

УДК 519.816 : 303.832.24

АНАЛИЗ КАЧЕСТВА И ДОСТОВЕРНОСТИ ЭКСПЕРТИЗ ПРОЕКТОВ, ВЫПОЛНЕННЫХ ПРЕДСТАВИТЕЛЯМИ ФЕДЕРАЛЬНОГО РЕЕСТРА ЭКСПЕРТОВ (ФРЭ)*

О.В. Викулов¹,
Ю.А. Шамсутдинов¹,
Ю.В. Капральный²

¹ ГОСУДАРСТВЕННЫЙ ЦЕНТР ЭКСПЕРТИЗЫ
В СФЕРЕ НАУКИ И ИННОВАЦИЙ НИИ РИНКЦЭ

² НАЦИОНАЛЬНЫЙ ФОНД
СОЦИАЛЬНО-ЭКОНОМИЧЕСКОГО
РАЗВИТИЯ РЕГИОНОВ

В статье рассмотрен вопрос оценки качества и достоверности экспертиз проектов, проводимых методом оценочного анкетирования. На реальном примере проведен сравнительный анализ полученных статистик и предложен критерий достоверности экспертиз, а также метод разграничения экспертных заключений на группы.

Ключевые слова: экспертиза научных проектов, компетентность эксперта, достоверность экспертиз, критерий достоверности, статистический закон распределения, выборочная статистика.

ИСХОДНЫЕ ПОЛОЖЕНИЯ

Оценка была проведена на конкретном примере по результатам экспертизы множества проектов (168 проектов из области «Информационные технологии» и 500 проектов по междисциплинарным темам). Каждый проект оценивался независимо двумя экспертами, которые оценивали их путем выбора баллов из предлагаемых списков и, в итоге, формировали суммарный балл, ограниченный диапазоном 0–100. Итоговая оценка по проекту определялась как среднее арифметическое двух оценок. То есть, проект оценивался по единственному параметру – итоговому баллу. Учитывая, что целью экспертизы является выбор ограниченного количества проектов из всех рассмотренных на основе их балльной оценки, данный анализ

* Статья подготовлена в рамках государственного задания №2015/Н7 Министерства образования и науки РФ

THE ANALYSIS OF QUALITY AND
RELIABILITY OF EXAMINATIONS
OF THE SCIENTIFIC PROJECTS
EXECUTED BY REPRESENTATIVES
OF THE FEDERAL REGISTER
OF EXPERTS (FRE)

O.V. VIKULOV, Y.A. SHAMSUTDINOV,
Y.V. KAPRALNY

In article the question of an assessment of quality and reliability of the expertizes of projects which are carried out by method of estimated questioning is considered. On a real example the comparative analysis received the statistician is carried out and the criterion of reliability of examinations, and also a method of differentiation of expert opinions on groups is offered.

KEYWORDS: examination of scientific projects, competence of the expert, reliability of examinations, criterion of reliability, statistical law of distribution, selective statistics.

имеет целью выявление случаев либо некачественно проведенной экспертизы, либо необходимости выполнения по ним дополнительных уточняющих работ. Это позволило на финальном этапе оценивать результаты конкурса более объективно, минимизировав возможные ошибки в оценке проектов.

ИСХОДНЫЕ ПРЕДПОСЫЛКИ ДЛЯ ВЫРАБОТКИ ПРОЦЕДУРЫ АНАЛИЗА

Проекты, представленные на экспертизу, имели большой разброс тем. Эксперты, выполнявшие экспертизу конкретных проектов, обладали знаниями, зачастую не точно соответствующими темам проектов. Их отношение к работе было различным и трудно прогнозируемым и, тем более, трудно поддающимся количественной оценке. Естественное различие в качестве проектов сопровождалось таким же различием в уровне знаний экспертов и их специфики оценки вообще (например, при одинаковом отношении к каче-

ству проекта у двух экспертов они могли оценить ее в баллах по разному, так как «хорошо» для одного это 70 баллов, для другого 90, и т. д.).

Все перечисленные факторы не позволяли априорно иметь критерий или эталон проекта, относительно оценки которого можно было строить шкалу сравнения проектов по качеству, не говоря об оценке качества их экспертизы. Такой критерий должен был быть выработан «внутри» полученных результатов, что и было сделано.

Основной задачей экспертизы по данным первичных результатов являлась оценка ее достоверности, имеющая цель определить границы, разделяющие проекты на достоверно оцененные и проекты, по которым достоверность вызывала сомнение, и которые требовали проведения дополнительной экспертизы. Ключевым моментом данного исследования стал выбор метода и определение в рамках его критерия оценки достоверности экспертизы.

ВЫБОР МЕТОДА АНАЛИЗА ЭКСПЕРТИЗ

Для анализа данных был применен метод, базирующийся на теории проверки статистических гипотез. В силу специфики природы исходных данных, связанной с участием в процессе человека, ее подходы были использованы нетрадиционным образом.

При классическом подходе проверка гипотезы строится на вычислении критерия (статистики) и сравнения его с критическим, по результатам чего производится принятие гипотезы, в рассматриваемом здесь приложении – отнесение выборки экспертиз к «достоверно оцененным». Данный подход, в силу сложившейся традиции его применения, предполагает достаточно «жесткое» разделение данных (и только на две группы), поскольку, значимость, как правило, выбирается на уровне (1–5) %, то есть, (95–99) процентного уровня достоверности принятого решения. Кроме того, сам выбор уровня значимости достаточно произволен. Однако статистические параметры результатов экспертиз, в силу участия в них человека и, как следствие, большой варибельности в полученных данных, могут существенно расходиться по заданным критериям с требованиями к их однородности (то есть, соответствию выдвинутой гипотезы, в классическом понимании). Как частный случай может оказаться так, что они все, или их большая часть, могут не попасть в группу «достоверно оцененных», что приведет к некорректному или завышено строгому решению задачи разделения. Целесообразно применение более «мягкого» подхода. С целью учета этого фактора был применен метод, при котором оценке подвергались не сами значения критериев, а характер и величины их варибельности. Поиск характерных участков, где этот параметр существенно менялся, а значит, менялся и характер данных, а именно, их однородность и связанная с этим параметром достоверность экспертизы,

О. В. ВИКУЛОВ,
Ю. А. ШАМСУТДИНОВ, Ю. В. КАПРАЛЬНЫЙ
АНАЛИЗ КАЧЕСТВА И ДОСТОВЕРНОСТИ ЭКСПЕРТИЗ
ПРОЕКТОВ, ВЫПОЛНЕННЫХ ПРЕДСТАВИТЕЛЯМИ
ФЕДЕРАЛЬНОГО РЕЕСТРА ЭКСПЕРТОВ (ФРЭ)

позволяет определять границы не на основе проверки их соответствия «абстрактным» (табличным) критическим значениям, а на основе критических значений, которые сформированы, образно говоря, самими данными и имеют, таким образом, «естественное происхождение» и непосредственную привязку к конкретному массиву данных.

Вторым отличием от классической теории служило то, что в методе было предложено разделение экспертиз на 3 группы – «достоверно оцененные», «условно достоверно оцененные» и «недостоверно оцененные», что хорошо согласовано с современными психологическими представлениями о процессе принятия решения экспертом, в котором также предполагаются три различные интерпретации принятого им решения. Соответственно, такое представление данных более объективно отражало специфику работы эксперта и должно было быть учтено. Уместно пояснить данный аспект более детально.

МОДЕЛЬ ЭКСПЕРТА

Представление эксперта-аналитика [1] на базе модели когнитивного управления, которая близка к традиционному представлению эксперта, принятому в теории экспертных оценок или науке управления, основана на формальных методах. Эта модель рассматривала в качестве схемы представления знаний человека (эксперта), некие понятия и теоретические модели, явно или неявно «навязанные» ему в процессе познаний. Однако это не всегда согласуется с практической деятельностью экспертов, так как никак не учитывает специфических черт деятельности эксперта, в частности, его компетентности и успешности, что было отмечено в работе [2], в которой данная модель была изменена с учетом следующего.

Специалист является экспертом в некоторой области, если он:

- обладает совокупностью эвристик, рутинное применение которых в стереотипных для него ситуациях обеспечивает относительно успешное решение задач;
- способен с относительно высокой степенью успеха отличать нестереотипные (и пограничные) ситуации из своей области компетенции и в случаях, когда его стереотипные эвристики неадекватны ситуации – осуществлять творческое решение задачи.

В предложенной модели, таким образом, предполагается, что знание, которое формирует эксперт, действующий в области стереотипных для него ситуаций, является наиболее достоверным, тогда как в случае нестереотипных ситуаций достоверность может снижаться. Выделяется и третий аспект. За рамками областей более или менее результативной (по достоверности) деятельности, в которой человек выступает как эксперт благодаря репродуктивному или творческому

переносу профессиональных знаний и умений, лежат области с меньшей достоверностью результатов, в которых возрастает роль немотивированных эвристик. Грубо, это область «обыденного мышления», когда эксперт уже перестает быть экспертом. Применению неадекватных эвристик способствует [3]:

- недостаток времени или познавательных ресурсов для исчерпывающего анализа ситуации, неопределенность, информационная сложность;
- специфика действий стереотипов, запускаемых автоматически без достаточных оснований (предполагается механизм условного рефлекс);
- «факторы веры» – принцип обязательства или занятой позиции, принцип авторитета, научные парадигмы и нормы научных направлений и школ.

Таким образом, модельное разбиение областей деятельности эксперта по областям, отличающимся достоверностью результатов, можно представить следующим образом (рис. 1).

Такое представление процесса принятия решения экспертом логично транслировать на результат его деятельности и анализировать качество его экспертиз исходя из трех возможных вариантов – «достоверно оцененные» (аналог «рутинных действий»), «условно достоверно оцененные» (аналог «творческой деятельности») и «недостоверно оцененные» (аналог «обыденного мышления»).

Данный подход был учтен при использовании предложенного метода.

ВЫБОР ПАРАМЕТРА ДЛЯ ОЦЕНКИ КАЧЕСТВА ЭКСПЕРТИЗЫ

Баллы экспертов, включая средние баллы, не могли выступать в качестве параметра оценки достоверности экспертизы в силу того, что они, по определению, являются критериями оценки качества самой работы и не имеют отношения к качеству ее экспертизы.

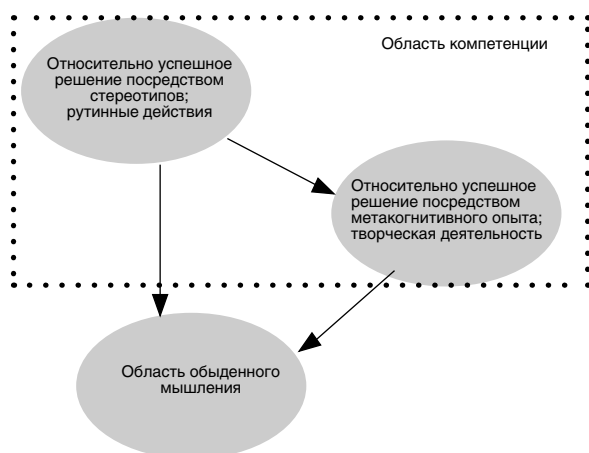


Рис. 1.

Модель компетентности эксперта

Оценочным параметром, таким образом, может выступать только расхождение в оценках экспертов по каждой конкретной работе – Δ , являющейся случайной величиной. Допустимость статистической обработки этого параметра требует пояснения и обоснования, поскольку правомерность использования статистических методов должно предполагать использование в качестве исходной информации данных одной природы, в то время как исходные значения расхождений получены на экспертизах, выполненных как на множестве «разнородных» по теме проектах, так и с привлечением множества «разнородных» экспертов и, на первый взгляд, это выглядит разнородной информацией.

Обоснованием могут служить следующие рассуждения. Несмотря на то, что каждый проект является тематически узкоспециализированным, в совокупности с проводящими экспертизу учеными «по этой же специальности», этот тандем – «проект – пара экспертов», позволяет абстрагироваться от уникальности (узкой специализации) темы проекта, поскольку этому соответствует такая же «уникальность» специалистов. Это позволяет свести ситуацию с проведением экспертизы «множество разных проектов с привлечением множества разных экспертов» (случай разнородных данных) к ситуации проведения экспертизы «однотипных проектов однотипными экспертами» (случай однородных данных). Этот ход рассуждений можно довести до логически завершенного – выполнение «экспертизы множества проектов (все по одной тематике) двумя экспертами в области этой тематики» или выполнение «экспертизы одного проекта множеством пар экспертов, где все специалисты в области тематики проекта» (оба случая эквивалентны), когда природа однородности расхождения представляется еще более очевидной.

Таким образом, выбранный параметр расхождения допустимо использовать для оценки качества экспертиз.

ИНТЕРПРЕТАЦИЯ ПАРАМЕТРА РАСХОЖДЕНИЯ

Параметр расхождения образуется как расхождение в баллах, выставленных двумя экспертами по одному и тому же проекту. Соответственно, величина этой разницы является мерой рассогласованности мнений экспертов или, интерпретируя ее иначе, величиной «неопределенности» при принятии решения одним условным экспертом, когда он выставляет балл, равный среднearифметическому значению баллов двух экспертов, дополняя это решение величиной расхождения – Δ , выражая, таким образом, свою меру неопределенности или сомнения относительно принятого решения. В этом случае мы видим, что величина расхождения может выступать параметром, характеризующим результат принятого экспертом решения в соответствии с принятой его моделью:

- если Δ мала, условный эксперт принимал решение по «совокупности эвристик, рутинное применение которых обеспечило ему относительно успешное решение задачи» и, таким образом, эту экспертизу можно рассматривать как «достоверно оцененную»;
- если Δ среднее, условный эксперт принимал решение по «стереотипным эвристикам для неадекватной ситуации – осуществляя творческое решение задачи» и, т. о., эту экспертизу можно рассматривать как «условно достоверно оцененную»;
- если Δ большое, решение было принято на уровне «обыденного мышления» и, т. о., эту экспертизу можно рассматривать как «недостоверно оцененную».

Таким образом, параметр расхождения отражает качество экспертизы и может быть использован в методе их разграничения по данному правилу.

ОПРЕДЕЛЕНИЕ ВЗАИМОСВЯЗИ ОДНОРОДНОСТИ ДАННЫХ С КАЧЕСТВОМ ЭКСПЕРТИЗЫ И ФОРМУЛИРОВАНИЕ МЕТОДА РАЗГРАНИЧЕНИЯ ГРУПП ЭКСПЕРТИЗ

Анализ материалов статистической теории в приложении к оценке экспертизы, выполняемой на условиях, которые были сформированы в постановке задачи, не дает однозначных рекомендаций по выбору конкретных критериев и методов их применения. В этом смысле имеется определенная свобода, как в выборе критериев, так и методов. Выборку расхождений, как и выборку значений баллов экспертов, можно рассматривать в самом общем случае как произвольно распределенные случайные величины, то есть не подчиняющиеся какому-либо закону распределения (например, нормальному), поэтому целесообразным видится поиск непараметрических критериев, свободных от привязки к конкретным законам распределения. Среди таких критериев наиболее мощными считаются критерии Смирнова и Лемана-Розенблата, менее мощным, но часто используемым, Манна-Уитни (Вилкоксона) и др. [6].

Данные критерии могут быть использованы при проверке на однородность двух независимых выборок, в качестве которых могут быть рассмотрены данные оценок двух экспертов, на основе которых вычислены значения расхождения. Логика применения критериев здесь может быть следующая. Принимаем, что данные оценок, полученные на экспертизах, имеющих меньшее значение расхождений, являются более однородными. Это может быть объяснено следующим образом. В идеальном случае мнения экспертов совпали и на всей выборке $i = 1, N$, величина $\Delta_i = 0$. Тогда функции распределения, полученные по рассмотренным экспертизам, однозначно считаются однородными. По мере роста расхождений (при вовлечении в рассмотрение данных, полученных на проектах

О. В. ВИКУЛОВ,
Ю. А. ШАМСУТДИНОВ, Ю. В. КАПРАЛЬНЫЙ
АНАЛИЗ КАЧЕСТВА И ДОСТОВЕРНОСТИ ЭКСПЕРТИЗ
ПРОЕКТОВ, ВЫПОЛНЕННЫХ ПРЕДСТАВИТЕЛЯМИ
ФЕДЕРАЛЬНОГО РЕЕСТРА ЭКСПЕРТОВ (ФРЭ)

с большим расхождением) логично предположить, что однородность данных будет ухудшаться, соответственно, величина статистики оценки однородности должна расти. В свою очередь, величина расхождений является характеристикой достоверности экспертизы и, таким образом, мы имеем прямую зависимость однородности с достоверностью.

Тогда алгоритм разбиения экспертиз на группы может быть выполнен следующим образом.

Вся совокупность данных (две выборки оценок экспертов, табл. 1) ранжируется по величине расхождения и поэтапно вычисляется статистика на выборочных значениях интервала расхождений, которые выбираются начиная от окрестностей в пределах нулевого значения до максимального.

Здесь данные ранжированы согласно значениям Δ_i , где $\Delta_i < \Delta_i + 1$, соответственно, к выборке «m» относятся экспертизы, имеющие расхождение в оценках менее или равные величине Δ_m , к выборке «k» – экспертизы, попадающие в интервал $\Delta_m < \Delta_i < \Delta_k$, и т. д. В общем случае, величина интервала расхождений может быть выбрана одинаковой для всех выборок. На каждой выборке вычисляется статистика по одному из критериев (например, Смирнова) – Q_m, Q_k, \dots, Q_n .

Предполагаемый характер изменения статистики, таким образом, согласно рассмотренной модели принятия решения экспертом и прямой зависимости оценки однородности со степенью достоверности, может тяготеть либо к варианту монотонного роста статистики (вариант 1), либо к варианту скачкообразных изменений значений статистики (вариант 2), происходящих на границах между разными областями достоверности принятых решений (рис. 2). Возможны промежуточные варианты.

То есть, определение границ возможно либо в окрестностях резкого изменения величины статистики, отслеживаемой по ее приращению при переходе от одной выборки к другой (вариант 2), либо, если явно выраженных участков не наблюдается, границы определяются классическим образом, то есть путем сравнения статистик с критическим значением с заданным уровнем значимости, определенным заказчиком экспертизы (вариант 1). Например, для разделения экспертиз (при втором варианте) на «достоверно» и «условно-достоверно» оцененные может быть выбран уровень значимости на уровне 80% доверительного интервала, а для разделения «условно-достоверных» и «недостоверных» 95% – интервал.

Можно допустить, что принятые предположения, в первую очередь, о прямой зависимости роста неоднородности данных с ростом параметра расхождения, могут в практических приложениях не проявиться явно или полученные результаты окажутся противоречивыми. Тогда дополнительным анализом может быть оценка данных на базе параметрических критериев, например, коэффициента корреляции, вычисленном

ТАБЛИЦА 1.
Выборки оценок экспертов

Проек-ты	R ₁	R ₂	Δ _i	Выборка «m»	Выборка «k»	...	Выборка «N»	Значение критерия					
								«m»	«k»	...	«N»		
~	R _{1~}	R _{2~}	Δ ₁	1				Q _m					
~	R _{1~}	R _{2~}	Δ ₂	2									
...									
~	R _{1~}	R _{2~}	Δ _m	m				Q _k					
~	R _{1~}	R _{2~}	Δ _{m+1}		1								
...								
~	R _{1~}	R _{2~}	Δ _k		k			Q _~					
~	R _{1~}	R _{2~}	Δ _{k+1}			1							
...							
~	R _{1~}	R _{2~}	Δ _{...}			~		Q _z					
~	R _{1~}	R _{2~}	Δ _{...+1}				1						
...						
~	R _{1~}	R _{2~}	Δ _N				z						

СТАТИСТИКА

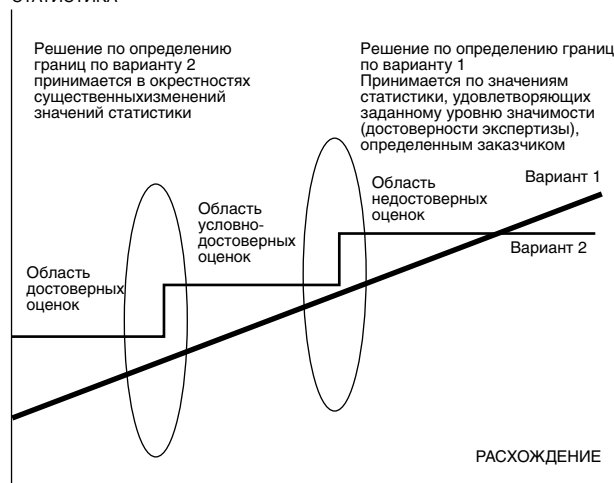


Рис. 2.
Предполагаемые (гипотетические) варианты изменения статистики

на соответствующих выборках, что разрешало бы возможные противоречия и повышало достоверность полученных результатов.

АНАЛИЗ РЕЗУЛЬТАТОВ ЭКСПЕРТИЗЫ

Исходные данные результатов экспертизы имеют вид, представленный в табл. 2, где Δ_i = |R_{1i} - R_{2i}|, величина расхождения в оценках экспертов по модулю.

Был проведен анализ 2 экспертиз:

- экспертизы 168 проектов в области естественных и технических наук (далее - выборка 1);

ТАБЛИЦА 2.

Данные первичной экспертизы

Проекты	Балл эксперта 1 (R ₁)	Балл эксперта 2 (R ₂)	Средний Балл (R)	Расхождение (Δ _i)
1	R ₁₁	R ₂₁	R ₁	Δ ₁
2	R ₁₂	R ₂₂	R ₂	Δ ₂
...
N	R _{1N}	R _{2N}	R _N	Δ _N

- экспертизы 500 проектов в междисциплинарных областях (далее - выборка 2).

Обе экспертизы проводились по одной и той же методике.

В качестве непараметрических критериев использовались следующие: Смирнова, омега-квадрат (Лемана-Розенблатта) и Вилкоксона (Манна-Уитни). Выбор этих методов преследовал цель получить более обоснованную и достоверную оценку выдвинутого в теоретической части метода анализа экспертиз, поскольку два первых из перечисленных статистических методов проверки гипотез относятся к «мощным», третий же к часто используемому, что определило их выбор. Привлечение всех трех, с другой стороны, учитывало тот факт, что каждый метод имеет свои особенности применения и, чтобы их нивелировать, проверка данных выполнялась по всем трем методам.

Для возможности использования дополнительно параметра анализа – коэффициента корреляции (КК), распределение исходных оценок было проверено на соответствие нормальному закону по критерию Колмогорова [1].

Оценка на нормальность распределения данных выполнялась по формуле

$$\sqrt{n}D_n(\theta^*) = \sqrt{n} \sup_x |F_n(x) - F(x; \theta^*)|,$$

где $F_n(x)$ и $F(x; \theta^*)$ – соответственно, коммулятивные эмпирические и теоретические частоты. По полученному значению максимума непосредственно значение критерия вычислялось по формуле

$$\lambda = \frac{D}{\sqrt{n}},$$

который и сравнивался с критическим на предмет проверки гипотезы H_0 , то есть, соответствия распределения оценок экспертов нормальному закону распределения. В отличие от «классического» критерия Колмогорова, где в параметре $\theta = (m, \sigma^2)$ используются истинные значения математического ожидания и дисперсии, был использован параметр $\theta^* = (\bar{x}, s^2)$, где эти величины были заменены на выборочные. То есть, примененный критерий был «типа Колмогорова» и поэтому квантили статистик были выбраны с учетом этого и имели меньшие значения, по сравнению с «классическими». Величины квантилей для разных уровней значимости приведены в табл. 3.

Полученные значения статистики на группах данных сведены в табл. 4.

Как видно, практически на всех интервалах данных, величина $Q_{i,j}$ для обеих выборок имеет значения, меньше квантиля на уровне значимости 0,05 (за исключением одного значения для 2 выборки), поэтому, с небольшим допущением, с данным уровнем значимости мы можем принять гипотезу, согласно которой

ТАБЛИЦА 3.

Величины квантилей для разных уровней значимости

Уровень значимости α	0,15	0,1	0,05	0,025	0,01
Квантили для $\sqrt{n}D_n(\theta^*)$	0,775	0,819	0,895	0,955	1,035

ТАБЛИЦА 4.

Статистика на группах данных

Значение критерия, $Q_{i,j}$	Интервалы, $\Delta_{i,j}$				
	$\Delta_{0,10}$	$\Delta_{10,20}$	$\Delta_{20,30}$	$\Delta_{30,45}$	$\Delta_{45,60}$
Выборка 1	0,71	0,58	0,29	0,81	0,67
Выборка 2	0,68	0,87	0,33	0,74	1,20

О.В. ВИКУЛОВ,
Ю.А. ШАМСУТДИНОВ, Ю.В. КАПРАЛЬНЫЙ
АНАЛИЗ КАЧЕСТВА И ДОСТОВЕРНОСТИ ЭКСПЕРТИЗ
ПРОЕКТОВ, ВЫПОЛНЕННЫХ ПРЕДСТАВИТЕЛЯМИ
ФЕДЕРАЛЬНОГО РЕЕСТРА ЭКСПЕРТОВ (ФРЭ)

наши данные согласованы с нормальным законом распределения и применение параметра - коэффициента корреляции для дополнительного анализа допустимо.

ОБРАБОТКА ДАННЫХ

Значения двухвыборочной статистики Смирнова [4] вычислялись по выражению

$$\lambda' = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \times \max |F_{n_1}(x) - F_{n_2}(x)|,$$

где $F_{n_1}(x)$ и $F_{n_2}(x)$ – эмпирические функции распределения, построенные по двум выборкам с объемами n_1 и n_2 .

Значения двухвыборочной статистики по критерию типа омега-квадрат (Лемана-Розенблатта) [5] вычислялись по формуле

$$A = \frac{n_1 n_2}{n_1 + n_2} \int_{-\infty}^{\infty} (F_{n_1}(x) - F_{n_2}(x))^2 dH_{n_1+n_2}(x),$$

где $H_{n_1+n_2}(x)$ – эмпирическая функция распределения, построенная по объединенной выборке:

$$H_{n_1+n_2}(x) = \frac{n_1}{n_1 + n_2} F_{n_1}(x) + \frac{n_2}{n_1 + n_2} F_{n_2}(x).$$

Статистика Z двухвыборочного критерия Манна-Уитни [7] определялась следующим образом

$$Z = \frac{U - EU}{\sqrt{DU}} = \frac{R_2 - \frac{n_2(n_2+1)}{2} - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2}{12} (n_1 + n_2 + 1)}},$$

где U – число инверсий, EU – среднее число инверсий, DU – дисперсия инверсий.

Полученные значения статистик на указанных интервалах сведены в табл. 5.

В таблице λ' , A_1 и Z_1 – статистики Смирнова, Лемана – Розенблатта и Манна-Уитни для 1-й выборки, с индексом два – для 2-й.

Объединенный график для выборки 1 представлен на рис. 3, для выборки 2 – на рис. 4.

Предварительный анализ данных по графикам показывает несостоятельность критерия Манна-Уитни – сильный разброс значений и нулевое значение на выборке 1 на интервале 45–60 (резкое падение на выборке 2), что существенно расходится с величинами статистик, полученных по двум другим критериям. Следует отметить, что и в работе [4] подчеркивалось, что именно два критерия – Смирнова и Лемана-Розенблатта являются состоятельными в отличии от множества других, включая критерий Манна-Уитни. Исходя из этого, дальнейший анализ проводился с использованием результатов применения критериев Смирнова

ТАБЛИЦА 5.

Статистика на указанных интервалах

Критерий	$\Delta_{i...j}$				
	$\Delta_{0...10}$	$\Delta_{10...20}$	$\Delta_{20...30}$	$\Delta_{30...45}$	$\Delta_{45...60}$
λ'_1	0,101	0,106	0,353	0,308	0,600
λ'_2	0,039	0,1	0,065	0,141	0,260
A_1	0,063	0,049	0,706	0,090	1,130
A_2	0,052	0,151	0,069	0,231	0,215
Z_1	0,130	0,087	3,010	0,564	0,000
Z_2	0,252	0,562	0,011	0,860	0,197

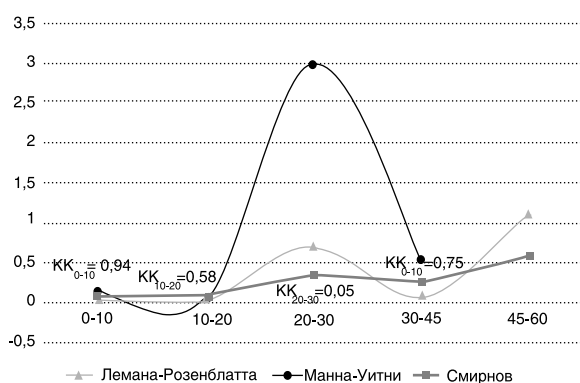


Рис. 3.

Общий график статистик Лемана-Розенблатта, Манна-Уитни и Смирнова на выборке 1

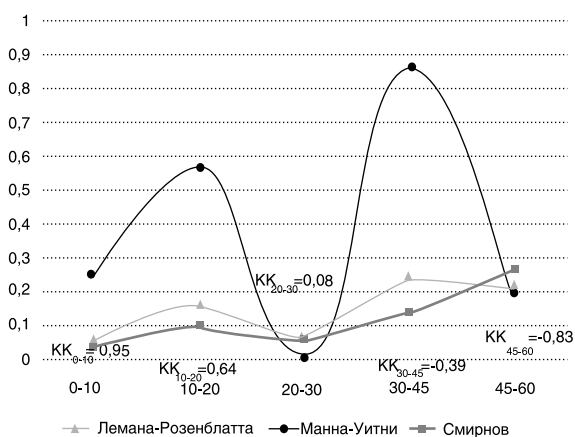


Рис. 4.

Общий график статистик Лемана-Розенблатта, Манна-Уитни и Смирнова на выборке 2

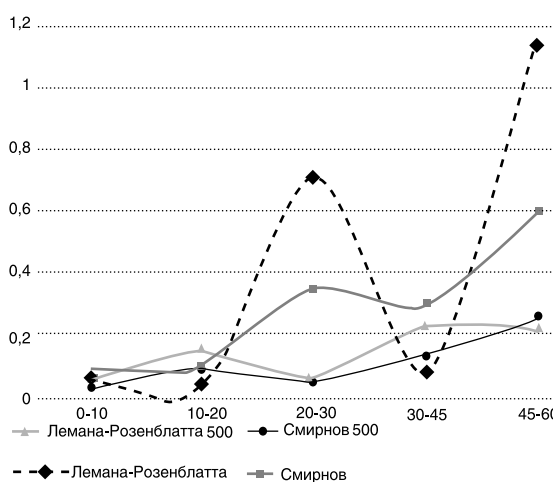


Рис. 5.

Общий график статистик Смирнова и Лемана-Розенблатта для двух выборок

и Лемана-Розенблатта, а анализ по критерию Манна-Уитни не проводился.

Обобщенный график для обеих выборок, представленный на рис. 5, отражает данные только по двум критериям – Смирнова и Лемана-Розенблатта.

АНАЛИЗ ДАННЫХ

Из анализа графиков на рис. 3, 4 и 5 мы видим, что на обеих выборках проявляются три характерных участка, на границе которых происходят практически синхронные и существенные изменения статистик по обоим критериям.

Для выборки 1 это интервалы расхождений: (0–20), (20–30–45), (более 30–45).

Для выборки 2: (0–10), (10–30–45), (более 30–45).

Не будем пока акцентировать внимание на факте смещенности данных у обеих выборок (график на рис. 5 это хорошо выявляет) и то, что статистики для второй выборки оказались меньшего значения. Здесь только заметим, что эти расхождения не противоречат

предложенному методу и имеют логичное объяснение, которое будет представлено ниже.

Визуальный анализ графиков показывает три явно выраженных участка, неоднородность данных в которых, оцененных по величине рассогласования, существенно отличается. Можно сказать, что через эти данные рассмотренная модель эксперта проявляется прямым образом.

Так, при анализе данных выборки 1 мы наблюдаем, что в проектах, где расхождение в мнении экспертов не превысило 20 баллов, значение статистики (по обоим критериям) минимально, что свидетельствует об однородности данных, или, как следует из предыдущих рассуждений, о достоверности экспертиз, то есть, эти экспертизы могут быть отнесены к «достоверно оцененным». Соотнося эту оценку с моделью эксперта, в этом случае условный эксперт принимал решение в «области стереотипных для него ситуаций, в которых оценка считается наиболее достоверной». Для второй выборки этот интервал расхождений составил 0–10 баллов.

В экспертизах с расхождением в пределах 20–40 (второе значение взято примерно средним по рассмотренному интервалу 30–45) видно нарушение однородности данных, т. е. значение статистики возрастает, при этом с явной границей справа, где ее значения несколько снижаются. Соотнося эту оценку с моделью эксперта видно, что здесь у условного эксперта «стереотипные эвристики уже неадекватны ситуации – он осуществлял творческое решение задачи». Для второй выборки это интервал 10–30.

Наконец в экспертизах, имеющих наибольшие значения расхождений (более 40) видна явная неоднородность в распределении оценок, поскольку величины статистик растут и имеют наибольшие абсолютные значения и по модели эксперта это, очевидно, область «обыденного мышления», когда эксперт уже перестает быть экспертом. Для второй выборки это интервал – от 30 баллов и выше.

Дополнительным фактором, свидетельствующим о непротиворечивости предложенному подходу полученных результатов, служат величины коэффициентов корреляции, также оцененных на данных интервалах. Как видно (рис. 3, 4), на проектах, оцененных достоверно, $КК \sim 0,95$, что свидетельствует о некоторой зависимости оценок экспертов, что может быть объяснено тем, что информация о проекте (не важно какого качества была его содержательная часть) была исчерпывающей, знания экспертов соответствовали темам проектов и субъективные факторы сказались минимально, так что мнения практически совпали, то есть мнения оказались зависимы через очевидность ожидаемого результата экспертизы. На проектах, отнесенных к оцененным «недостоверно», $КК \sim -0,80$, что так же свидетельствует об зависимости мнений экспертов, только отрицательной. Эта зависимость

О. В. ВИКУЛОВ,
Ю. А. ШАМСУТДИНОВ, Ю. В. КАПРАЛЬНЫЙ
АНАЛИЗ КАЧЕСТВА И ДОСТОВЕРНОСТИ ЭКСПЕРТИЗ
ПРОЕКТОВ, ВЫПОЛНЕННЫХ ПРЕДСТАВИТЕЛЯМИ
ФЕДЕРАЛЬНОГО РЕЕСТРА ЭКСПЕРТОВ (ФРЭ)

имеет противоположную (в отличие от рассмотренной на достоверно оцененных проектах) природу. Если попытаться интерпретировать этот результат с позиции модели эксперта, оценившего проект на основе «обыденного мышления», то оно сформировано на основе того, что на мнение одного из экспертов (или, возможно, обоих) как раз повлияли факторы, способствующие проявлению неадекватных эвристик (недостаток времени или познавательных ресурсов для исчерпывающего анализа ситуации, неопределенность, информационная сложность, специфика действий стереотипов, запускаемых автоматически без достаточных оснований, принцип обязательства или занятой позиции, принцип авторитета, научные парадигмы и нормы научных направлений и школ). То есть, эта зависимость имеет в своей основе «неадекватность» или противоречивость оценки, что и проявилось в отрицательном значении $КК$. В области «условно достоверных экспертиз» $КК$ имеет промежуточное значение.

Поясним логичность обнаруженных отличий в выявленных интервалах и поведении статистик на различных выборках. Как указывалось, ранее, выборка 1 была сформирована на 168 оценках экспертиз проектов по естественно научным и техническим темам, 2 – на 500 оценках экспертиз проектов по междисциплинарным темам.

1. По поводу меньших значений величин статистик на выборке из 500 проектов (выборка 2). Каждый балл, выставленный экспертом может рассматриваться как некая сумма составляющих. Например, итоговый бал суммировался из баллов, выставленным по 7–10 критериям (статус научного руководителя, научный задел, уровень коллектива, пр.), в каждом критерии учитывались дополнительные параметры оценки (количество патентов, статей, индекс цитирования, пр.). Если допустить вклад составляющих баллов в итоговую оценку как равноценный, то согласно центральной предельной теореме при росте числа наблюдений распределение частот конечной наблюдаемой величины стремится к нормальному распределению. То есть, распределения оценок у обеих выборок (первого и второго эксперта) имея свои, отличные вероятности распределений, при увеличении данных стремятся к нормальному закону, тем самым, появляется дополнительный фактор, который повышает их однородность в общем смысле – обе выборки «нормализуются». В этом смысле уменьшение величин статистик, отражающих величину однородности данных, на всем интервале расхождений выглядит закономерным.

2. Смещенность данных. Проекты второй выборки (по междисциплинарным темам) имели гораздо больший разброс в мнениях экспертов, поскольку для экспертиз привлекались эксперты из смежных областей, и у них естественным образом было больше экспертиз, попавших в интервалы с большим расхождением. Это

прямым образом отразилось в смещении графиков, так, что области «условно достоверных» и «недостаточно оцененных» имели оценку неоднородности уже на меньших интервалах расхождений.

Таким образом объясняются как меньшие значения статистик на большей выборке, так и смещенность оценки однородности (достоверности), связанная с субъективными факторами (в данном случае, с другим характером проектов, представленных на экспертизу). В целом же, метод показал согласованные результаты в обоих случаях.

В итоге можно констатировать, что предложенный метод показал полное согласие с представлением результата экспертиз в виде явно выраженных трех групп, каждая из которых характеризуется (в смысле «достоверности») определенным значением статистики, которое в определенных окрестностях анализируемого параметра (расхождения) относительно существенно меняется, определяя границы разделения групп. Исходя из этого, к каждой группе были применены следующие действия:

– проекты со статусом «достоверно оцененные» были приняты как прошедшие экспертизу безусловно и не требующие дополнительных проверок;

– проекты со статусом «недостаточно оцененные» были рекомендованы к проведению дополнительной экспертизы, что и было сделано;

– проекты со статусом «условно достоверно оцененные» были отправлены экспертам на доработку и уточнения.

ЗАКЛЮЧЕНИЕ

Таким образом, анализ экспертизы проектов, представленных на экспертизу, был выполнен в несколько этапов.

На первом этапе проведения основной (первичной) экспертизы по методу оценочного анкетирования, по разработанной методике были определены границы, согласно которым проекты были разделены на 3 группы, к каждой из которых была применена своя процедура обработки: оценки по «условно достоверно оцененным» проектам были просто уточнены, путем уточнения экспертами своих решений, по «недостаточно оцененным проектам» была проведена дополнительная экспертиза по методу «комиссий».

На заключительном этапе данные по всем экспертизам были объединены и представлены вместе с проектами на конкурс.

ЛИТЕРАТУРА

1. АБРАМОВА Н. А., ВАССУНОВ И. В. О влиянии формализации на адекватность когнитивной модели управленческой ситуации // Труды 5-й международной конференции «Когнитивный анализ и управление развитием ситуаций» CASC' 2005. М.: ИПУ РАН, 2005. С. 47–51.
2. АБРАМОВА Н. А. О проблеме рисков из-за человеческого фактора в экспертных методах и информационных технологиях // Проблемы управления. 2007. № 2. URL: <http://cyberleninka.ru/article/n/o-probleme-riskov-iz-za-chelovecheskogo-faktora-v-ekspertnyh-metodah-i-informatsionnyh-tehnologiyah> (дата обращения: 12.08.2014).
3. АЛЕКСЕЕВА Н. П. Введение в статистические методы. Учебное пособие по прикладной статистике. СПбГУ, 2010.
4. ОРАОВ А. И. Математика случая: Вероятность и статистика – основные факты. Уч. пособие. М.: МЗ-Пресс, 2004.
5. ОРАОВ А. И. Состоятельные критерии проверки абсолютной однородности независимых выборок // Заводская лаборатория. Диагностика материалов. 2012. Т. 78. № 11. С. 66–70.
6. ОРАОВ А. И. Непараметрические критерии согласия Колмогорова, Смирнова, омега-квадрат и ошибки при их применении // Научный журнал КубГАУ. 2014. № 97 (03).
7. ЧЕЛДАНИ Р. Психология влияния. СПб.: Питер, 2001. 270 с.

Виколов Олег Владимирович,

д.т.н., профессор, зам. директора Государственного центра экспертизы в сфере науки и инноваций НИИ РИНКЦЭ,

☎ 123995, г. Москва, ул. Антонова-Овсеенко, д.13, стр.1,
тел.: +7 (499) 259-52-64, e-mail: vikulov@extech.ru,

Шамсутдинов Юрий Аслахович,

к.т.н., начальник отдела РИНКЦЭ,

☎ 123995, г. Москва, ул. Антонова-Овсеенко, д.13, стр.1,
тел.: +7 (499) 259-52-64, e-mail: jurisp@extech.ru,

Капральный Юрий Викторович,

вице-президент Национального фонда социально-экономического развития регионов,

☎ 105187, г. Москва, ул. Мироновская, д. 38,
тел.: +7 (495) 918-34-30, e-mail: kapralniy@mail.ru