ИНТЕРВАЛЬНАЯ МАТЕМАТИКА И РАСПРОСТРАНЕНИЕ ОГРАНИЧЕНИЙ
МКВМ-2004 РАБОЧИЕ СОВЕЩАНИЯ С. 147–158

Задачи анализа с неопределенными данными – интервальность и/или случайность?

А.П. Вощинин*

В подавляющем большинстве прикладных задач исследователь имеет дело с неточными исходными данными (табл. 1), неопределенность которых порождается различными факторами.

Таблица 1

№	Данные	Источник неопределенности и неточности
1	Результаты измерений	Вариабельность, шумы, ошибки измерения (систематические и случайные), ошибки округления
2	Прогнозные данные	Незнание, неопределенность, неполнота информации, методические ошибки, ошибки округления и дискретизации
3	Экспертные оценки	Субъективность, незнание, ошибки округления

В зависимости от источника неточности и неопределенности данных в настоящее время используются различные модели описания неопределенных данных, включая вероятностную, нечеткую и интервальную модели. Каждая из этих моделей имеет свою парадигму, опирается на соответствующий теоретический аппарат, имеет свои методы анализа и область применения.

Вероятностно-статистическая модель [1–3] основана на предположении, что изучаемая переменная x является случайной величиной с заданной функцией плотности вероятности $f(x)$. Эту модель можно отнести к классическим моделям из-за ее глубокой теоретической и методической проработанности для многих важных приложений, включая регрессионный, корреляционный, факторный и дисперсионный анализ. Вероятностно-статистическая модель до сих пор является наиболее популярной и широко используется в многочисленных приложениях. Несомненным преимуществом, облегчающим ее применение, является наличие многочисленных статистических пакетов, среди которых Статистика, Statgraph, GPSS и др.

"Нечеткая модель" [4, 5] основана на понятии нечеткого множества S, которое описывается парой – нечеткой переменной x и ее функцией принадлежности $\mu_{\mathrm{S}}(\mathrm{x})$. Формально нечеткое множество записывается в виде

$$S = \{(x, \mu_S(x)) \colon x \in X,\ \ 0 \le \mu_S(x) \le 1\}, \tag{1}$$

где x означает возможные значения нечеткой переменной в заданной области X; $0 \le \mu_S(x) \le 1$ – функция принадлежности, задающая степень принадлежности конкретного значения x нечеткому множеству S. Обычно функ-

*ФГУП "ЦНИИАТОМИНФОРМ", Москва.

ция принадлежности $\mu_S(x)$ задается экспертным путем на основе информации об источниках неопределенности переменной x. Понятие "степень принадлежности" в некотором смысле аналогично понятию вероятности, в частности, $\mu_S(x_1) = 0$ означает, что значение x_1 наверняка не принадлежит множеству S, а $\mu_S(x_2) = 1$ – что значение x_2 достоверно принадлежит множеству S. Однако функция принадлежности в отличие от плотности вероятности $f(x)$ не удовлетворяет условию нормировки.

В *интервальной модели* [6–9] неопределенность параметра x описывается границами его возможных значений в виде $[x] = [x_{\min}; x_{\max}]$. В отличие от теории вероятности внутри интервала $[x]$ не задается никакой вероятностной меры, т. е. все значения внутри интервала предполагаются равновозможными (не путать с равновероятными).

Ниже в докладе рассматривается ряд прикладных задач, в которых применяется как вероятностно-статистическая, так и интервальная модели и анализируются ситуации, в которых интервальная модель представляется более осмысленной и корректной.

Неопределенность единичного измерения

Пусть в результате измерения получена точечная оценка $\hat{x}$ неизвестной измеряемой величины x.

Вероятностная модель измерения последовательно реализуется в экспериментальной физике. При этом ошибка измерения рассматривается как случайная, нормально распределенная величина с нулевым средним и известной дисперсией σ_x^2. Неопределенность переменной x для заданной доверительной вероятности (обычно ее принимают равной 0.95) описывается в форме доверительного интервала

$$\hat{x} - 2\sigma_x \leq x \leq \hat{x} + 2\sigma_x. \tag{2}$$

В метрологии в отличие от приведенного выше подхода принята интервальная модель неопределенности. Предполагается, что измерение $\hat{x}$ получено с помощью неточного прибора с известной абсолютной ошибкой измерения Δ, которая включает как систематическую, так и случайную погрешности. (Случай известной относительной ошибки сводится к рассматриваемому). Предполагается, что для любого $\hat{x}$ выполняется условие $|x - \hat{x}| \leq \Delta$, которое естественным образом задает интервал неопределенности в виде.

$$[x] = [\hat{x} - \Delta; \hat{x} + \Delta] = [x_{\min}; x_{\max}]. \tag{3}$$

Если есть необходимость учесть другие факторы неточности, например ошибки округления, исследователь может расширить интервал неопределенности (3).

Сравнивая две приведенных модели, можно заметить, что модель (2) не позволяет учесть факторы неопределенности, не связанные со случайной вариабельностью, включая систематические ошибки измерения, ошибки округления. Кроме того, постулируемое в вероятностной модели нормальное распределение, которое задает неограниченный диапазон величины x, на практике часто оказывается неадекватным, например, для заведомо положительных переменных. Интервальная модель позволяет учесть любые факторы неопределенности.

Выборка повторных наблюдений

Допустим теперь, что имеется выборка $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n$ повторных наблюдений величины x. При анализе выборочных данных необходимо решить две задачи:

- проверить однородность выборки, т. е. отсутствие в ней выбросов;
- найти наиболее точную выборочную оценку математического ожидания x_0 величины x и оценить ее точность.

В рамках вероятностно-статистической модели в предположении, что выборка взята из нормального распределения с дисперсией σ_x^2 точечная оценка величины x и ее среднеквадратичное отклонение находится по формуле

$$\bar{x} = \sum \hat{x}_i/n, \qquad \sigma_{\bar{x}} = \sigma_x/\sqrt{n}. \tag{4}$$

Считается, что оценка (4) является несмещенной и состоятельной, т. е. $\sigma_x/\sqrt{n} \to 0$, $\bar{x} \to x_0$ при $n \to \infty$. Нетрудно заметить, что в рамках модели с постулируемым нормальным распределением само понятие “выброса” становится не совсем корректным.

В рамках интервальной модели предполагается заданным интервал неопределенности каждого выборочного наблюдения

$$[x]_i = [x_i - \Delta x;\, x_i + \Delta x],$$

где Δx – известная ошибка измерения или выборочного наблюдения. Совокупность интервальных измерений образует “интервальную выборку”

$$\{[x]_1 \ldots [x]_i \ldots [x]_n\}.$$

При этом неопределенность выборочных значений определяется не только случайной изменчивостью, но и наличием систематических составляющих погрешности. Представим ошибку Δx в виде суммы

$$\Delta x = \Delta x_1 + \Delta x_2, \tag{5}$$

где значение Δx_1 определяет переменную, а Δx_2 – постоянную составляющую неопределенности.

Если выборка однородна, то каждое интервальное наблюдение содержит неизвестное истинное значение x$_0$, т. е. $x_0 \in [x]_i$, $\forall\ i = 1, \ldots, n$.

Если существует интервальной наблюдение $[x_j]$ такое, что оно не пересекается с другими интервальными наблюдениями, его естественно рассматривать как выброс и исключить из выборки для обеспечения ее однородности.

Введем понятие интервального выборочного среднего $[\bar{x}]$, которое определяется как пересечение выборочных интервальных измерений

$$[\bar{x}] = [\max_i(x_i^-);\ \min_i(x_i^+)] = [\bar{x}]. \tag{6}$$

Можно показать, что интервальное среднее $[\bar{x}]$, найденное по однородной выборке, обладает следующими свойствами:

- Средняя точка интервала $[\bar{x}]$ совпадает со статистической выборочной оценкой медианы. Ширина интервала $[\bar{x}]$ определяет точность оценки с учетом случайных и систематических составляющих.
- При отсутствии систематической составляющей ($\Delta x_2 \equiv 0$) и увеличении числа выборочных значений $[\bar{x}] \to x_0$ при $n \to \infty$. В общем случае среднее арифметическое (4) не принадлежит интервальному среднему $[\bar{x}]$. В ходе численного моделирования установлено, что при равномерном распределении выборочных измерений ширина интервального среднего $[\bar{x}]$ убывает со скоростью $1/n$, тогда как стандартное отклонение статистического среднего убывает со скоростью $1/\sqrt{n}$. Установлено также, что
- Если в выражении (5) обе составляющие ошибки не равны нулю, то

  $$\underset{n\to\infty}{[\bar{x}]} \to [x_0 - \Delta_2(x);\ x_0 + \Delta_2(x)], \tag{7}$$

  т. е. при любом числе измерений полу ширина интервального среднего будет не меньше значения Δx_2. Это свойство делает бессмысленным увеличение числа измерений, если ширина интервального среднего соизмерима со значением Δx_2. Этот факт обычно игнорируется при статистическом анализе и арифметическое среднее трактуется как состоятельная оценка.

Сравнивая два подхода, применяемых при анализе выборочных измерений, можно констатировать, что в рамках статистической модели неопределенность выборочных значений, не связанная со случайностью, игнорируется. При этом становится необоснованным утверждение о состоятельности оценок и возможности достижения любой заданной точности оценки выборочного среднего путем увеличения числа повторных измерений.

Ошибки косвенных измерений и эмпирических формул

Данные прямых измерений часто используются для нахождения по известным формулам так называемых косвенных измерений. Кроме того, в технических, экологических и экономических приложениях наряду с объективными законами широко используются так называемые “эмпирические модели”, которые включают экспериментальные, неточно задаваемые переменные. Обе задачи приводят к необходимости определения ошибки результирующего значения функции $y = f(x_1, x_2, \ldots, x_k)$ с неточно заданными параметрами x_i.

Применение вероятностной модели для решения этой задачи сопряжено со значительными трудностями, так как в этом случае необходимо найти распределение случайной величины y при заданном совместном распределении в общем случае зависимых случайных величин x_i, $i = 1, \ldots, k$. Результирующее распределение находится как свертка интегралов. Сравнительно просто находится распределение суммы независимых случайных величин. Однако, уже для отношения даже двух переменных – это нетривиальная операция, при нормальном распределении приводящая к разрыву, в нулевой точке знаменателя.

В рамках интервального анализа исследуемая функция неточных переменных записывается в виде интервальной функции $[y] = f([x_1], [x_2], \ldots, [x_k])$, конечные границы которой находятся как

$$[y] = [y_{\min}; y_{\max}] = [\min_{\vec{x}\in[X]} f(\vec{x}),\ \max_{\vec{x}\in[X]} f(\vec{x})], \quad (8)$$

где $[X]$ – прямоугольная гиперпризма, образованная интервалами $[x_i]$.

Для значительного класса функций, включая функции линейные по интервальным параметрам или позиномы вида

$$[x_1]^{\beta_1} \cdot [x_2]^{\beta_2} \cdot \ldots \cdot [x_k]^{\beta_k}, \quad (9)$$

границы интервала неопределенности $[y]$ могут быть записаны в аналитическом виде, в том числе и в форме, привычной для метрологов

$$[y] = [y_{\min} = y_{\text{ср}} \cdot (1 - \delta_y);\ y_{\max} = y_{\text{ср}} \cdot (1 + \delta_y)], \quad (10)$$

где $y_{\text{ср}}$ – средняя точка интервала неопределенности, $\delta_y \cong \sum_{i=1}^{k} \beta_i \delta_{x_i}$ – относительная ошибка результата, δ_{x_i} – относительные ошибки измерения переменных x_i.

Позиномы (9) широко используются в многочисленных приложениях в качестве эмпирических моделей. При этом для одного и того же явления часто существует несколько альтернативных эмпирических формул, отличающихся как структурой, так и числом переменных. Например, для оценки пространственно-временных характеристик аварийного загрязнения водотоков предложено более 10 различных эмпирических формул, которые отличаются количеством учитываемых переменных. Входящие в них переменные, которые включают скорость течения, глубину и ширину реки, ее гидравлический радиус и уклон, коэффициент шероховатости русла и др., измеряются с большими погрешностями [16].

В этих условиях включение переменной в модель с одной стороны увеличивает ее полноту, т. е. снижает неопределенность результата, а с другой увеличивает ошибку модели, связанную с неточностью измерения переменных. Допустим, что модель (9) содержит полный перечень k переменных. Тогда если в модель включено $m < k$ переменных возникает ошибка из-за ее упрощения и общую ошибку результата можно записать в виде суммы

$$\delta_y(m) = \delta_{y1}(m) + \delta_{y2} = \delta_{y1}(m) + \sum_{i=1}^{m} \beta_i \delta_{x_i}, \quad (11)$$

где первое слагаемое связано с неполнотой модели, а второе – с ошибками измерения переменных. Для величины $\delta_{y1}(m)$ выполняются естественные условия $\delta_{y1}(m) = 0$ при $m = k$ и $\delta_{y1}(m) = 1$, т. е. имеет место максимальная неопределенность при m=0.

Допустим, что вклад каждой переменной в модели одинаков. Тогда при одинаковых ошибках измерения переменных $\delta_{xi} = \delta_x$ и степенях β_i формула (11) приобретает простой вид

$$\delta_y(m) = (1 - m/k) + m \cdot \delta_x. \tag{12}$$

При этом зависимость общей ошибки результата за счет неполноты и неточности измерений от числа переменных имеет минимум, равный $\delta_{\min y} = 1-(1-k \cdot \delta_x)/(1 + k \cdot \delta_x)$ в точке $m_{\text{opt}} = k/(1 + k \cdot \delta_x)$.

Число переменных m_{opt} можно считать оптимальным для эмпирической модели, так как дальнейшее увеличение числа переменных будет приводить к увеличению общей ошибки модели. При $k = 10$ и $\delta_x = 0.05$ оптимальное число переменных для рассматриваемой модели не превышает $6 \div 7$.

Сглаживание данных: прямые, неявные и обратные функции

Наиболее часто в приложениях для сглаживания экспериментальных данных, полученных в результате n-измерений и заданных в виде $(n \times m)$-матрицы X и $(n \times 1)$-вектора Y используется линейно параметризованные функции вида [2, 3]

$$y = \beta_1 f_1(\boldsymbol{x}) + \ldots + \beta_j f_j(\boldsymbol{x}) + \ldots + \beta_m f_m(\boldsymbol{x}) + e, \tag{13}$$

где y – измеряемая выходная переменная, $\boldsymbol{x} = (x_1 \ldots x_k)^{\mathrm{T}}$ – вектор k входных переменных, f_j – известные базисные функции, β_j – неизвестные коэффициенты.

В рамках статистического подхода для нахождения оценок параметров обычно используется множественный регрессионный анализ, основанный на методе наименьших квадратов [1–3]. Точечные оценки параметров и их ковариационная матрица находятся по формулам как

$$\boldsymbol{B} = (\boldsymbol{F}^T \boldsymbol{F})^{-1} \boldsymbol{F}^T \boldsymbol{Y}, \qquad \boldsymbol{D}(\boldsymbol{B}) = \sigma^2 (\boldsymbol{F}^T \boldsymbol{F})^{-1}, \tag{14}$$

где $\boldsymbol{B}$ – вектор оценок неизвестных параметров β_j, $\boldsymbol{F} = \{f_j(\boldsymbol{x}_i)\}$ – $(n \times m)$-матрица значений базисных функций в n опытах, $\boldsymbol{Y}$ – вектор наблюдений выходной переменной.

Известно, что при нормальной аддитивной ошибке оценки (14) обладают целым рядом оптимальных свойств, в частности, являются несмещенными, эффективными и наилучшими в классе линейных оценок. Распределение оценок подчиняется нормальному закону, что позволяет при заданной доверительной вероятности α определить доверительные интервалы для оценок параметров.

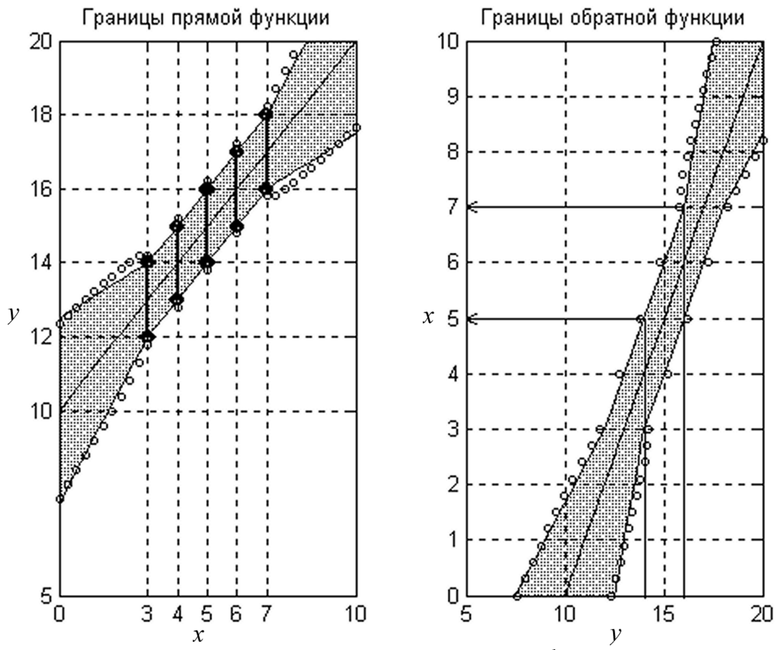
Сложившаяся парадигма регрессионного анализа, восходящая к Фишеру, включает два главных постулата: разделение переменных на входные (независимые) и выходные (зависимые) переменные; присутствие аддитивных, нормально распределенных случайных ошибок только в выходной переменной. Именно при этих допущениях получены наиболее сильные результаты регрессионного анализа, связанные с оптимальностью найденных оценок, их несмещенностью, состоятельностью и т. п.

Вместе с тем, на практике существует широкий класс прикладных задач, когда один или одновременно оба постулата регрессионного анализа нарушаются. На практике модель со случайной, аддитивной нормально распределенной ошибкой является скорее исключением, чем правилом. Обычно ошибки неизбежны как в выходной, так и во входных переменных. При этом их величина часто зависит от текущего значения переменной. Часто нет оснований считать источником неопределенности данных случайность и, следовательно, приписывать ошибкам определенную плотность вероятности. Например, при измерении неточными приборами, анализе ошибок квантования и округления, обработке изображений и т. п. более естественно описывать неопределенность в терминах интервала возможных значений.

В картографии и при обработке фотоизображений отношения между переменными описываются неявной функцией $f(x, y, b_1, \ldots, b_m) = 0$. При этом отсутствует разделение переменных на входные и выходные. Делались многочисленные попытки смягчить достаточно жесткую систему приведенных выше допущений. Например в конфлюэнтном анализе допускается наличие ошибок как на входе так и на выходе. Однако в связи со сложностью аналитического представления функции правдоподобия в основном разработаны эвристические методы, в частности связанные с группированием данных (метод инструментальных переменных).

В рамках интервального подхода задачи сглаживания решаются в предположении, что экспериментальные наблюдения заданы в интервальной форме, при этом допускаются ошибки как во входных, так и в выходной переменной [9–15].

Поясним идею сглаживания интервальных данных на простом примере, рис. 1,а, где вертикальными жирными линиями изображены пять интервальных наблюдений величины y, полученных в эксперименте при $x \in [3; 7]$.



а б

Рис. 1

Для сглаживания интервальных данных и нахождения средней линии, а также верхней и нижней границ интервала неопределенности используются линейные функции. При этом находятся в диапазоне $x \in [3; 7]$ коэффициенты уравнений границ рассчитываются методом наименьших квадратов соответственно по верхним и нижним граничным значениям интервальных наблюдений, а коэффициенты средней линии – по средним точкам интервальных наблюдений. Уравнение средней линии имеет вид $y = 10 + x$.

Очевидно, что в диапазоне $x \in [3; 7]$ через интервальные наблюдения можно провести бесчисленное множество прямых, однако существуют две "экстремальные", одна из которых проходит через точки с координатами (3; 12) и (7; 18), а другая – через точки (3; 14) и (7; 16). Эти прямые порождают более широкий интервал неопределенности вне диапазона $x \in [3; 7]$, ширина тем больше, чем ближе значение x к краям диапазона $[0; 10]$.

Границы результирующего интервала неопределенности в диапазоне $x \in [0; 10]$ (на рис. 1,а выделен темным цветом) описываются линейными сплайн-функциями с тремя участками. Интервальный подход к данной задаче позволяет достаточно просто решить задачу аппроксимации этих сплайн-функций гладкими аналитическими функциями из класса конических сечений. Для решения задачи кроме имеющихся 5 интервальных наблюдений генерируются также дополнительные точки на участках $x \notin [3; 7]$, которые рассчитываются по найденным уравнениям прямых. Аналитические уравнения границ ищутся в форме *неявной функции* [11]

$$x^2 = \beta_1 y^2 + \beta_2 yx + \beta_3 x + \beta_4 y + \beta_5. \tag{15}$$

Векторы коэффициентов уравнения (15), найденные методом наименьших квадратов для данных рис. 1,а, составили:

$$\text{для нижней границы: } \vec{\beta}_{\min} = [-1.05; 2.12; -19.16; 18.35; -78.85] \tag{16}$$

$$\text{для верхней границы: } \vec{\beta}_{\max} = [-1.05; 2.12; -24.36; 23.48; -129.84]. \tag{17}$$

Для целей прогноза уравнения (16) и (17) приводятся к каноническому виду и записываются в виде двух явных функций прогноза $y_{\min} = \varphi(\vec{\beta}_{\min}, x)$, $y_{\max} = \varphi(\vec{\beta}_{\max}, x)$. Полученные предсказанные значения, которые отмечены на рис. 1,а круглыми маркерами, обеспечивают достаточно высокую точность описания границ интервала неопределенности.

В теории измерений важное место занимает *задача калибровки*, т. е. построения градуировочной характеристики сенсора по неточным данным эксперимента. При этом для простейшего линейного случая, когда функция преобразования сенсора имеет вид линейной функции $y = \beta_1 + \beta_2 x$, необходимо найти определить обратную модель $x = a_1 + a_2 y$ и оценить коридор ее ошибок.

В рамках статистической парадигмы эта задача относится к классу наиболее сложных и слабо формализуемых. Ее решению посвящен отдельный раздел статистики – теория калибровки. Трудности вызываются следующими причинами. Если по данным эксперимента, найдены несмещенные статистические оценки коэффициентов β_i прямой модели, то при переходе к обратной модели это свойство теряется, кроме того, возникают теоретические сложности при построении доверительного интервала. Если же рассчитываются оценки коэффициентов обратной модели, то зависимой становится переменная y, измеряемая с ошибками, что приводит к нарушению предпосылок регрессионного анализа.

В рамках интервального анализа задача решается достаточно просто даже при наличии ошибок в обеих переменных x и y, так как при наличии аналитических уравнений границ переход к обратной модели не вызывает никаких трудностей [12, 13]. Если данные рис. 1,а воспринимать как данные калибровочного эксперимента, то, имея уравнения средней линии и границ, легко получить искомую градуировочную характеристику, т. е. обратную модель и ее границы, представленные на рис. 1,б. В частности уравнение градуировочной характеристики имеет вид $x = -10 + y$. Коридор ошибок сенсора залит темным цветом.

На рис. 1,б наглядно проявляется один феномен, который не учитывается при статистическом подходе, связанный с рабочим диапазоном сенсора, в котором ошибка определена на основе исходных данных, а не по прогнозным значениям. Проясним этот момент.

Для прямой модели рис. 1,а в рабочем диапазоне переменной $x \in [3; 7]$ имела место постоянная абсолютная ошибка измерения переменной y, равная ± 1. Для обратной модели постоянная абсолютная ошибка измерения переменной x, равная ± 1, сохраняется лишь в диапазоне $x \in [5; 7]$, который отмечен на рис. 1,б стрелками, которому соответствует сравнительно узкий рабочий диапазон сенсора $y \in [14; 16]$. Вне этого диапазона ошибки измерения x стремительно возрастают.

Этот факт отражает объективную реальность информационной достоверности результатов внутри и вне пределов диапазонов калибровочного эксперимента.

В качестве преимущества интервального подхода к задаче калибровки необходимо отметить также тот факт, что в отличие от статистического подхода могут быть использованы оба варианта расчета: построение прямой модели и ее обращение или непосредственное построение обратной модели. Причем оба варианта приводят к одним и тем же результатам.

Оценка экономических рисков инвестиционных проектов

Для оценки эффективности инвестиционного проекта в Методических рекомендациях [17], основанных на методологии UNIDO, рекомендуется ряд интегральных показателей, среди которых: чистый доход (ЧД), чистый дисконтированный доход (ЧДД), внутренняя норма доходности (ВНД), срок окупаемости (Ток), индекс доходности дисконтированных затрат (ИДДЗ). Предполагается, что ожидаемые показатели эффективности проекта вычисляются с учетом факторов риска и неопределенности.

В соответствии с [17] *риск* для инвестиционных проектов трактуется как совокупность трех составляющих: событие, связанное с риском; ущерб (величина денежной суммы, подвергаемой риску) и *вероятность риска*. Для оценки риска инвестиционных проектов используется так называемый сценарный анализ, когда проект оценивается для конечного числа различных сценариев, каждому из которых приписывается определенная вероятность p_i. Интегральный показатель ожидаемой эффективности проекта $Э_{\text{ож}}$, риск неэффективности проекта $P_{э}$ и средний ущерб $У_{э}$ от реализации неэффективных проектов вычисляются по формулам

$$Э_{\text{ож}} = \sum Э_i p_i, \qquad P_{э} = \sum p_i^-, \qquad У_{э} = \sum p_i^- Э_i / P_{э}. \tag{18}$$

Суммирование в двух последних формулах ведется только для сценариев с отрицательными интегральными эффектами $Э_j$.

Учитывая, что случайность является только одним и не всегда главным источником неопределенности, которая в основном связана с незнанием, неполнотой и неточностью информации в последнее десятилетие в зарубежной литературе предложен ряд новых моделей к описанию неопределенности в терминах не вероятности, а возможности неблагоприятного события. При этом вместо классической вероятностной модели (18) находит более адекватная модель экономического риска, записываемая в виде

Ущерб = неопределенность × денежные потери при неблагоприятном событии.

При этом для описания неопределенности предлагаются разные подходы и теоретические методы, среди которых теория нечетких множеств, робастный байесовский подход, теория неточных вероятностей, модель интервальной вероятности и модель с ограниченными вероятностями. Последние три модели в этом списке рассматриваться как симбиоз вероятностного и интервального подхода.

Ниже рассматривается интервальная модель риска [15, 16, 18], концепция которой основана на преобразовании входных интервальных показателей проекта в выходные интервальные значения интегральных критериев его эффективности.

Интервальная модель экономического риска для проектов, реализованная в программе *Интервал-Инвест* включает следующие этапы:

1. В интервальной форме $[z_{\min}; z_{\max}]$ задаются возможные границы параметров внешних условий и технико-экономических показателей и на их основе программно формируются два экстремальных сценария – пессимистический и оптимистический. В каждом из них в различных комбинациях фигурируют верхние и нижние граничные значения параметров $z_{\min}; z_{\max}$. Вводятся предельно допустимые, пороговые значения d_i для критериев эффективности. Применительно к ЧДД $d = 0$ и при ЧДД < 0 результат проекта по этому критерию трактуется как отрицательный.

2. По известным формулам рассчитываются интервальные критерии $[x_i]$ эффективности проекта. Показано, что значения критериев x_i, которые определяются через кумулятивные значения финансовых потоков, при любых комбинациях исходных данных внутри границ $z_{\min}; z_{\max}$ будут принадлежать интервалам $[x_i]$.

3. Рассчитывается риск (возможность) получения отрицательного результата, когда $x_i < d_i$ (или $x_i > d_i$ для критериев, связанных с затратами).

При расчетах риска r неблагоприятного события $x < d$ при заданном интервальном значении критерия $[x]$ используется следующий достаточно естественный подход:

- если $d \geq x_{\max}$, то $R = 1$;
- если $d \in [x]$, то риск вычисляется как отношение длины отрезка, на котором отношение выполняется, к ширине всего интервала $[x]$, т. е. по формуле
$$R = (d - x_{\min})/(x_{\max} - x_{\min}). \tag{19}$$

Необходимо отметить, что в случае постулирования равномерного распределения на интервале $[x]$ полученные по формуле (19) значения будут совпадать со значением вероятности соответствующего события.

Ниже в табл. 2 приведены результаты расчета критериев эффективности и экономических рисков проекта АЭС, полученные по программе *Интервал-Инвест*. Пороговые значения критериев составляли: для ЧД и ЧДД $d = 0$ (не менее, чем), для ВНД $d = 0.25$ (не менее, чем), для ИДДЗ $d = 1.2$ (не менее, чем), для тарифа безубыточности $d \leq 2$ (не более, чем).

Таблица 2

Показатели эффективности / сценарии	Пессим.	Базовый	Оптим.	Риск
1. Чистый доход, млн. долл.	−1393.1	1222.2	3287.0	0.30
2. ЧДД, млн. долл.	−2373.3	−855.1	521.8	0.82
3. ВНД, %	0.0	2.5	7.5	1.00
4. ИДДЗ	0.36	0.76	1.34	0.86
5. Тариф безубыточности, цент/кВт·ч	5.53	2.98	1.87	0.53
6. Срок окупаемости, лет	Не окуп.	Не окуп.	27	
7. Дисконтированные затраты, млн. долл.	3556	2673	2484	

В экономической литературе рекомендуются следующие градации риска: $0 \div 0.1$ – минимальный; $0.1 \div 0.3$ – малый; $0.3 \div 0.4$ – средний; $0.4 \div 0.6$ – высокий; $0.6 \div 0.8$ – максимальный; $0.8 \div 1$ – критический. В соответствии с этой градацией для исследуемого в условиях неопределенности проекта АЭС критические значения риска получены по всем критериям за исключением ЧД (0.3) и тарифа безубыточности (0.53). Это позволяет сделать вывод о неэффективности проекта при сделанных допущениях относительно возможных сценариев реализации проекта и его технико-экономических характеристиках.

Заключение

Приведенные иллюстративные примеры показывают, что применение интервального анализа позволяет снять многие проблемы и методические сложности, возникающие при решении прикладных задач статистическими методами. В рамках интервального анализа неопределенность исходных данных может иметь разные источники и природу. Интервал неопределенности позволяет описать широкий класс неопределенных, неоднозначных, вариабельных и неточных исходных данных. Значения ошибок в исходных данных могут колебаться в широких пределах. Результаты, полученные с помощью в рамках парадигмы интервального анализа, имеют ясную и четкую интерпретацию в терминах интервалов и областей неопределенности. Основной проблемой интервального анализа является корректное определение интервалов неопределенности на основе различной исходных данных при наличии различных источников неопределенности переменной. Приведенные результаты дают основание сделать заключить, что применение интервального анализа к широкому спектру прикладных задач с ограниченными ошибками и неопределенностью в данных имеет больших перспективы.

Список литературы

[1] Вентцель Е.С. Теория вероятности. – М.: Физматгиз, 1962. – 563 с.

[2] Химмелблау Д. Анализ процессов статистическими методами. – М.: Мир, 1973. – 948 с.

[3] Дрейпер Н., Смит Г. Прикладной регрессионный анализ. – М.: Финансы и Статистика, 1986. – 345 с.

[4] Bellman G., Zadeh L. Decision making in fuzzy environment // Management Science. – 1970. – Vol. 17, № 1.

[5] Zimmerman H. Fuzzy Sets. Decision making and expert systems. – Dordrecht: Kluwer Acad. Publishers, 1987.

[6] Moore R. E. Interval analysis. – Englewood, Cliffs, N.Y.: Prentice-Hall, 1966.

[7] Алефельд Г., Херцберг Ю. Введения в интервальные вычисления. – М.: Мир, 1987. – 370 с.

[8] Калмыков С.А., Шокин Ю.И., Юлдашев З.Х. Методы интервального анализа. – Новосибирск: Наука, 1986.

[9] Шарый С.П. Интервальные алгебраические задачи и их численное решение. – Дисс. … докт. физ.-мат. наук. – 2000. – 322 с.

[10] Вощинин А.П., Бочков А.Ф., Сотиров Г.Р. Интервальный анализ данных как альтернатива регрессионному анализу // Заводская Лаборатория. – 1990. – № 7. – С. 76–81.

[11] Вощинин А.П. Метод анализа данных с интервальными ошибками в задачах проверки гипотез и оценивания параметров неявных и линейно параметризованных функций // Заводская Лаборатория. – 1998. – №7.

[12] Вощинин А.П., Скибицкий Н.В. Интервальный метод калибровки // Системы и датчики. – 2000. – № 7.

[13] Voschinin A., Skibitsky N. Interval calibration model of multisensor systems // IDAACS'2003. – Lviv, Ukraine, 2003.

[14] Вощинин А.П. Интервальный анализ данных: развитие и перспективы // Заводская лаборатория. – 2002. – № 1.

[15] Voschinin A., Tyurin A. Interval identification of time series using readings with bounded errors // 5th International Scientific-Technical Conf. ProcessControl, Pordubice. – 2002.

[16] Gorsky V., Shvetzova-Shilovskaya T., Voschnin A. Risk assessment of accident involving environmental high-toxicity substances // J. of Hazardous Materials. – 2000. – № 78.

[17] Виленский П.Л., Лившиц В.Н., Смоляк С.А. Оценка эффективности инвестиционных проектов (теория и практика). – М.: Изд-во "Дело", 2001.

[18] Вощинин А.П., Тюрин А.В. Интервальные модели оценки эффективности инвестиционных проектов электроэнергетики // Тр. междунар. конф. Control 2003. – Москва, 2003.