УДК 624.046.5

А.А. Соловьева[1], В.А. Смирнов[2]
[1]Вологодский государственный университет,
[2]НИУ Московский государственный строительный университет

МЕТОДЫ ОЦЕНКИ ПЛОТНОСТИ ВЕРОЯТНОСТЕЙ ДЛЯ РЕЗЕРВА НЕСУЩЕЙ СПОСОБНОСТИ ЭЛЕМЕНТОВ СТАЛЬНЫХ ПОКРЫТИЙ ПРИ СТАТИСТИЧЕСКИХ ИСПЫТАНИЯХ

В статье представлены алгоритмы оценки плотности вероятностей для резерва несущей способности элементов стальных покрытий при статистических испытаниях на основе метода ядерной оценки плотности вероятностей и метода k-ближайших соседей. При стохастическом анализе надежности методов статистических испытаний имеется существенный недостаток – при сложном детерминированном расчете конструкции здания или сооружения, реализация большого числа испытаний потребует высоких вычислительных мощностей и длительного времени реализации. При этом оценка вероятности отказа по частоте будет иметь низкую точность, а подбор функции распределения по малой выборочной совокупности данных для резерва прочности может привнести большую долю субъективности в расчет. Непараметрические подходы к оценке плотности распределения вероятностей позволяют комплексно проанализировать проблему анализа распределения вероятностей для резерва несущей способности элементов стальных покрытий при статистических испытаниях. По результатам численных примеров можно сделать вывод о том, что при малом числе статистических данных существует значительный разброс в доверительных оценках параметров и функций распределения вероятностей. Для моделирования такого рода неопределенностей могут быть использованы p-блоки (probability box).

Надежность, статистические испытания, неопределенность, метод k-ближайших соседей, метод ядерной оценки, стальные конструкции покрытий, безопасность, анализ данных.

Для применения вероятностных методов расчета требуется формирование моделей базисных переменных. Обычно выделяют две группы базисных переменных: первые определяют модели воздействия (например, нагрузку), вторые – модели сопротивления (например, несущую способность).

В фундаментальной монографии по теории надежности В.Д. Райзера и О.В. Мкртычева [1] отмечается, что преимуществом метода статистических испытаний является его простота и универсальность. В то же время метод обладает существенным недостатком – при оценке малых вероятностей отказа с приемлемой достоверностью может потребоваться достаточно большое число испытаний. И если на каждом испытании выполняется сложный детерминированный расчет, то общая потребность времени расчета становится непомерно большой и метод становится неэффективным.

Для решения данной проблемы разрабатываются методы вероятностного анализа надежности элементов зданий и сооружений при неполной статистической информации [2, 3].

После реализации метода статистических испытаний, как правило, имеется выборочная совокупность данных (например, максимальные напряжения в элементе при каждой реализации):

$$\tilde{x} = \{x_1,\ x_2, \ldots ,\ x_n\}, \tag{1}$$

где n – число проведенных статистических испытаний.

При наличии такого подмножества данных в дальнейшем, как правило, используют два подхода:

– если выборка достаточно большая или вероятность безотказной работы мала, то надежность оценивают вероятностью по частоте – путем отношения количества успешных генераций, где напряжения не превысили предельные, к общему числу генераций;

– если выборка невелика, то по параметрам выборки строится вероятностная модель на базе известного распределения вероятностей, как правило, нормального. Вероятность безотказной работы рассчитывается известными методами теории надежности сооружений [1, 2].

Однако подбор распределения вероятностей для второго случая представляет собой субъективный выбор аналитики данных. Существует ряд непараметрических подходов к оценке плотности распределения вероятностей, использование которых позволит комплексно проанализировать проблему анализа распределения вероятностей для резерва несущей способности элементов стальных покрытий при статистических испытаниях. В настоящей работе рассмотрены методы ядерной оценки плотности и k-ближайших соседей в контексте оценки плотности распределения резерва несущей способности элементов стальных покрытий.

Восстановить плотность вероятностей резерва несущей способности элемента можно с использованием ядерной оценки плотности (или метод окна Парзена – Розенблатта) [4, 5]. Ядерная оценка плотности вероятностей – это непараметрический способ оценки плотности вероятности случайной величины. Пусть имеется выборочная совокупность данных случайной величины в виде вычисленного резерва несущей способности элемента строительной конструкции путем статистических испытаний: $\tilde{x} = \{x_1,\ x_2, \ldots\ ,\ x_n\}$. Тогда функция $\hat{f}_X(x)$ в форме оценки плотности вероятностей величины $f_X(x)$ может быть записана в виде выражения:

$$\hat{f}_X(x) = \frac{1}{h \cdot n} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \qquad (2)$$

где n – число испытаний/измерений случайной величины $\tilde{x}$; h – сглаживающий параметр, называемый «ширина полосы» (bandwidth), принимаемый $h > 0$; K – ядро (ядерная функция), которое представляет собой неотрицательную функцию.

Существуют различные виды ядерных функций K:

- прямоугольная:

$$K(t) = \begin{cases} 0{,}5 \ \ \textit{если} \ |t| < 1 \\ 0 \ \ \textit{если} \ |t| \geq 1 \end{cases},$$

- треугольная:

$$K(t) = \begin{cases} 1 - |t| \ \ \textit{если} \ |t| < 1 \\ 0 \ \ \textit{если} \ |t| \geq 1 \end{cases},$$

- параболическая (Епанечникова):

$$K(t) = \begin{cases} \dfrac{3}{4\sqrt{5}}\left(1 - \dfrac{1}{5}t\right)^2 \ \ \textit{если} \ |t| < \sqrt{5} \\ 0 \ \ \textit{если} \ |t| \geq \sqrt{5} \end{cases},$$

- нормальная и др. [6].

В связи с тем, что в строительном проектировании множество случайных величин имеют по форме распределение вероятностей близкое к нормальному (логнормальное, логистическое, распределение экстремальных значений и др.), то и в качестве ядерной функции зачастую применяется стандартная нормальная (гауссовская) функция в виде

$$K(t) = \phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}.$$

Для оценки сглаживающего параметра, «ширины полосы» (bandwidth) h распределений, близких к нормальным, используется следующее уравнение:

$$h = 1{,}06 \cdot n^{\frac{-1}{5}} \cdot S_x, \qquad (3)$$

или правило Сильвермана [7]:

$$h = 0{,}9 \cdot n^{\frac{-1}{5}} \cdot \min\left(S_x,\ \frac{IQR}{1{,}34}\right), \qquad (4)$$

где S_x – среднеквадратическое отклонение выборочной совокупности данных; n – число данных в выборочной совокупности; IQR – межквартильный размах: принимается равной разнице между 75-м и 25-м процентилями (между третьим и первым квартилями $IQR = Q_3 - Q_1$).

В выражение для оценки плотности (2) также могут быть добавлены веса для выборочных значений x_i, если данные были получены различными измерительными приборами или используется информация от различных экспертов:

$$\hat{f}_X(x) = \frac{1}{h \cdot n} \sum_{i=1}^{n} w_i \cdot K\left(\frac{x - x_i}{h}\right), \qquad (5)$$

где w_i – вес значения x_i, $\sum_{i=1}^{n} w_i = 1$.

В [8] приводится формула оценки плотности распределения случайной величины методом ядерной оценки при нормальной ядерной функции $K(t)$ в виде:

$$\hat{f}_X(x) = \frac{1}{S_x \cdot n \cdot \sqrt{2\pi}} \sum_{i=1}^{n} \exp\left[\frac{-(x - x_i)^2}{2 \cdot S_x^2}\right]. \qquad (6)$$

Оценка параметра S_x при малой выборке также бывает затруднительна. В [8] приведен следующий подход к назначению среднеквадратического отклонения в случае малых выборок: среднеквадратическое отклонение определяют из условия

$$\int_{-\infty}^{a} \hat{f}_X(x,\ S_x)dx + \int_{b}^{+\infty} \hat{f}_X(x,\ S_x)dx = p(\alpha,\ n), \qquad (7)$$

где $p(\alpha,\ n) = 1 - \alpha^{\frac{1}{n}}$, α – доверительная вероятность; $a = x_{\min} - \dfrac{1}{2N - 2}(x_{\max} - x_{\min})$ – нижняя граница изменчивости случайной величины; $b = x_{\max} + \dfrac{1}{2N - 2}(x_{\max} - x_{\min})$ – верхняя граница изменчивости случайной величины; $x_{\min}$ и $x_{\max}$ – минимальное и максимальное значение в выборке $\tilde{x} = \{x_1,\ x_2, \ldots\ ,\ x_n\}$.

Альтернативным вариантом для оценки плотности распределения вероятностей резерва прочности является метод k-ближайших соседей (k-nearest neighbors, KNN) [9]:

$$\hat{f}_X(x) = \frac{k}{n} \cdot \frac{1}{\dfrac{\pi^{\frac{p}{2}}}{\Gamma(p/2 + 1)} \|x - x_k\|}, \qquad (8)$$

где n – количество данных выборочной совокупности; p – размерность пространства; $\Gamma(.)$ – гамма функция;

k – назначаемый параметр KNN-метода, как правило, обозначающий порядковый номер ближайшего числа к заданному.

Рассмотрим пример. Пусть по результатам статистических испытаний определялся резерв предельной узловой нагрузки на стальную ферму покрытия при изменчивости снеговой нагрузки, собственного веса конструкций покрытия и предела текучести стали [10]. По результатам 8-ми статистических испытаний были получены следующие значение резерва несущей способности:

$$g \in \{0.5,\ 1.2,\ 1.0,\ 1.1,\ 0.8,\ 1.4,\ 1.1,\ 1.5\}\ \text{кН}.$$

Необходимо оценить плотность вероятности в точке $g = 0$.

Выборочные значения статистических параметров функции предельного состояния: среднее $\overline{g} = 1{,}075$ кН, стандартное отклонение $S_g = 0{,}320$ кН. Доверительные интервалы при малом объеме статистической информации: $m_g = \left[\underline{m}_g,\ \overline{m}_g\right] = \left[0{,}750;\ 1{,}400\right]$ кН, $S_g = \left[\underline{S}_g,\ \overline{S}_g\right] = \left[0{,}064;\ 0{,}576\right]$ кН.

В случае использования нормального распределения по выборочным значениям, значение плотности вероятности в точке 0 составит: $f_{norm}(0) = 0{,}0044$. В случае использования нормального распределения при пессимистическом выборе значений ($m_g = \underline{m}_g$, $S_g = \overline{S}_g$) из доверительного интервала: $f_{norm}(0) = 0{,}2980$.

Метод k-ближайших соседей дает следующую оценку значению плотности распределения вероятностей в точке $g = 0$:

при k=1: $\hat{f}_g(0) = 0{,}125$;

при k=2: $\hat{f}_g(0) = 0{,}156$.

Оценки плотности распределения вероятностей методом ядерной оценки при нормальной ядерной функции по (6):

при $S_g = 0{,}320$ кН: $\hat{f}_g(0) = 0{,}055$;

при $S_g = 0{,}576$ кН: $\hat{f}_g(0) = 0{,}157$.

Из полученных результатов можно сделать вывод о том, что при малом числе статистических данных существует значительный разброс в доверительных оценках параметров и функций распределения вероятностей. Для моделирования такого рода неопределенностей могут быть использованы p-блоки (probability box), которые формируют две граничные функции распределения, внутри области которых находится действительная функция распределения вероятностей.

На основе выражений для плотности распределения вероятностей (7) и (8) по указанному алгоритму можно вычислить вероятность отказа элемента конструкции стального покрытия по формуле:

$$P_f = \int_{-\infty}^{0} \hat{f}_g(g)\,dg\,. \tag{9}$$

Представленный подход может быть использован при анализе надежности стальных покрытий зданий и сооружений [11], когда метод статистических испытаний при сложном детерминированном расчете позволяет получить ограниченное число статистических данных об исследуемом параметре.

Литература

1. Мкртычев, О. В. Теория надежности в проектировании строительных конструкций / О. В. Мкртычев, В. Д. Райзер. – Москва : АСВ, 2016. – 905 с.

2. Соловьев, С. А. Неклассические методы анализа надежности строительных конструкций / С. А. Соловьев. – Вологда : Вологодский государственный университет, 2022. – 135 с.

3. Соловьев, С. А. Вероятностная оценка промышленной безопасности при неполной статистической информации / С. А. Соловьев // Безопасность труда в промышленности. – 2020. – № 9. – С. 88–93.

4. Rosenblatt, M. Remarks on Some Nonparametric Estimates of a Density Function / M. Rosenblatt // Annals of Mathematical Statistics. – 1956. – № 27. – P. 832–837.

5. Parzen, E. On Estimation of a Probability Density Function and Mode / E. Parzen // Annals of Mathematical Statistics. – 1962. – № 33. – P. 1065–1076.

6. Węglarczyk, S. Kernel density estimation and its application / S. Węglarczyk // ITM Web of Conferences. EDP Sciences. – 2018. – Vol. 23. – P. 00037.

7. Silverman, B.W. Density Estimation for Statistics and Data Analysis / B. W. Silverman. – London : Chapman & Hall/CRC, 1986. – 45 p.

8. Pradlwater, H. J. The use of Kernel densities and confidence intervals to cope with insufficient data in validation experiments / H. J. Pradlwater, G.I. Schueller // Computer Methods and Applied Mechanics and Engineering. – 2008. – № 197. – P. 2550–2560.

9. Yen-Chi Chen. Lecture 7: Density Estimation: k-Nearest Neighbor and Basis Approach. – URL: https://faculty.washington.edu/yenchic/18W_425/Lec7_knn_basis.pdf (дата обращения: 06.12.2024). – Text : Electronic.

10. Соловьева, А. А. Вероятностные модели случайных величин в строительном проектировании / А. А. Соловьева, С. А. Соловьев. – Вологда : Вологодский государственный университет, 2024. – 144 с.

11. Соловьев, С. А. Моделирование случайной статической нагрузки на покрытия сооружений при неполной статистической информации / С. А. Соловьев // Строительная механика инженерных конструкций и сооружений. – 2020. – Т. 16, № 4. – С. 243–249.

A.A. Solovyova[1], V.A. Smirnov[2]
[1]Vologda State University,
[2]Moscow State University of Civil Engineering

METHODS OF PROBABILITY DENSITY ESTIMATION FOR RESERVE OF STEEL STRUCTURAL ELEMENTS LOAD-BEARING CAPACITY UNDER STATISTICAL TESTS

The paper presents algorithms of probability density estimation for the reserve of load-bearing capacity of steel elements under statistical tests based on the kernel probability density estimation method and k-nearest neighbors (KNN) method. At stochastic reliability analysis of statistical testing methods there is a significant disadvantage – at complex deterministic calculation in structural design, realization of a large number of tests will require high computing power and longtime of realization. In this case, the estimation of failure probability by frequency will have low accuracy, and the selection of the distribution function from a small sample of data for the strength reserve may introduce a large share of subjectivity into the calculation. A nonparametric approach to the estimation of the probability distribution density function allows analyzing comprehensively the problem of probability distribution analysis for the bearing capacity reserve of steel elements under statistical tests. According to the results of numerical examples, it can be concluded that at small number of statistical data there is a significant scatter in confidence estimates of parameters and probability distribution functions. For modeling of such uncertainties p-boxes (probability box) can be used.

Reliability, statistical tests, uncertainty, k-nearest neighbors method, kernel, steel structures, safety, data analysis.