УДК 519.23

МЕТОДИКА СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ ОЦЕНОК ПРИ ИССЛЕДОВАНИИ МАССИВА УЧЕБНЫХ ПОСОБИЙ

© Н. А. Буренин, В. Н. Петрушин, Е. В. Никульчев, д.т.н., доцент, МГУП, Москва, Российская Федерация

У роботі запропоновано методику статистичного аналізу експертних оцінок при дослідженні масиву навчальних посібників для установ загальної освіти. Методика припускає отримання оцінки якості видань, задовільної за заданою величиною довірчого інтервалу і надійності.

The author proposes the methodology of statistic analysis of the expert assessments during the investigation of the textbooks for the secondary education institutions. According to the methodology it is assumed that it is possible to obtain the results of quality assessment of the editions to be satisfactory by the preset value of the confidence interval and reliability.

Оценка эффективности использования учебных изданий для учреждений общего образования во многом определяется их качеством и доверием к ним экспертов, таких как учащиеся, учителя, методисты.

Отметим, что издательское дело является слабопараметризуемой областью, обладающей неформализуемыми особенностями применения, дидактическими и психологическими факторами.

Наиболее приемлемым решением при проведении экспертизы учебных изданий является использование статистических методов анализа экспертных оценок. Это позволяет использовать особенности человеческого восприятия недетерминированного в жестких рамках правил, а также математический аппарат для оценки достоверности и надежности полученных результатов.

Интегральное экспертное оценивание представляет собой процесс измерения, который можно определить как процедуру субъективного сравнения объектов по выбранным показателям и формирования групповой интервальной оценки.

Анализ предметной области выявил следующие требования к формированию параметров оценок для использования специалистами в области книгоиздания:

— эксперт должен иметь возможность дать различимые оценки в диапазоне от 0 до 10;

— выборка должна быть случайной;

— эксперты, дающие оценку, не должны контактировать друг с другом;

— необходимо обеспечить объем пробной выборки, достаточный для достоверности статистических оценок.

Целесообразно выделять следующие стадии экспертного опроса:

1) формулировка лицами, принимающими решения, цели экспертного опроса;

2) разработка подробного сценария проведения сбора и анализа экспертных оценок, включая конкретный вид экспертной информации;

3) проведение сбора экспертной информации;

4) анализ экспертной информации;

5) интерпретация полученных результатов и подготовка заключения.

Для анализа экспертных оценок при исследовании массива учебных пособий в части учебных изданий для учреждений общего образования разработана методика, которая позволит иметь оценку качества изданий, удовлетворительную по заданной величине доверительного интервала и надежности.

Разработанная методика состоит из пяти шагов.

1. Определение размерности пространства оценок. Предполагается рассмотрение оценок групп экспертов в качестве независимых координат метрического пространства. Группы формируются по социальным признакам и количественно оценивается кластерным анализом.

Кластерный анализ является анализом качественным, группировка в котором осуществляется по одинаковым или сходным признакам. В данном случае сформированы три разнородных, независимых групп экспертов. Поэтому, размерность пространства оценок с вероятностью, близкой к единице, равна 3:

$$\lim_{p\to 1} N(a_n) = 3,$$

где a_n — величина оценок; N — размерность пространства; p — вероятность. Вероятность трехмерности пространства равна единице в случае отсутствия ошибок в формировании группы экспертов.

2. Оценка распределения рассматриваемых случайных величин с применением статистических критериев Пирсона, Колмогорова-Смирнова и др., для выбора наиболее надежной оценочной функции в рассматриваемой задаче.

Учитывая конечность диапазона оценок, их распределение может быть аппроксимировано распределением Бернулли или β-распределением.

Как известно, распределение Бернулли имеет вид

$$P_n(k) = C_n^k\, p^k (1-p)^{n-k},$$

где n — число оценок; k — число выделенных оценок; p — вероятность выделенных оценок; C_n^k — число сочетаний оценок $C_n^k = n!/(k!(n-k)!)$.

Плотность вероятностей бета-распределения имеет вид

$$f(x) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \text{для всех } x \in [0,1], \\ 0, \text{для всех } x \notin [0,1], \end{cases}$$

где x — нормированное значение оценки для нормированной длины интервала; нормировка осуществляется длиной интервала в границах экспертных оценок [–δ, 10 + δ], где δ — бесконечно малая величина; гамма-функция определяется в соответствии с интегралом Эйлера 2-го рода

$$\Gamma(z)=\int_0^{\infty} x^{z-1}e^{-x}dx.$$

Оценка из предложенных распределений случайных величин осуществляется с применением статистических критериев Колмогорова-Смирнова, Пирсона [см. напр. 1]. Преимуществом пользуется та, надежность которой ближе к единице.

Критерий согласия Колмогорова применяется для проверки правильности подбора теоретического распределения. Для его применения нужно найти максимальную по модулю разность между выборочной (опытной) функцией распределения $\tilde{F}(x)$ и теоретической F(x):

$$D=\max_{\forall x}|F(x)-\tilde{F}(x)|,$$

по которой вычисляется $\lambda = D\sqrt{y}$, которая сравнивается с квантилем λ-распределения Колмогорова (квантили приведены в статистических таблицах в специальной литературе). Здесь y — объем выборки. Если величина λ не превосходит квантиля λ_p, то с доверительной вероятностью p статистическую гипотезу можно принять. Если же $\lambda > \lambda_p$, то теоретическое распределение подобрано неверно.

Для применения критерия согласия Колмогорова нужно построить на одном графике теоретическую функцию распределения $F_x(x)$ и выборочную $\tilde{F}(x)$, а затем применить сам критерий. График эмпирической функции распределения $\tilde{F}(x)$ представляет ступенчатую линию: это ломаная со ступеньками высотой 1/y в точках с абсциссами y_i.

В критерии согласия Пирсона сравниваются между собой теоретические и эмпирические числа попаданий в интервалы. Эмпирические числа попаданий в эти интервалы y_j сравниваются с теоретическим числом попаданий yp_j, где p_j — вероятность попадания нашей величины в j-й интервал. Теоретическое распределение можно считать подобранным верно с доверительной вероятностью p, если суммарная квадратичная относительная разность между теоретическим и практическим числом попаданий в каждый интервал будет не очень большой. При этом должно выполняться условие

$$\sum_{i=1}^{k}\frac{(y_j-yp_j)}{yp_j}\le\chi_p^2(k-m),$$

причем нужно, чтобы все $yp_j \geq 5$. Здесь m — число ограничений, равное числу параметров выбранного закона распределения плюс 1. Следовательно, необходимо построить таблицу результатов, в которую занесены номера интервалов, границы интервалов a_j и b_j, вероятность попадания в интервал p_j, теоретическое число попаданий yp_j и практическое число попаданий

y_j. Теоретическая вероятность попадания в j-й интервал подсчитывается по формуле

$$p_j = F(b_j) - F(a_j),$$

где $F_x(x)$ — подобранная теоретическая функция распределения.

Далее рассчитывается сумма элементов предпоследней формулы (статистика Пирсона) и сравнивается с квантилем χ^2-распределения Пирсона при заданной доверительной вероятности.

3. Оценка достоверности среднего значения высказываний экспертов по каждому вопросу во всех кластерах на основе выбранных распределений средних значений. В случае недостоверности полученных средних предполагается разбиение первоначально сформированных кластеров на более мелкие, что приведет к изменению размерности пространства оценок.

4. Формирование интегральных оценок в многомерном и одномерном вариантах на основании выбранной надежности результатов и принятых распределений.

В связи с независимостью кластеров объединяющих экспертов, нет необходимости учета коррелированности. Таким образом, интегральная оценка распределения представляет собой произведение трех независимых функций распределения

$$p = \prod_{i=1}^{3} p_i .$$

5. Выделение главных компонентов интегральных оценок и разработка вероятностных рекомендаций для лиц, принимающих решения осуществляется на основе анализа составляющих дисперсии.

Методика реализована в виде программного средства, работающего в среде разработки инженерных и научных приложений MATLAB.

Разработанная методика применена при выполнении государственного контракта Ф-384 (Договор № П 840.2 от 20.11.07 с Федеральным институтом развития образования).

1. Иглин С. П. Теория вероятностей и математическая статистика на базе MATLAB / Иглин С. П. — Харьков : ХПИ, 2006.

Надійшла до редакції 29.06.10