

МЕТОДИКА ИДЕНТИФИКАЦИИ ЗАКОНА РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ ГРАФОАНАЛИТИЧЕСКИМ МЕТО- ДОМ

к.т.н. С.В. Рудаков, О.Ю. Дубийчук
(представил д.т.н., проф. А.М. Крюков)

Предлагается методика компьютерного определения закона распределения случайной величины для выборки небольшого объема графоаналитическим методом, в котором построение аппроксимирующей функции выполняется по методу наименьших квадратов. Приведен пример определения закона распределения по результатам ресурсных испытаний секций высоковольтных конденсаторов.

Постановка проблемы. Известно [1], что одним из эффективных методов идентификации законов распределения (ЗР) является графоаналитический метод, который позволяет оценить вид ЗР для выборки небольшого объема. С другой стороны, с помощью этого метода возможно оценить эффективность предложенной методики идентификации ЗР.

Суть метода состоит в следующем. На координатной сетке, оси которой закодированы в соответствующем масштабе для конкретного ЗР (вероятностная бумага), наносятся экспериментальные точки. Если эти точки «ложатся» на одну прямую, то их распределение согласуется с этим конкретным ЗР.

Анализ литературы. Если априорно известно, что выборка может подчиняться одному из нескольких ЗР [1, 2], то такие построения выполняются на разных вероятностных бумагах, каждая из которых соответствует своему закону распределения [3]. Наилучшая степень близости прямой, проведенной через экспериментальные точки к теоретическим значениям, и является признаком наилучшего согласия теоретического ЗР с экспериментальными данными.

Недостатком метода является допущение о визуальном проведении прямой через экспериментальные точки, что может привести к ошибкам в интерпретации экспериментальных результатов прямой.

Цель работы – повышение достоверности методики идентификации вида ЗР случайной величины графоаналитическим методом.

Методика идентификации ЗР графоаналитическим методом. Ре-

ализация методики идентификации графоаналитическим методом осуществляется в четыре этапа:

- кодируются оси координат для каждого из возможных законов распределения;
- экспериментальные точки наносятся на преобразованные оси в соответствующем масштабе;
- через экспериментальные точки методом наименьших квадратов проводится прямая и уравнение этой прямой принимается теоретической функцией распределения;
- сравнивается теоретическое и экспериментальное распределение по критерию Смирнова-Колмогорова.

Рассмотрим построение преобразованных координатных осей для нескольких законов распределения, отличающихся существенно формой кривой плотности вероятности.

Экспоненциальный закон распределения.

Для определения выражений для Y и X (осей преобразованных координат) возьмем логарифм от левой и правой частей выражения $\ln(1 - F) = -\lambda x$. Если $Y = \ln(1 - F)$, то получим линейную функцию вида $Y = -\lambda x$. Принимая $X = x$, замечаем, что в координатах (Y, X) экспоненциальный ЗР изображается прямой. Таким образом, если экспериментальные точки легли на одну прямую вида $Y = -\lambda x$, то можно однозначно утверждать, что их распределение подчинено экспоненциальному закону.

Нормальный закон распределения.

Так как связь между квантилем u нормального ЗР и значением случайной величины x линейна $u = \frac{x - \bar{x}}{\sigma}$, то, полагая $Y = u$, а $X = x$, получим координатную плоскость для построения прямой, соответствующей нормальному ЗР. Следует отметить, что для *нормально-логарифмического закона распределения* вид “ u ” меняется $u = \frac{\ln x - \ln \bar{x}}{\sigma[\ln x]}$. Тогда $Y = u$, а $X = \ln x$.

Равномерный закон распределения.

Для равномерного ЗР справедливо следующее соотношение $Y = y$, $X = x$.

Распределение Симпсона.

Для построения осей Y, X воспользуемся тем обстоятельством, что плотность вероятности имеет следующий вид (рис. 1) и описывается такими выражениями:

$$\begin{aligned} f &= kx + b_1; & x_1 \leq \bar{x} \leq x; \\ f &= -kx + b_1, & \bar{x} \leq x \leq x_n; \end{aligned} \quad (1)$$

$$\bar{x} = \frac{x_1 + x_n}{2}. \quad (2)$$

Из рис. 1 следует, что $\operatorname{tg} \alpha = k$. Тогда высота треугольника

$$h = (\bar{x} - x_1) \operatorname{tg} \alpha = k(\bar{x} - x_1), \quad (3)$$

а площадь заштрихованной области S_1 равна

$$S_1 = 1/2(\bar{x} - x_1)h = \frac{k(\bar{x} - x_1)^2}{2}. \quad (4)$$

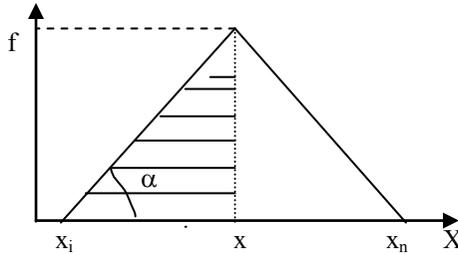


Рис. 1. Вид функции плотности вероятности по закону Симпсона

С другой стороны, учитывая, что f – плотность вероятности, и она симметрична относительно высоты h , то $S_1 = 0,5$ (численно равна функции распределения в средней точке выборки).

Тогда
$$\frac{k(\bar{x} - x_1)^2}{2} = 0,5, \quad (5)$$

откуда
$$k = \frac{1}{(\bar{x} - x_1)^2}, \quad (6)$$

или с учетом (2):

$$k = \frac{4}{(x_n - x_1)^2}.$$

Коэффициент b_1 определим из условия $x = x_1, f = 0$:

$$0 = kx_1 + b_1; \quad b_1 = -kx_1 = -\frac{4x_1}{(x_n - x_1)^2}.$$

Функция распределения на участке $(x_1; \bar{x})$ имеет вид

$$F = \int_{x_1}^x (kx + b_1) dx = \frac{kx^2}{2} + b_1x - \frac{kx_1^2}{2} - b_1x_1 = \frac{2x^2}{(x_n - x_1)^2} - \frac{4x_1 \cdot x}{(x_n - x_1)^2} - \frac{2x_1^2}{(x_n - x_1)^2} + \frac{4x_1^2}{(x_n - x_1)^2} = 2 \left[\frac{x - x_1}{x_n - x_1} \right]^2. \quad (7)$$

Если обозначить $Y = \sqrt{0,5F}$, $X = x$, (8)

то Y будет линейной функцией X на интервале $x_1 \leq \bar{x} \leq x_n$.

Найдем вид Y для интервала $\bar{x} \leq x \leq x_n$. Коэффициент k не изменится. Найдем коэффициент b_2 из условия $x = x_n$, $f = 0$:

$$0 = -kx_n + b_2; \quad b_2 = \frac{4x_n}{(x_n - x_1)^2}. \quad (9)$$

Функция распределения на участке $\bar{x} \leq x \leq x_n$ имеет вид

$$F = \int_{\bar{x}}^x (-kx + b_2) dx + 0,5 = -\frac{kx^2}{2} + b_2x + \frac{k\bar{x}^2}{2} - b_2\bar{x} + 0,5 = \frac{2x^2}{(x_n - x_1)^2} + \frac{4x \cdot x_n}{(x_n - x_1)^2} + \frac{2\bar{x}^2}{(x_n - x_1)^2} - \frac{4x_n\bar{x}}{(x_n - x_1)^2} + 0,5 = -2 \left[\frac{x - x_n}{x_n - x_1} \right]^2 + 1. \quad (10)$$

Преобразуем (10) к виду $\sqrt{0,5(1-F)} = \frac{x - x_n}{x_n - x_1}$.

Обозначим $Y = \sqrt{0,5(1-F)}$. (11)

Тогда Y будет линейной функцией X на интервале $\bar{x} \leq x \leq x_n$.

Таким образом, для закона распределения Симпсона ось ординат Y строится по формулам (8) и (11).

Арсинусоидальный закон распределения.

Учитывая, что функция распределения случайной величины изменяется по закону $\arcsin x$, а именно $F = \arcsin x$, то взяв \sin от обеих частей равенства, получаем $\sin F = x$. Если принять $Y = \sin F$, а $X = x$, то Y будет линейной функцией x .

Функция распределения Вейбулла достаточно часто применяется на практике для описания распределения отказов систем со слабыми звеньями и имеет вид

$$1 - F = e^{-\left(\frac{x}{a}\right)^b}, \quad (12)$$

где a – параметр масштаба, b – параметр формы.

Если принять $Y = \ln[-\ln(1 - F)]$, а $X = \ln x$, то Y будет линейной функцией X .

Построение теоретических функций распределения осуществляется путем наилучшего приближения к экспериментальным точкам. Так как оси Y и X в кодированных значениях являются осями с равномерной шкалой, то построение прямых через экспериментальные точки на каждой из «вероятностных бумаг» можно осуществить методом наименьших квадратов [2].

При этом, если через экспериментальные точки провести прямую $Y = aX + b$, то значения a и b определяются следующими выражениями:

$$a = \frac{-\frac{1}{n} \left(\sum_{i=1}^n X_i \right) \cdot \left(\sum_{i=1}^n Y_i \right) + \sum_{i=1}^n X_i \cdot Y_i}{\sum_{i=1}^n X_i^2 - \frac{1}{n} \cdot a \sum_{i=1}^n X_i}; \quad b = \frac{1}{n} \left(\sum_{i=1}^n Y_i \right) - \frac{1}{n} \cdot a \sum_{i=1}^n X_i, \quad (13)$$

где n – объем выборки.

После нахождения уравнений прямых для каждой из вероятностных бумаг определяется наибольшее отклонение D экспериментальных данных от найденной прямой $D_j = \max |(F_{Ti} - F_i)_j|$, где F_{Ti} , F_i – соответственно значения теоретической и экспериментальной функции распределения в каждой i -й точке, соответствующей i -му экспериментальному значению случайной величины; j – номер закона распределения.

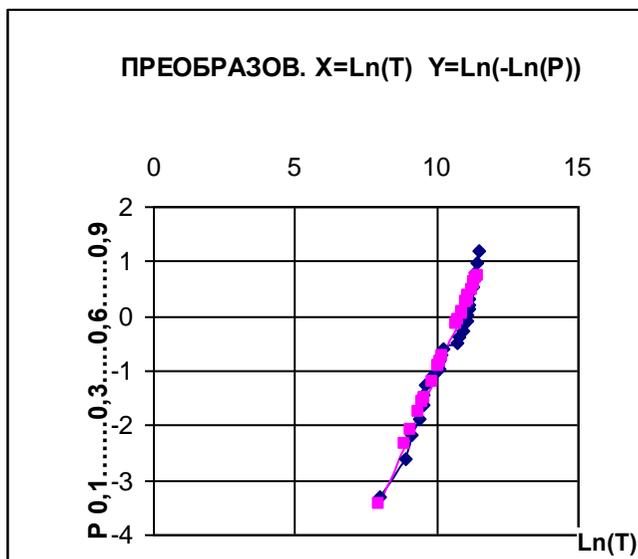
Тот закон, у которого значение D_j окажется минимальным, будет точнее остальных соответствовать экспериментальному распределению. Оценка достоверности согласия выбранного закона осуществляется по критерию Смирнова-Колмогорова [4] для заданного уровня значимости по величине $D_j \sqrt{n} < \lambda$, где λ определяется из справочных таблиц.

Методика реализована в редакторе электронных таблиц *Excel Microsoft* с автоматическим выбором лучшего согласующего ЗР по величине D_j и построением графиков для всех законов распределения.

Рассмотрим пример обработки результатов ресурсных испытаний конденсаторной изоляции. При испытаниях влияния давления на секции высоковольтных импульсных конденсаторов при их запрессовке [5] получены следующие данные по ресурсу (табл. 1). В таблице N – номер опыта, T – число импульсов, приводящее к пробоем секции.

На рис. 2 представлены результаты обработки данных по двум законам распределения при сравнении пяти законов распределения, наиболее часто встречающихся при вычислениях результатов ресурсных испытаний электрической изоляции, по критерию Смирнова-Колмогорова.

Из них лучшим является ЗР Вейбулла ($D_j = 0,112$), далее следует нормальный ЗР ($D_j = 0,126$), равномерный ($D_j = 0,13$), нормально-логарифмический ($D_j = 0,146$), экспоненциальный ($D_j = 0,256$). По оси ординат слева отложена числовая ось, соответствующая значениям функции распределения $F = P$, справа – числовая ось в кодированных координатах (равномерная шкала). Более темные квадратики на графиках – это экспериментальные точки, светлые квадратики – точки теоретической прямой.



а

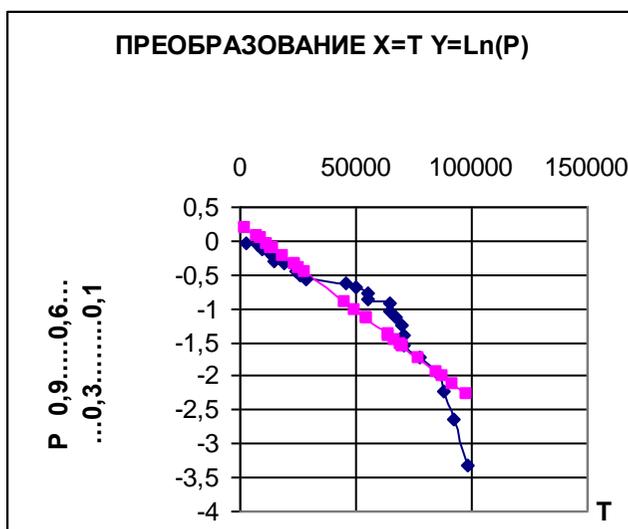


Рис. 2. Результаты обработки экспериментальных данных:
а – ЗР Вейбулла; б – экспоненциальный ЗР

Таблица 1

Результаты испытаний влияния давления на секции
высоковольтных импульсных конденсаторов при их запрессовке

N	1	2	3	4	5	6	7	8	9	10
T	10050	10226	10403	10760	10819	10899	11000	11020	11117	11218
N	11	12	13	14	15	16	17	18	19	20
T	11264	11333	11863	11908	12586	13435	13544	13832	13964	14100
N	21	22	23	24	25	26	27	28		
T	14499	14710	15373	15920	16129	16129	17890	25063		

Выводы. 1. Уточнена методика идентификация ЗР графоаналитическим методом путем построения теоретической зависимости ЗР на вероятностной бумаге с помощью метода наименьших квадратов.

2. Разработан компьютерный вариант обработки экспериментальных данных с одновременным сравнением и выбором лучшего из пяти законов распределения.

3. Приведен пример обработки экспериментальных данных ресурсных испытаний секций высоковольтных импульсных конденсаторов.

ЛИТЕРАТУРА

1. Сотсков Б.С. *Основы теории и расчета надежности элементов и устройств автоматики и вычислительной техники.* – М.: Наука, 1990. – 270 с.
2. Корн Г., Корн Т. *Справочник по математике.* – М.: Наука, 1970. – 722 с.
3. Рунион Р. *Справочник по непараметрической статистике. Современный подход.* – М.: Финансы и статистика, 1982. – 198 с.
4. Кремер Н.Ш. *Теория вероятностей и математическая статистика: Учебник для ВУЗов.* – М.: ЮНИТИ-ДАНА, 2002. – 543 с.
5. Rudakov V.V., Albova I.M., Gorelov V.S. *The term of service life of high voltage pulsed capacitors depending on the embedding degree.* // *Prog. 9-th Intern. Symp. H.V.E. – GRAZ (Austria).* – 1995. – sub.1. – P. 1904 (1 – 4).

Поступила 28.08.2003

РУДАКОВ Сергей Валерьевич, канд. техн. наук, преподаватель кафедры ХВУ. В 1996 году окончил ХВУ. Область научных интересов – моделирование случайных процессов.

ДУБИЙЧУК Олег Юрьевич, аспирант НТУ «ХПИ». В 2002 году окончил НТУ «ХПИ». Область научных интересов – надежность высоковольтной изоляции.
