СЕКЦИОННЫЕ ДОКЛАДЫ
«МОДЕЛИРОВАНИЕ И АНАЛИЗ БЕЗОПАСНОСТИ И РИСКА В СЛОЖНЫХ СИСТЕМАХ»

УДК 519.24

НАДЕЖНОЕ МОДЕЛИРОВАНИЕ В ЗАДАЧАХ УПРАВЛЕНИЯ ИНВЕСТИЦИОННЫМ РИСКОМ В УСЛОВИЯХ НЕПОЛНОЙ ИНФОРМАЦИИ

Б.С. Добронец, О.А. Попова (Красноярск)

Введение

В задачах оценки и управления рисками важное значение отводится проблеме надежного моделирования. С позиции надежности результатов моделирования пользователь должен иметь систему убедительных доказательств, что полученным результатам можно доверять. При этом ему необходимы знания о свойствах полученного решения, о степени доверия результатам моделирования в виде статистических и других аргументов, подкрепленных интерпретируемыми и визуальными доказательствами. Убедительные доказательства могут быть обеспечены многими факторами.

С позиции численного моделирования к надежности математических моделей, имеют прямое отношение требования корректной постановки вычислительной задачи, учет всех ограничений и неопределенностей, в том числе присутствующих в данных, обеспечение связи численных расчетов и аналитических (символьных) преобразований, обоснование выбора вычисленного метода, позволяющего получить нужный результат с приемлемыми затратами вычислительных ресурсов, выбор языковых средств для описания алгоритмов, и многое другое.

Данная ситуация актуализируется, если исследуемая проблема характеризуется высоким уровнем неопределенности, неполнотой входной информации, влиянием случайных факторов, поведенческой нестабильностью и так далее. В этом аспекте, можно утверждать, что неопределенность несет свою долю риска и имеет свою цену.

В рамках данной статьи с целью получения дополнительных оснований к надежности результатов моделирования в условиях неполной и неопределенной информации рассматривается подход, основанный на применении вычислительного вероятностного анализа (ВВА). Данный подход позволяет количественно оценивать неопределенности [1] и реализует основной тезис новой парадигмы вычислительного моделирования «Distributions are the Numbers of the Future» (Распределения есть числа будущего) [2].

Парадигма «Distributions are the Numbers of the Future» начала формироваться как новое направление в анализе данных с начала 80-х годов прошлого века. В самом общем случае суть ее состоит в том, что вычислительная модель работает с распределениями как с обычными числами. В рамках этой парадигмы стали развиваться новые направления анализа данных такие, как символьный и функциональный [3-5]. Основная проблема такого подхода состоит в необходимости разработки алгоритмов оценки функций от распределений, в том числе вероятностных арифметик над распределениями.

Одной из первых работ, отвечающих этой парадигме в России, является работа [6]. В этой статье гистограммы рассматриваются как математические объекты или гистограммные числа, для которых предлагается гистограммная арифметика, реализованная авторами для задач сетевого планирования. Важно отметить, что на возможность оценки рисков с помощью гистограммной арифметики указывает также работа [7].

Отметим, что новая парадигма, прежде всего, определяет важность разрабатывать и применять к обработке и анализу входных данных численные методы над распределениями, при этом эмпирические данные прежде, чем стать входными данными модели, уже на этапе предобработки представляются в виде распределений, и моделирование осуществляется уже над распределениями, подобно числам. При этом важно, чтобы и результаты моделирования, как выходные данные модели, имели вид распределений, что крайне важно для прогнозной аналитики в задачах управления рисками.

На примере задачи вывода на рынок нового товара мы рассмотрим новый подход к оценке инвестиционных рисков в условиях неполной информации. Для этого используем модель Гомперца и применим технику, разработанную на основе вычислительного вероятностного анализа, который основан на построении распределений второго порядка и численных операциях над ними. Такой подход позволит получить результаты моделирования инвестиционного риска в виде вероятностного поля значений рисковых показателей.

Количественная оценка неопределенности

В настоящее время интенсивно развивается группа методов имеющая общее название «Количественная оценка неопределенности» (Uncertainty Quantification (UQ)) [8]. Методы количественной оценки неопределенности играют ключевую роль в снижении влияния неопределенностей, как в процессах оптимизации, так и в процессах оценки рисков и принятии решений. Основная цель UQ – это построение выходных распределений результатов численного моделирования y. В рамках UQ мы предполагаем, что исследуемая модель зависит от некоторого числа параметров. В отличие от обычных методов численного моделирования UQ предполагает использование знание распределений входных параметров. К UQ можно отнести методы Монте-Карло (МК), при всех своих положительных качествах МК имеет крайне низкую скорость сходимости.

Для получения адекватных оценок о статистических характеристиках результатов моделирования исходную задачу необходимо решить несколько тысяч раз. Стремление снизить число операций привело к использованию в рамках UQ полиномиальное разложение хаоса (ПРХ). Число решений исходной задачи в случае ПРХ растет по экспоненте от числа параметров. Конкурентом методам, основанным на ПРХ, является вычислительный вероятностный анализ (ВВА). Основа ВВА – вычисление вероятностных расширений, т.е. построение законов распределений функций $f(x_1,\ldots,x_n)$ от случайных аргументов. И хотя в общем в случае для построения вероятностных расширений число обращений к вычислению f растет по экспоненте от числа параметров, в ВВА существует подходы позволяющие преодолеть этот барьер.

Функции плотности вероятности случайных величин x, y, z будем обозначать фонтом bold $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$. В тех случаях, когда надо указать непосредственно значение $\boldsymbol{f}$ в некоторой точке ξ, будем использовать обозначение

$$\boldsymbol{z}(\xi) = \boldsymbol{f}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)(\xi).$$

Оценка функций распределения в условиях неполной информации

Оценка функций распределения по эмпирическим данным всегда содержит неточности. Особенно это актуально в условиях недостаточного объема информации. Это приводит к неопределенности вероятностных оценок. Существуют различные подходы оценки неточной вероятности, такие как вероятности второго порядка [9].

Предположим, что у нас есть оценка вероятности P некого события. Когда мы можем оценить границы $P_1 \le P \le P_2$, то можно говорить об интервальной оценке вероятности. Такие интервальные оценки для гистограмм приводят к интервальным гистограммам. В тех случаях, когда для P существует вероятностная оценка, можно говорить о вероятности второго порядка. Таким образом, вероятности второго порядка – это вероятностные оценки самих вероятностей.

В вероятностном пространстве (Ω, F, P) случайный процесс представляет собой набор случайных величин

$$\{a(x,\omega), x \in D, \omega \in \Omega\}. \tag{1}$$

Термин «случайное поле» обычно относится к случайному процессу, принимающему значения в евклидовом пространстве R^d $d = 1,2,3$. Случайное поле можно посмотреть двумя способами:

– для фиксированного $x \in D, a(x,\cdot)$ является случайной величиной в Ω;

– для фиксированного $\omega \in \Omega, a(\cdot,\omega)$ является реализацией случайного поля в D.

Одним из подходов, моделирующим неопределенность вероятностных оценок, стало использование кусочно-полиномиальных моделей распределений второго порядка [12].

Определение 1.[10] Распределение второго порядка $\boldsymbol{f}^{(2)}$ — случайное поле $f(x,\omega)$, $x \in D, \omega \in \Omega$ заданное на $D \subset R$, где (Ω, F, P) — вероятностное пространство. Обладает следующими свойствами: для фиксированного $\omega \in \Omega$, $f(\cdot,\omega)$ является функцией распределения f_{ω}.

Оценка рисков инвестиционных проектов

Рассмотрим задачу оценки рисков инвестиционного проекта выпуска нового продукта. Внедрение новых продуктов на рынок может обеспечить конкурентное преимущество, а также долгосрочную финансовую отдачу от инвестиций. Однако выход нового продукта на рынок является длительным процессом с высоким риском неудачи, которая во многом определяется неполнотой информации, существенно влияющей на надежность полученных результатов. Поэтому применение методов количественной оценки неопределенности в условиях неполной информации представляет актуальную задачу для управления рисками и принятия управленческих решений.

Для оценки привлекательности инвестиционных проектов можно использовать показатели, такие как чистая текущая стоимость (NPV) и внутренняя норма доходности (IRR). Чистая приведенная стоимость (NPV) это формула, используемая для определения текущей стоимости инвестиций по дисконтированной сумме всех денежных потоков, полученных от проекта [11].

NPV – это сумма приведенных к текущему моменту времени чистых денежных потоков по инвестиционному проекту. Данный показатель определяется по следующей формуле:

$$NPV = \sum_{i=1}^{T} \frac{CN_i}{(1+d)^i},$$

где T – расчетный срок инвестиционного проекта в годах; CN_i – денежные потоки за i-й период времени; d – ставка дисконтирования.

Денежные потоки $CN_i = N(t_i)C(t_i, N(t_i))$ можно оценить, зная $N(t_i)$ – число продаж в i-й период времени и цену C_i. Цена в общем случае является функцией от

продаж $C(t_i, N(t_i))$ и определяется менеджерами [11]. Инвестиционный проект признается эффективным в случае, если $NPV > 0$.
IRR – расчетная ставка дисконтирования, при которой чистый дисконтированный доход (NPV) равен нулю. IRR определяется из уравнения

$$\sum_{i=1}^{T} \frac{CN_i}{(1+IRR)^i} = 0.$$

Инвестиционный проект признается эффективным в случае, если IRR $> d$.

Модель Гомперца

Рассмотрим задачу оценки привлекательности инвестиционного проекта выпуска новой продукции. Для нахождения NPV необходимо знать объемы продаж и цены нового продукта. Многие компании называют оценку объемов продаж новых продуктов одной из самых сложных проблем прогнозирования, с которыми они сталкиваются. Такое прогнозирование предполагает «слепой прыжок в будущее», потому что исторических данных, которые могли бы служить ориентиром, мало или совсем нет. Высокая степень неопределенности, связанная с внедрением новых продуктов, только усложняет их прогнозирование. Неточные прогнозы могут привести к ошибочным бизнес-решениям, что может привести к избыточным запасам продуктов или значительному их дефициту. С другой стороны, большинству методов машинного обучения и статистического моделирования требуется определенный объем информации, чтобы сделать (надежный) прогноз.

В работе предлагается, как один из способов оценки объемов продаж новых продуктов в виде распределений, специальные модели, например, модель Гомпертца.
Модель Гомпертца (Gompertz model) [12] позволяет прогнозировать распределение продаж за период и совокупных продаж в течение жизненного цикла продукта. Модель Гомпертца разработана путем экспоненциального наклона модели. Благодаря параметру наклона эта новая модель является гибкой и способна описывать более широкий диапазон форм по сравнению с существующими моделями. Модель с наклоном Гомпертца (tilted-Gompertz model) в большинстве случаев дает более точные квантильные и точечные прогнозы.

Распределение имеет три параметра: λ, σ и ρ это параметры масштаба, наклона и формы соответственно. Вывод модели основан на использовании распределения $F(t), t \in [0,T]$, где:
F(t) – доля возможных покупателей, купивших продукт к моменту времени *t* от совокупности возможных покупателей;
$N(t)$ – количество покупателей к моменту *t*;
$E[N(t)] = mF(t)$ – ожидаемое количество покупателей к моменту *t*, $m = N(t_n)$ – общее число проданных товаров.

Будем считать, что человек покупает новый продукт в неопределенный момент времени. Предполагается, что времена покупок независимы и одинаково распределены в соответствии с функцией распределения F. Следовательно, количество покупателей к моменту t – N(t), а $E[N(t)]$ = mF(t). Чтобы соответствовать этому представлению о процессе, нужно максимизировать вероятность наблюдения данных: число покупок каждый i-й день $y_1, y_2, ..., y_n$, выбирая m и параметры функции F, где каждый

$$y_i = N(t_i) - N(t_{i-1}).$$

$N(t_n) = m$ – общее число проданных товаров. Мы предполагаем, что нам известен временной ряд покупок каждый день $y_1, y_2, ..., y_n$, взятый при продажах аналогичного продукта.

Прогнозы на основе модели Гомпертца могут быть сделаны до запуска продукта на основе данных анализа конкурентов, особенно данных о продажах продукта. Распределение имеет три параметра: λ, σ и ρ, это параметры масштаба, наклона и формы соответственно.

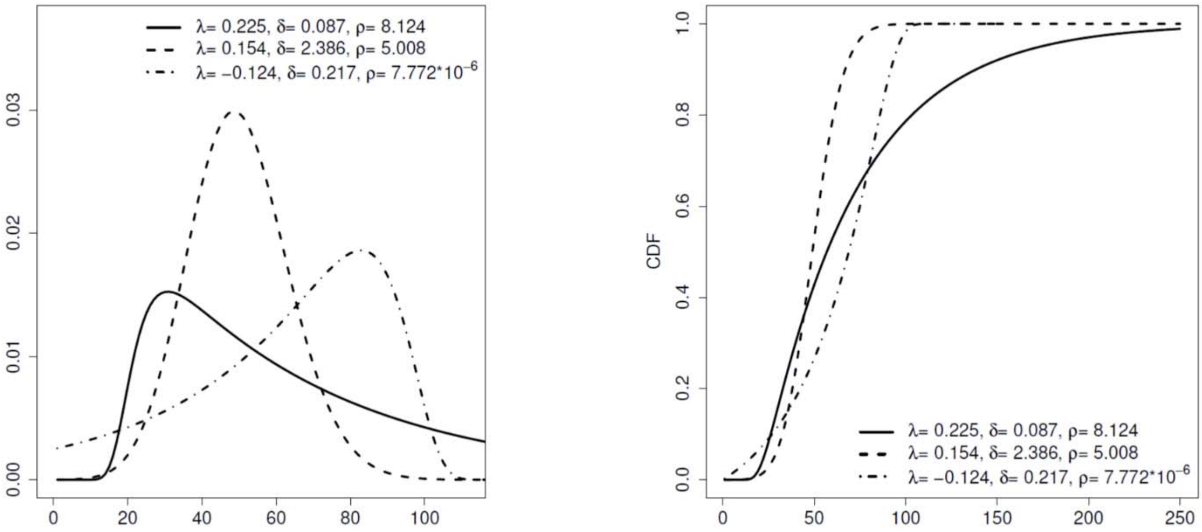

Рис. 1. Функции плотности вероятности и функции распределения
при различных значениях λ, σ, и ρ

На рис. 1 представлены вид функции плотности вероятности и функции распределения модели Гомпертца при различных значениях λ, σ, и ρ. Таким образом можно построить прогноз по числу продаж нового продукта. Число продаж Y_i в момент времени t_i определяется, как

$$Y_i = Y(t_i, \lambda, \sigma, \rho) = \left(N(t_i) - N(t_{i-1})\right), i = 2, \dots, n.$$

Ниже рассмотрим основные этапы построения надежных оценок Y_i.

Построение надежных оценок модели Гомпертца

Пусть известны значения моментов времени $t_j, j = 1, \dots, M,$ такие что $N(t_j) - N(x_{t-1}) = N(t_n)/M$ равные доли всех проданных товаров, где $N(t_n)$ — число всех проданных товаров. Для данного ряда $y_1, y_2, \dots, y_n$ значения t_j — случайные величины. Тогда, поскольку F функция распределения, то в этом случае можно определить $z_i = F(x_i), i = 1,\dots,M$. Заметим, что $z_j, j = 1,\dots,M$ – равномерно распределенные случайные величины на $[0,1]$. Если $z_1 \le z_2 \le \dots \le z_n$, тогда $z_k -$ k-я порядковая статистика и математическое ожидание $\mathsf{M}[z_k] = k/(M+1)$.

Заметим, если бы мы могли вместо математических ожиданий $z_k = k/(M+1)$ использовать неизвестные точные значения z_j, то можно применяя, например, метод наименьших квадратов (МНК) определить константы λ, σ, и ρ

$$\sum_j \left(z_j - N\left(t_j, \lambda, \sigma, \rho\right)\right)^2 \to min.$$

каждому набору $\{z\}_k$ соответствует свой набор λ_k, σ_k и ρ_k. Поскольку, $\{z\}$ – порядковые статистики и для них известна вся необходимая дополнительная информация, то перебирая все возможные наборы $\{z\}$, можно оценить функции распределений λ, σ, ρ и, соответственно, построить надежную оценку распределения F в виде распределения второго порядка $F^{(2)}$ [13].

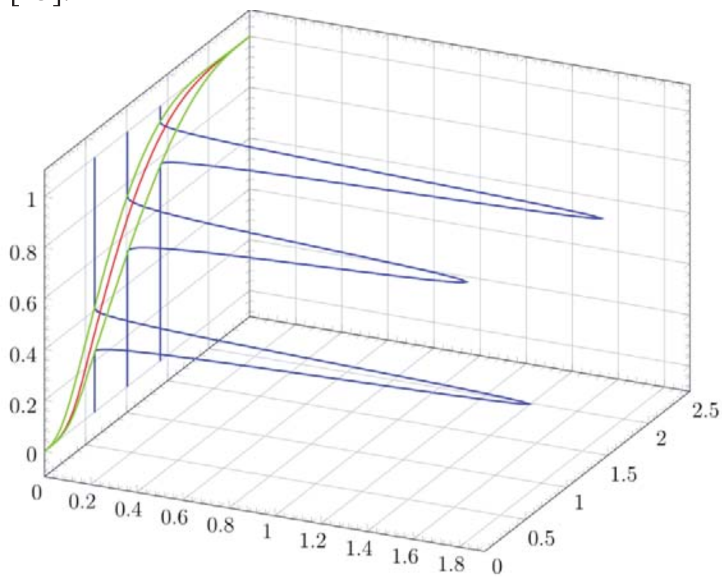

Рис. 2. Надежная оценка функции распределения модели Гомпертца
в виде распределения второго порядка

На рис. 2 представлена надежная оценка функции распределения модели Гомпертца, в виде распределения второго порядка. Красная линия – математическое ожидание распределения второго порядка, синие линии – функции плотности вероятности распределения второго порядка $F^{(2)}(t)$ при фиксированных значениях t, зеленые линии – границы носителей $F^{(2)}(t), t \in [0, 2.5]$.

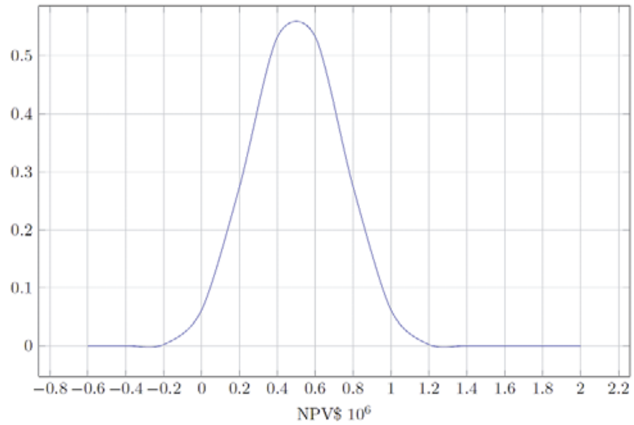

Рис. 3. Функция плотности вероятности NPV

Построение распределений числа продаж

На основе распределения второго порядка построим вероятностное распределение числа продаж Y от времени как вероятностное расширение функции продаж

$$\boldsymbol{Y}_i = \boldsymbol{Y}(t_i, \boldsymbol{\lambda}, \boldsymbol{\sigma}, \boldsymbol{\rho}).$$

Зная $\boldsymbol{Y}_\mathrm{i}$ мы можем построить и функции распределения вероятности денежных потоков CN_i. Таким образом, используя вероятностные арифметики можно построить функции плотности вероятности NPV и IRR.

На рис. 3 показана функция плотности вероятности NPV. Носитель NPV [$ -

0.704 106, $ 2.07 106].

Зная функцию плотности вероятности NPV можно подсчитать вероятность риска, как неудачи, так и значительных прибылей. Подсчитав интеграл под функцией плотности вероятности NPV от $-\infty$ до 0 можно узнать риск неудачи, значение интеграла от 1.0 до ∞ определит риск получения прибыли свыше $1.0 \cdot 10^6$ условных единиц.

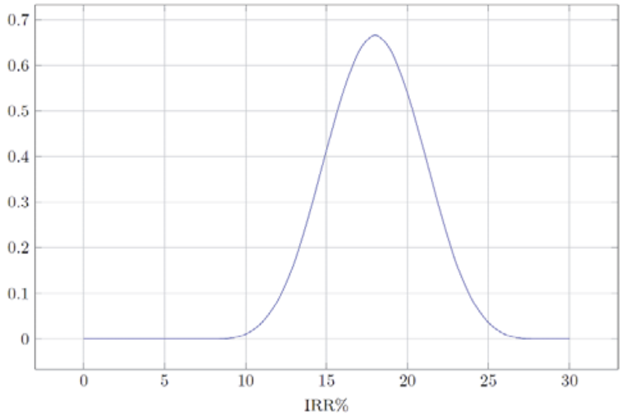

Рис. 4. Функция плотности вероятности IRR

На рис. 4 показана функция плотности вероятности показателя IRR. Носитель IRR расположен на отрезке [5%, 30%].

Заключение

Модельный пример показал, что высокая степень неопределенности инвестиционных проектов, связанная с внедрением новых продуктов, может быть преодолена за счет применения новых подходов, соответствующих новым парадигмам.

Применение распределительной модели Гомперца и метода построения распределений второго порядка позволило в условиях неполной информации получить вероятностное поле возможных значений оценок показателей инвестиционного риска, тем самым существенно расширив систему знаний для принятия управленческих решений.

Знание прогноза функций плотности вероятности продаж позволяет, используя вычислительный вероятностный анализ, построить функции плотности вероятности показателей инвестиционных проектов, такие как NPV и IRR, что дает возможность оценить риски инвестиционных проектов. Зная функцию плотности вероятности NPV, можно подсчитать вероятность риска как неудачи, так и значительных прибылей

Литература

1. **Добронец Б.С., Попова О.А.** Вычислительный вероятностный анализ: модели и методы. Сибирский федеральный университет, Институт космических и информационных технологий. Красноярск, 2020.
2. **Schweizer B.** Distributions are the numbers of the future. In Proceedings of the mathematics of fuzzy systems meeting (Naples, Italy), (1984). pp. 137-149.
3. **Billard L. and Diday E.** Symbolic Data Analysis: Conceptual Statistics and Data Mining (1st ed.), Chichester, John Wiley & Sons, 2006.
4. **Brito P., and Dias S.** Analysis of Distributional Data, Taylor & Francis Group, 2022
5. **Ramsay J.O. and Silverman B.W.** Functional Data Analysis. Springer 2nd ed., New York, 2005.
6. **Герасимов В.А., Добронец Б.С., Шустров М.Ю.** Численные операции гистограммной арифметики и их приложения // Автоматика и телемеханика. 1991. № 2. С. 83-88.

7. **Moore R.E.** Risk analysis without Monte Carlo methods // FreiburgerIntervall-Berichte, No. 84/1, 1984, pp. 1-48.
8. **Tennoe S., Halnes G., Einevoll G.T.** Uncertainpy: A Python Toolbox for Uncertainty Quantification and Sensitivity Analysis // Computational Neuroscience. Front. Neuroinf. (2018).
9. **Augustin T., Coolen F., Cooman G., Troffaes M.** Introduction to Imprecise Probabilities. 2014 John Wiley & Sons, Ltd.
10. **Dobronets B. Popova O.** Second-order distributions: construction, operations, applications // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. 61. С. 61-68.
11. **Лукашов А. В.** Метод Монте-Карло для финансовых аналитиков: краткий путеводитель // Управление корпоративными финансами. 2007. Т. 1, № 19. С. 22–39.
12. **Guo X.** Probabilistic Forecasting in Decision-Making: New Methods and Applications. 2020. Doctoral thesis (Ph.D), UCL (University College London). P. 154.
13. **Добронец Б.С., Попова О.А.** Надежные оценки эмпирических распределений в условиях малых выборок // XXVII Международная конференция по мягким вычислениям и измерениям (SCM’2024). 2024. Т.1, С. 40-43.