75,8	6,7
		П4.3	73	40	67	60,0	10,1
ЭАИ-9	ЭГ-5	П5.1	83	68	73	74,7	4,4
		П5.2	62	52	62	58,7	3,3
	ЭГ-6	П6.1	75	96	55	75,3	11,8
		П6.2	86	89	22	65,7	21,9
		П6.3	65	65	28	52,7	12,3
	ЭГ-7	П7.1	88	88	85	87,0	1,0
	ЭГ-8	П8.1	84	79	84	82,3	1,7
		П8.2	46	43	30	39,7	4,9
	ЭГ-9	П9.1	84	78	77	79,7	2,2
		П9.2	79	74	78	77,0	1,5
	ЭГ-10	П10.1	94	36	81	70,3	17,6
		П10.2	87	41	80	69,3	14,3
		П10.3	29	41	56	42,0	7,8
	ЭГ-11	П11.1	89	81	81	83,7	2,7
П11.2		81	84	81	82,0	1,0	
ЭАИ-11	ЭГ-12	П12.1	45	15	36	32,0	8,9
		П12.2	35	47	31	37,7	4,8

Видно, что математическое ожидание и среднеквадратическое отклонение экспертных оценок переменны не только для разных экспертных анкет, но даже и в рамках одного ЭАИ для разных пар «экспертная группа – проект». Кроме того, не наблюдается видимой корреляции между оценками, поставленными разным ОЭ одними и теми же экспертами. Другими словами, по выборке невозможно определить, кто из входящих в ЭГ экспертов дал более точную оценку. В экспертной группе ЭГ-4 оценки «эксперта 3» являются медианами во всех трех выборках, однако это, скорее всего, случайность, так как в других группах подобная закономерность не прослеживается.

Критерии принятия решения и пороговые значения оценок

Рассмотрим критерии принятия решения и пороговые значения оценок ЭО на примере научно-технических проектов.

Оценивая проект, эксперт руководствуется определенными критериями, которые так же, как и показатели, устанавливаются экспертной анкетой. Если под показателем понимается какое-либо свойство проекта, измеримое по заданной шкале, то критерий – это правило, в соответствии с которым показатель получает ту или иную оценку. Типичным критерием является «наличие и достаточность» чего-либо, относящегося к характеристикам проекта. Это могут быть научная новизна, квалификация и опыт исполнителей, научно-технический задел, лабораторно-исследовательская база и т. п. В других случаях критерием может выступать достижение или недостижение показателем некоторого порогового значения.

В процессе выполнения ЭАИ одним из часто задаваемых экспертами вопросов является вопрос о так называемом проходном балле, т. е. о пороговом значении суммарной оценки, выше которого проект может считаться «хорошим» и рекомендоваться к реализации. По сложившейся традиции считается, что, если проект набрал больше половины максимально возможного количества баллов, его целесообразно скорее поддержать, чем отклонить. Однако здесь имеется подводный камень: если в пределах анкеты оцениваемые показатели (группы показателей) имеют одинаковую значимость, то набрать 50 % и более баллов может проект, имеющий одновременно максимально высокие баллы по одним показателям и «нулевые» или близкие к ним – по другим. Избежать подобных перекосов можно двумя способами:

– введением требования о равномерности достижения 50%-го порога всеми группами показателей;

– ранжированием показателей по степени важности (значимости) с назначением им весовых коэффициентов.

В последнем случае итоговая оценка проекта представляется *взвешенной суммой*, слагаемыми которой являются оценки по показателям, умноженные на соответствующие весовые коэффициенты этих показателей. Весовые коэффициенты показателям (группам показателей) могут устанавливаться при составлении экспертных анкет до начала ЭАИ, однако нужно учитывать то обстоятельство, что степень важности тех или иных показателей может различаться для проектов, тематически относящихся к различным областям науки и техники. В связи с этим представляется целесообразным оставить решение данного вопроса на усмотрение самих экспертов, участвующих в ЭАИ.

В простейшем случае можно предложить экспертам просто упорядочить показатели (группы показателей) в убывающем порядке по степени их важности. Полученные таким образом упорядочения (ранжировки) могут различаться, в связи с чем возникает задача нахождения «среднего» упорядочения, которое можно считать мнением ЭГ. Таким упорядочением будет являться медиана Кемени, которую можно вычислить одним из известных методов (венгерский метод, метод ветвей и границ). Используя найденное «среднее» упорядочение и присвоив его элементам весовые коэффициенты, можно затем пересчитать все оценки по отдельным показателям и получить итоговые оценки проекта как взвешенные суммы частных оценок. Данные оценки, на наш взгляд, будут более точными, чем простые суммы, что в конечном счете повысит эффективность экспертизы.

Заключение

Повышение достоверности и точности результатов экспертизы различных типов ОЭ, в том числе научно-технических и инновационных проектов, является актуальной задачей, решение которой достигается комплексом организационно-технических мер, включающих, в частности, применение научно обоснованных методов обработки экспертных оценок.

Особенностями ЭАИ, организуемых ФГБНУ НИИ РИНКЦЭ по заданиям Минобрнауки России, являются: пакетный характер экспертных заданий; большой объем пакетов (до нескольких сотен проектов), жестко фиксированные сроки, отводимые на экспертизу; проведение экспертизы в один этап; использование преимущественно балльной системы оценок. С учетом этого, а также исходя из временных и финансовых ограничений, наиболее удобной формой организации экспертизы в ИС ФРЭ является формирование экспертных групп в составе трех экспертов, работающих независимо друг от друга, с последующей обработкой экспертных оценок и формированием итоговых (сводных) экспертных заключений.

Используемые в экспертных анкетах оценочные шкалы являются порядковыми, предусматривающими оценку отдельных показателей проекта и проекта в целом в баллах. Анализ экспертных анкет показал, что требуется дальнейшее совершенствование оценочных шкал по ряду показателей в целях устранения неоднозначности оценок, выставляемых по этим показателям.

Получаемые в результате работы экспертных групп суммарные оценки проектов представляют собой нормально распределенные случайные числа, в связи с чем наилучшей средней оценкой, отражающей мнение экспертной группы, является среднее арифметическое частных оценок. Повысить точность экспертных оценок можно путем увеличения численности экспертных групп, для чего целесообразно рассмотреть различные формы привлечения к экспертизе максимально возможного числа экспертов, обладающих требуемыми компетенциями.

Необходимо продолжение исследований по уточнению критериев оценки проектов по отдельным показателям, изучение вопроса целесообразности ранжирования показателей (групп показателей) по степени их важности для получения более точных и достоверных экспертных оценок.

Статья подготовлена в рамках научно-исследовательской работы, проводимой ФГБНУ НИИ РИНКЦЭ по государственному заданию № 075-01394-20-02 на 2020 г. Министерства образования и науки Российской Федерации в сфере экспертизы по проекту 730000Ф.99.1. БВ15АА00003.

Список литературы

1. Илющенко Р.Р., Мельник П.Б. Достоверность экспертных оценок при проведении экспертизы научно-технических и инновационных проектов в информационной системе Федерального реестра экспертов // *Инноватика и экспертиза*. 2020. № 1 (29). С. 46–58.
2. Мельник П.Б. Методика формирования экспертных пулов и групп для проведения экспертно-аналитических исследований // *Инноватика и экспертиза*. 2017. № 1 (19). С. 39–54.
3. Миронов Н.А., Марышев Е.А., Дивуева Н.А. Новые подходы к организационно-техническому обеспечению экспертизы конкурсных заявок на получение государственной поддержки в виде грантов Президента Российской Федерации молодым российским ученым – кандидатам и докторам наук и ведущим научным школам // *Инноватика и экспертиза*. 2020 г. № 1 (29). С. 59–67.
4. Панкова Л.А., Петровский А.М., Шнейдерман М.В. Организация экспертизы и анализ экспертной информации. М.: Наука, 1984.
5. Орлов А.И. Экспертные оценки // *Заводская лаборатория*. 1996. № 1. Т. 62. С. 54–60.
6. Орлов А.И. Нечисловая статистика. М.: МЗ-Пресс, 2004. С. 513.
7. Мушик Э., Мюллер П. Методы принятия технических решений / пер с нем.: Н.В. Васильченко, В.А. Душский. М.: Мир, 1990. С. 208.
8. Кемени Дж., Снелл Дж. Кибернетическое моделирование. Некоторые приложения / пер. Б.Г. Миркин. М.: Советское радио, 1972.

9. Бочаров П.П., Печинкин А.В. Математическая статистика: учеб. пособие. М.: РУДН, 1994.
10. Горский Д.П., Ивин А.А., Никифоров А.Л. Краткий словарь по логике / под ред. Д.П. Горского. М.: Просвещение, 1991. С. 208.
11. Сигорский В.П. Математический аппарат инженера. Киев: Техніка, 1975.
12. ГОСТ Р 8.736–2011. Государственная система обеспечения единства измерений. Измерения прямые многократные. Методы обработки результатов измерений. Основные положения. М., 2019. ОКС 17.020, ОКСТУ 0008.
13. Прохоров Ю.В. Математический энциклопедический словарь. М.: Советская энциклопедия, 1988.
14. Большаков В.Д. Теория ошибок наблюдений с основами теории вероятностей. М.: Недра, 1965.

References

1. Ilyushchenko R.R., Melnik P.B. (2020) *Dostovernost' ekspertnykh otsenok pri provedenii ekspertizy nauchno-tekhnicheskikh i innova-tсионnykh projektov v informatsionnoy sisteme Federal'nogo reestra ekspertov* [The reliability of expert assessments in the examination of scientific, technological and innovative projects in the information system of the Federal Roster of Experts] *Innovatika i ekspertiza* [Innovation and Expert Examination]. No. 1 (29). P. 46–58.
2. Melnik P.B. (2017) *Metodika formirovaniya ekspertnykh pulov i grupp dlya provedeniya ekspertno-analiticheskikh issledovaniy* [Methodology for the formation of expert pools and groups for expert and analytical research] *Innovatika i ekspertiza* [Innovation and Expert Examination]. No. 1 (19). P. 39–54.
3. Mironov N.A., Maryshev E.A., Divueva N.A. (2020) *Novye podkhody k organizatsionno-tekhnicheskomu obespecheniyu ekspertizy konkursnykh zayavok na poluchenie gosudarstvennoy podderzhki v vide grantov Prezidenta Rossiyskoy Federatsii molodym rossiyskim uchenym – kandidatam i doktoram nauk i vedushchim nauchnym shkolam* [New approaches to the organizational and technical support of the examination of competition applications for receiving state support in the form of grants from the President of the Russian Federation to young Russian scientists – Doctors and Ph.D.-s and leading scientific schools] *Innovatika i ekspertiza* [Innovation and Expert Examination]. No. 1 (29). P. 59–67.
4. Pankova L.A., Petrovsky A.M., Shneiderman M.V. (1984) *Organizatsiya ekspertizy i analiz ekspertnoy informatsii* [Organization of expert examination and analysis of expert information] *Nauka* [Nauka]. Moscow.
5. Orlov A.I. (1996) *Ekspertnye otsenki* [Expert estimates] *Zavodskaya laboratoriya* [Factory laboratory]. No. 1. T. 62. P. 54–60.
6. Orlov A.I. (2004) *Nechislovaya statistika* [Non-numeric statistics] *MZ-Press* [MZ-Press]. Moscow. P. 513.
7. Mushik E., Muller P. (1990) *Metody prinyatiya tekhnicheskikh resheniy. Per s nem. N.V. Vasil'chenko, V.A. Dushskiy* [Methods of making technical decisions. Translated from German: N.V. Vasilchenko, V.A. Dushsky] *Mir* [Mir]. Moscow. P. 208.
8. Kemeny J., Snell J. (1972) *Kiberneticheskoe modelirovanie. Nekotorye prilozheniya. Per. B.G. Mirkin*. [Cybernetic modeling. Some applications. Translated by B.G. Mirkin] *Sovetskoe radio* [Soviet radio]. Moscow.
9. Bocharov P.P., Pechinkin A.V. (1994) *Matematicheskaya statistika: ucheb. posobie* [Mathematical statistics: textbook] *RUDN* [RUDN]. Moscow.
10. Gorskiy D.P., Ivin A.A., Nikiforov A.L. (1991) *Kratkiy slovar' po logike. Pod red. D.P. Gorskogo* [A short dictionary of logic. Ed. D.P. Gorsky] *Prosveshchenie* [Education]. Moscow. P. 208.
11. Sigorskiy V.P. (1975) *Matematicheskii apparat inzhenera* [Engineer's mathematical apparatus] *Tekhnika* [Technika]. Kiev.
12. ГОСТ R 8.736–2011. (2019) *Gosudarstvennaya sistema obespecheniya edinstva izmereniy. Izmereniya pryamye mnogokratnye. Metody obrabotki rezul'tatov izmereniy. Osnovnye polozheniya* [State system for ensuring uniformity of measurements. Multiple direct measurements. Methods for processing measurement results. Basic provisions] *OKS 17.020, OKSTU 0008*. Moscow.
13. Prokhorov Yu.V. (1988) *Matematicheskii entsiklopedicheskiy slovar'* [Mathematical encyclopedic dictionary] *Sovetskaya entsiklopediya* [Soviet Encyclopedia]. Moscow.
14. Bolshakov V.D. (1965) *Teoriya oshibok nablyudeniy s osnovami teorii veroyatnostey* [The theory of observation errors with the foundations of the theory of probability] *Nedra* [NEDRA]. Moscow.