

## ОЦЕНИВАНИЕ ДОСТОВЕРНОСТИ ЭКСПЕРТНЫХ ОЦЕНОК

Евсеев Т.И., Телебук О.И., Гиль Н.Н.

*Белорусский национальный технический университет**Минск, Республика Беларусь*

При оценивании различных объектов эксперты часто расходятся во мнениях относительно решаемой проблемы. В результате этого возникает неопределенность результатов экспертного оценивания и, как следствие, снижение степени их достоверности.

Задачу экспертного оценивания можно рассматривать как задачу измерения, поэтому качество экспертного оценивания должно характеризоваться понятием точности. Однако экспертные оценки часто используются при решении слабоструктурированных задач, в которых истинное значение объекта практически недостижимо или относительно и зависит от сложившихся обстоятельств в конкретный момент времени. У таких величин, выражаясь языком математики, нет математического ожидания. В связи с этим получить оценку смещения экспертных оценок от истинного значения оцениваемого параметра объекта не всегда возможно и целесообразно. Другим параметром, характеризующим точность, может служить разброс или рассеяние экспертных оценок. Этот параметр в большей степени отражает согласованность мнений экспертов.

Возникает задача количественной оценки достоверности экспертных оценок как наиболее презентативный способ анализа точности. Логично установить количественные показатели достоверности экспертных оценок, которые будут показывать уровень доверия исследователей к полученным результатам.

При экспертном оценивании зачастую непостижимым является истинное значение оцениваемого параметра по ряду причин. Экспертные оценки в большинстве случаев используются там, где использование технических средств невозможно или затруднено в силу специфичности объектов измерения, причем такие объекты зачастую в принципе не имеют истинного значения, а лишь некую объективную оценку или заданное целевое значение. Таким образом, целесообразно говорить не о существовании истинного значения величины, а о существовании некоторого интервала значений, который может быть обоснованно приписан данной величине. Поэтому мы видим необходимость представления оцениваемых параметров экспертным методом с позиции случайных величин.

Нами предлагается новый подход к определению достоверности экспертных оценок с использованием понятий и положений теории не-

определенности, широко используемой на практике в области метрологии. Аргументация схожести и сущности задач метрологии и теории экспертного оценивания дает нам основание для обоснованного использования инструментов, используемых в метрологии для определения точности результатов измерения.

Экспертное оценивание ставит задачу получения экспертных оценок по рассматриваемому объекту (результат измерения) с помощью эксперта и принятой шкалы оценивания (средство измерения) и дальнейшей оценки достоверности полученных результатов (точность измерений).

Серия стандартов СТБ ИСО 5725 регламентирует подход к определению точности методов измерения с помощью понятий правильности и прецизионности. Правильность и прецизионность результатов измерения обусловлена соответственно случайными и систематическими эффектами. Случайные эффекты возникают из-за непредсказуемых и стохастических временных и пространственных изменений влияющих величин, вызывают изменения измеряемой величины при повторных наблюдениях, а математическое ожидание или ожидаемое значение таких изменений равняется нулю. Систематические эффекты остаются постоянными или закономерно изменяющимися при повторных измерениях одной и той же физической величины. Такие же эффекты присущи и результатам экспертного оценивания. Эксперт, склонный к систематическим ошибкам, выдает значения, которые устойчиво отличаются от истинного в сторону увеличения или уменьшения. Для коррекции систематических ошибок применяют поправочные коэффициенты или специально разработанные обучающие процедуры, которые позволяют снизить величину систематических ошибок за счет приобретения экспертом навыков оценивания. Случайные ошибки характеризуются величиной дисперсии.

Нами принята модель по выражению неопределенности результатов измерения, установленная техническим отчетом EURO-LAB №1/2007, для оценивания достоверности экспертных оценок на основании результатов комплексирования систематической и случайной составляющих экспертного оценивания. Таким образом, показателем достоверности экспертных оценок является их неопределенность

$$U = \sqrt{\sigma^2 + \Delta^2}. \quad (1)$$

За оценку прецизионности принимается среднеквадратичное отклонение  $\sigma$ , которое характеризует величину систематической составляющей. За оценку правильности принимается смещение оценки от истинного значения  $\Delta$ , которое характеризует величину случайной составляющей.

Для решения проблемы по оцениванию неопределенности экспертных оценок нами сформулирована следующая задача: определить соотношение  $\Delta$  и  $\sigma$

$$\Delta = (c_1 \dots c_2)\sigma \quad (2)$$

Зная величину соотношения (2), задача по определению неопределенности сводится к поиску одной неизвестной вместо двух. Численно определить соотношение (2) возможно лишь эмпирическим путем, поэтому нами был спланирован и реализован эксперимент экспертного оценивания.

Модель экспертного оценивания кратко изложена в таблице 1.

Таблица 1 – Основные положения модели экспертного оценивания

| Структурный элемент модели | Содержание структурного элемента   |
|----------------------------|--|
| Объект оценивания          | - окружность<br>- неподобный прямоугольник<br>- параллелепипед   |
| Параметр оценивания        | - площадь<br>- объем   |
| Шкала оценивания           | любая цифровая шкала от 0 до $+\infty$ на усмотрение респондента   |
| Метод оценивания           | модифицированный метод альтернатив   |
| Методика оценивания        | - каждый вид фигур предьявляется отдельно<br>- фигуры предьявляются в случайном порядке<br>- оценивание по каждой фигуре в отдельности без ознакомления с размерностями фигур в серии<br>- 4 серии по каждому виду фигур:<br>а) 7 фигур в серии<br>б) 9 фигур в серии<br>в) 11 фигур в серии<br>г) 15 фигур в серии<br>- количество опрашиваемых респондентов по каждой серии фигур - 10 |

На основании полученных результатов эксперимента экспертного оценивания мы сделали следующие заключения:

- 11-балльная экспертная шкала для оценивания качества объектов, характеризуемых одним или двумя свойствами, и 9-балльная экспертная шкала для оценивания качества объектов, характеризуемых тремя свойствами, наиболее приемлемы;

- эксперты склонны завышать свои оценки при увеличении количества предьявляемых фигур на оценивание;

- на величину смещения экспертных оценок оказывает влияние соотношение «количество оцениваемых объектов – размерность шкалы»;

- экспертная субъективная оценка отличается от объективной оценки или номинальной величины оцениваемого свойства в  $(1+k)$  раз в случае линейной зависимости  $\Delta$  от номинального значения;

- смещения экспертных оценок от номинальных величин оцениваемых параметров объектов, характеризуемых одним или двумя свойствами, получаемых (экспертных оценок) по 11-балльной шкале, имеют линейный характер увеличения при увеличении интенсивности оцениваемых свойств объекта, что характеризует 11-балльную экспертную шкалу как наиболее удобную с позиций респондента и модератора;

- величина отклонения экспертных оценок от действительных значений оцениваемого параметра уменьшается по мере увеличения числа оцениваемых фигур в серии и при этом не зависит от вида фигуры;

- величина СКО экспертных оценок уменьшается по мере увеличения числа оцениваемых фигур в серии и при этом не зависит от вида фигуры;

- при оценивании качества объекта по одному свойству следует отдавать предпочтение размерам шкал от 11 до 15 оценочных единиц;

- при оценивании качества объекта по двум свойствам следует отдавать предпочтение размерам шкал от 9 до 15 оценочных единиц;

- при оценивании качества объекта по трём свойствам следует отдавать предпочтение размеру шкалы в 15 оценочных единиц, при этом случайная составляющая практически равна систематической;

- 7-балльная шкала при оценивании качества объекта по одному или двум свойствам менее предпочтительна, т.к. более зависима от изменения величины оцениваемого параметра при большом превосходстве случайной составляющей над систематической.