УДК 621.019

Э. М. Фархадзаде, д-р техн. наук,
А. З. Мурадалиев, канд. техн. наук,
Ю. З. Фарзалиев, аспирант
Азербайджанский научно-исследовательский
и проектно-изыскательский ин-т энергетики
(Азербайджанская Республика, Az 1012 Баку, пр-т Зардаби 94,
тел.: (99412) 4316407, e-mail: fem1939@rambler.ru)

Повышение точности оценки показателей индивидуальной надежности энергоблоков

(Статью представил д-р техн. наук В. Г. Тоценко)

Оценка показателей индивидуальной надежности при механической классификации статистических данных связана с возможной существенной погрешностью по сравнению с действительными значениями. Предложен новый критерий контроля представительности выборки, позволяющий оценить значимость признаков при классификации данных.

Оцінка показників індивідуальної надійності при механічній класифікації статистичних даних пов'язана з можливим суттєвим відхиленням цих оцінок від дійсних значень. Запропоновано новий критерій контролю представництва вибірки, який дозволяє оцінити значимість ознак класифікації даних.

Ключевые слова: надежность, моделирования, индивидуальность, точность.

Одной из основных задач, возникающих при оценке надежности энергоблоков (ЭБ), является обоснование точности показателей индивидуальной надежности (ПИН).

Необходимость решения такой задачи возникает:

1) при классификации априорной информации о состояниях ЭБ по заданным разновидностям признаков, характеризующих ЭБ; например, требуется оценить ПИН 4ЭБ (признак — номер ЭБ, разновидность — 4) в 2005 г. (признак — год, разновидность — 2005 г. и т.д.); не все заданные признаки могут быть значимыми [1];

2) при сравнении ПИН двух и более ЭБ в задачах планирования сроков и объема предстоящих плановых ремонтов по техническому состоянию, оценке качества ремонта;

3) при прогнозировании ПИН;

4) при моделировании ПИН — отдельные ПИН представляют собой функцию ряда случайных величин (например, коэффициент готовности) и

потому их точность может быть оценена только в результате моделирования состояний на ЭВМ [2].

Энергоблок относится к категории сложных объектов. Одна из основных особенностей сложных объектов — их многомерность, т.е. для оценки ПИН приходится учитывать большое число признаков различной физической природы. Причем знаний о значимости этих признаков очень мало. При выборе разновидностей признаков возможны дублирующие и «шумящие» признаки. Объективное уменьшение числа признаков и их разновидностей позволяет сократить потери информации, ограничить интервал изменения оценок ПИН и тем самым повысить их точность.

При оценке ПИН ЭБ по статистическим данным диспетчерских журналов могут быть определены длительности состояний однотипных ЭБ за ряд лет. Собранная информация включает статистические данные о тысячах состояний ЭБ и позволяет оценить усредненные показатели надежности [3]. При исследованиях ПИН возникают вопросы о возможности совместного использования этих данных. Если для однотипных элементов, работающих в сходных условиях, такие вопросы не возникают, то с возрастанием сложности исследуемого объекта ответы на эти вопросы далеко не однозначны.

Энергоблок представляет собой ряд последовательно соединенных установок (парокотельная, паротурбинная, блочный турбогенератор, силовой трансформатор, выключатель), каждая из которых состоит из большого числа единиц оборудования, устройств, специализированных систем (маслоснабжения, релейной защиты и автоматики, охлаждения и др.). В свою очередь, можно выделить узлы оборудования, элементы и т.д. Оценки ПИН каждого из них различны. С увеличением длительности наблюдения за техническим состоянием ЭБ изменяются и характеристики распределения длительности состояний.

Изложенное свидетельствует о том, что статистические данные о длительности состояний по всем ЭБ не могут рассматриваться как генеральная совокупность. Эта трудно формализуемая, изменяющая свои статистические характеристики во времени, совокупность статистических данных может быть представлена лишь как некоторая совокупность данных, а статистические данные о конкретном ЭБ — как выборка из этой совокупности.

Требуется ответить на вопрос: насколько справедливо предположение о представительности выборки. В математической статистике подобные предположения именуются гипотезами. Теория проверки статистических гипотез включает ряд методов [4].

Обращаясь к данным о длительности состояний ЭБ, нетрудно заметить, что для проверки гипотезы о представительности выборки могут быть использованы непараметрические критерии, в частности критерий

Смирнова. Для этого необходимо сопоставить функцию распределения анализируемой выборки $F_m^*(x)$ и функцию распределения общей совокупности данных M за вычетом реализаций выборки $F_n^*(x)$, где $n = M - m$. Если в результате расчетов окажется, что распределения $F_m^*(x)$ и $F_n^*(x)$ различаются неслучайно, то можно утверждать, что выборка непредставительна. Применение непараметрических критериев позволяет сделать косвенное заключение о представительности выборки. Иначе говоря, предполагается, что если функции распределения $F_m^*(x)$ и $F_n^*(x)$ различаются неслучайно, то неслучайно различаются и распределения $F_m(x)$ и $F_M(x)$.

Рассмотрим разработанный метод непосредственной проверки предположения о представительности выборки.

Метод проверки представительности выборки. Пусть задана некоторая совокупность значений случайной величины, например длительность простоя в аварийном ремонте ЭБ. Поскольку в дальнейшем изложении будут использованы случайные величины, характеризующие количественный результат различных измерений или вычислений, символическое обозначение которых различно, в качестве упрощения условимся обозначать их буквами X, Y, Z, а их возможные реализации — соответственно буквами x, y, z. Обозначим данную совокупность через $\{x\}_M$, где M — число реализаций случайной величины x. Численное значение x зависит от r признаков, каждый из которых характеризуется r_i разновидностями, где i — порядковый номер признака, $1 \le i \le r$; обозначив через j порядковый номер разновидности признака, запишем $1 \le j \le r_i$. Например, длительность аварийного простоя ЭБ зависит от типа поврежденной установки, от характера отключения (внезапное или по аварийной заявке), от вида повреждения и т.д.

Из $\{x\}_M$ проведена выборка из m значений x. Например, выделены сведения о длительности восстановления внезапных отказов третьего ЭБ за 2005 г. Обозначим выборку значений x через $\{x\}_m$. Требуется оценить представительность выборки $\{x\}_m$.

Выборка считается представительной, если она выбрана случайно. Для решения этой задачи воспользуется леммой [1], в соответствии с которой для принятых выше обозначений «если случайная величина x имеет непрерывную функцию распределения $F(x) = P(X < x) = y$, то случайная величина Y имеет равномерное распределение $F(y)$ в интервале [0,1]». Равномерное распределение соответствует равенству $y = F(y)$, которое справедливо, так как если $F(x) = y$, то

$$y = P(X < x) = P(Y < y) = F(y).$$

Использование этой леммы было предложено А.Н. Колмогоровым и нашло отражение в оценке статистик $D_m(a)$:

$$D_m = \max(D_m^+, D_m^-);$$

$$D_m^+ = \max\left[\frac{i}{m} - F(x_i)\right]_m;$$
$$D_m^- = \max\left[F(x_i) - \frac{i-1}{m}\right]_m. \qquad (1)$$

В [5] этот прием использован для проверки гипотезы о показательном виде вероятности безотказной работы по результатам периодических испытаний на надежность небольших групп элементов. Разработанный метод контроля представительности выборки (КПВ) также основан на этом приеме и сводится к определенной последовательности вычислений.

1. Совокупность $\{x\}_M$ размещается в порядке возрастания и каждому значению x вариационного ряда совокупности сопоставляется оценка вероятности:

$$F_n^*(x_i) = \frac{i}{M}\,(i = \overline{1, M}).$$

2. Аналогично п. 1 выборка $\{x\}_m$ размещается в порядке возрастания и каждому значению x вариационного ряда выборки сопоставляется оценка вероятности:

$$F_m^*(x_i) = \frac{i}{m}\,(i = \overline{1, m}).$$

3. Определяется расхождение:

$$\Delta_{m,i} = \left|F_m^*(x_i) - F_M^*(x_i)\right| = \left|\frac{i}{m} - \frac{S_i}{M}\right|, \qquad (2)$$

где $i = \overline{1, m}$; S_i — порядковый номер τ_i в вариационном ряду совокупности случайных величин $\{x\}_M$.

4. Находится статистика:

$$\Delta_m = \max\{\Delta_{m,i}\}_m. \qquad (3)$$

5. Если фактически найденное наибольшее расхождение Δ_m^* окажется не менее критического значения статистики $\Delta_m(\alpha)$, то гипотеза о представительности выборки должна быть отвергнута, т.е. классификация статистических данных совокупности по заданному (заданным) признаку оправдана. В противном случае — классификация статистических данных совокупности приводит к неоправданному дроблению информации.

Экспериментальные исследования методом статистических испытаний на ЭВМ. Цель этих исследований — оценка распределений значений статистик Смирнова $D_{m,n}^+$, $D_{m,n}^-$ и $D_{m,n}$ и рекомендуемой статистики Δ_m, их сопоставление и оценка критических значений при $m \le M < 100$.

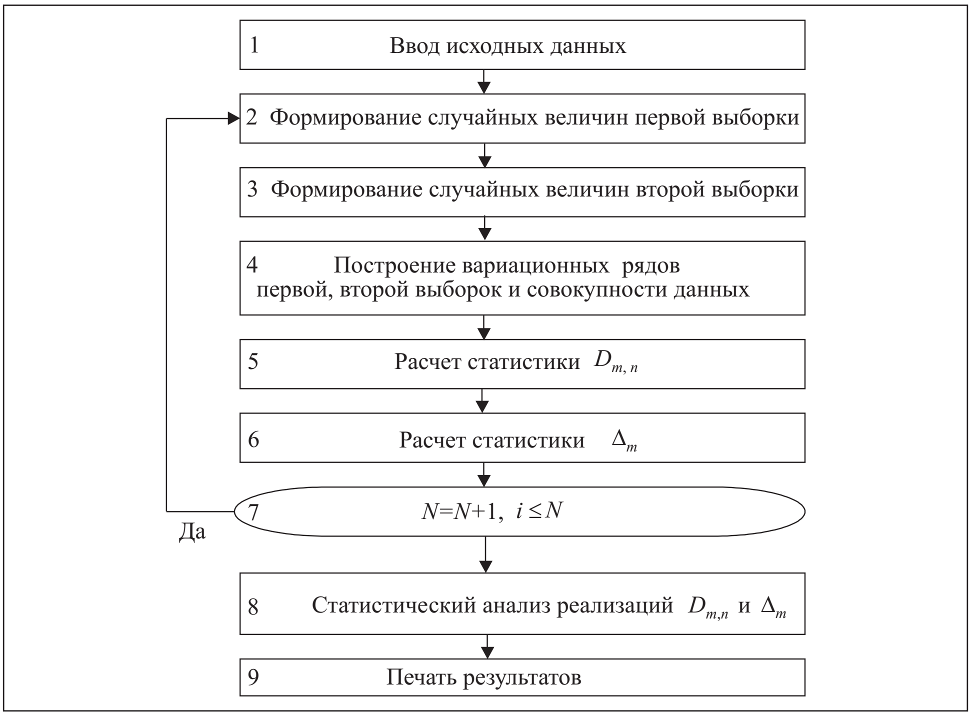

Рис. 1. Укрупненная блок-схема алгоритма оценки критических значений статистик $D_{m,n}$ и Δ_m

Блок-схема последовательности расчета критических значений статистик приведена на рис. 1.

Следует отметить главные различия статистик Смирнова и Δ_m: если статистики Смирнова определяются для различных независимых случайных выборок с числом случайных величин соответственно m и n, то статистика Δ_m оценивает наибольшее расхождение значений функций распределения $F_m^*(x)$ и $F_M^*(x)$ для одних и тех же значений $\{x_i\}_m$, определяющих выборку с числом случайных величин m.

На рис. 2 приведены графики расчета величины Δ_m, характеризующей представительность выборки, где задана совокупность значений случайной величины X, соответствующих распределению $F(x)$, из восьми данных $\{x\}_8$. Случайным образом из $\{x\}_8$ выбраны четыре ($m = 4$) значения X: $x_{2,1}$, $x_{2,2}$, $x_{2,3}$ и $x_{2,4}$. Каждому значению x_i $(i=\overline{1,8})$ соответствует определенное значение $y_i = F(x_i)$. Совокупность $\{y\}_8$ согласно упомянутой выше лемме соответствует равномерному закону $F_8(y)$, а $F_8^*(Y_i)=i/8$, где $i=\overline{1,8}$, также как и каждому j-му значению y_j $(j=\overline{1,4})$ выборки $\{y\}_4$ $F^*(Y_j)=j/4$. Реализация расхождений Δ определяется по единой формуле (2).

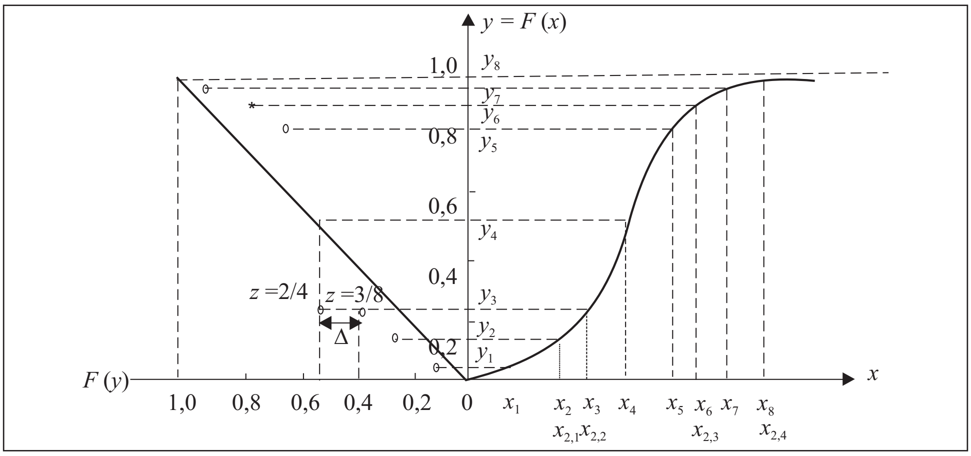

Рис. 2. Графики расчета величины Δ для гипотезы H_0

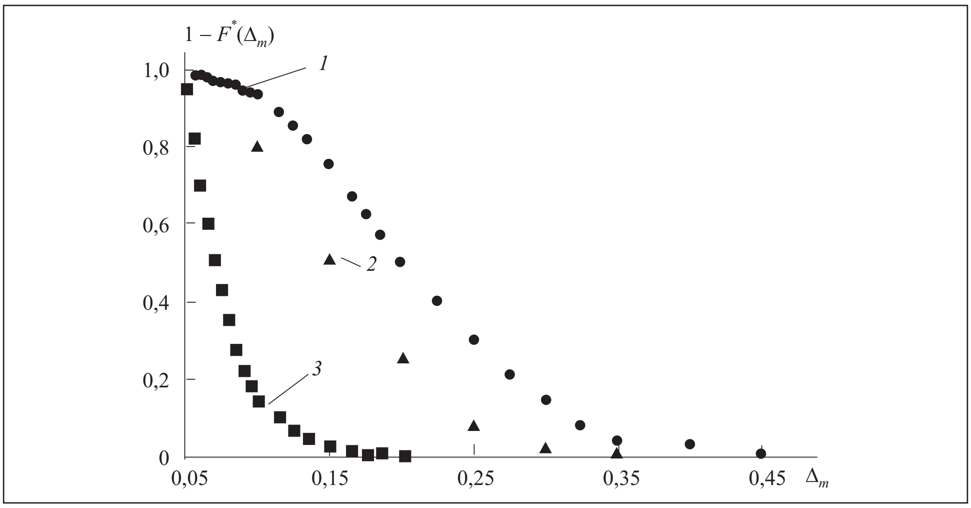

Рис. 3. Распределения статистики Δ_m в функции числа случайных величин выборки и совокупности данных: *1* — $m = 10$, $n = 200$; *2* — $m = 10$, $n = 20$; *3* — $m = 50$, $n = 200$

На рис. 3 приведены распределения статистики Δ_m в функции числа случайных величин выборки и числа совокупности данных. Как следует из рис. 3, с увеличением числа случайных величин представительной выборки m при фиксированном n, вероятность заданного отклонения распределения $\alpha(\Delta_m) = [1 - F_m^*(x)]$ существенно уменьшается, т.е. существенно уменьшается критическое значение статистики Δ_m. Изменение m оказывает более значительное влияние, чем изменение n.

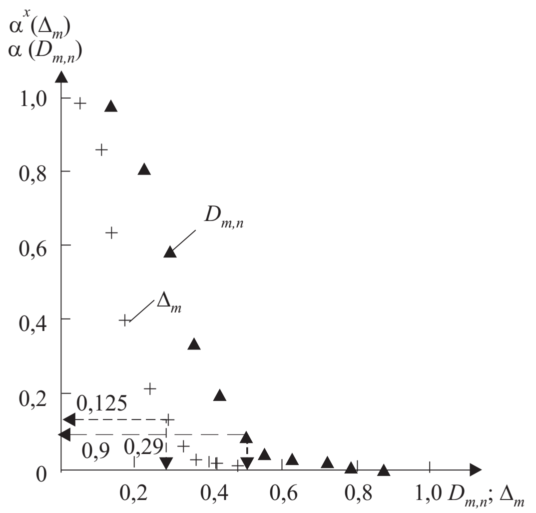

Рис. 4. Зависимость изменения расхождений распределений длительности простоя ЭБ в аварийном ремонте

Пример. Рассмотрим совокупность данных о длительности простоя в аварийном ремонте третьего ЭБ в 2005 г. и выборку этой совокупности, отражающей длительность отключения ЭБ по аварийной заявке. Вариационные ряды этих реальных случайных величин, а также результаты расчетов средних значений $x^*_{\text{ср}M}$, $x^*_{\text{ср}m}$, статистик Δ_m и Смирнова $D_{m,n}$ приведены в таблице. Следует заметить, что значения $M = 14$ и $m = 7$ выбраны преднамеренно, так как не искажая процесс КВП, они при существенном сокращении числа дискретизаций (в рассматриваемом примере для критериев Δ_m и $D_{m,n}$ они оказались равными 11), уменьшают громоздкость вычислений.

На рис. 4 приведены результаты моделирования распределений $[1-F^*(\Delta_m)]$ и $[1-F^*(D_{m,n})]$ для $M = 14$ и $n = 7$. С учетом дискретного характера этих распределений наиболее близкими при уровне значимости $\alpha = 0{,}1$ оказались вероятности 0,125 для статистики Δ_m и 0,092 для статистики Смирнова $D_{m,n}$. Соответствующие этим вероятностям критические значения следующие: Δ_m (0,125) = 0,29 и $D_{m,n}$ (0,092) = 0,50. Сопоставив эти значения с данными таблицы можно сделать вывод о том, что по критерию Δ_m выборка является представительной, так как $\Delta^*_m < \Delta_m(0{,}125)$, а по критерию Смирнова она непредставительна, поскольку $D^*_{m,n} > D_{m,n}(0{,}092)$. Результат сопоставления не изменится, если принять $\alpha^*(\Delta_m)$ и $\alpha\,(D_{m,n})$ равными примерно 0,2 или 0,05.

Этот пример свидетельствует о возможном различии решений. Согласно теории проверки статистических гипотез в таких случаях реко-

Последовательность расчета статистики Δ_m

$N(i)$	Исходные данные		$F^*_M(x_i)$	$F^*_M(x_i)$	$\|\Delta\|$	Результаты расчетов
	Совокупность $\{x_i\}_{14}$	Выборка $\{x_i\}_{14}$				
1	2,1	3,3	0,071	0,143	0,0	
2	3,3	3,6	0,142	0,285	0,00	$x^*_{срM} = 27$ ч
3	3,4	10,1	0,214	0,429	0,21	
4	3,6	22,4	0,285	0,571	0,21	$x^*_{срm} = 48$ ч
5	5,1	87,36	0,357	0,714	0,14	$\Delta^*_m = 0,21$
6	6,2	93,2	0,429	0,857	0,07	
7	6,4	113,47	0,500	1,000	0,00	
8	8,3		0,571			$M = 14$
9	10,1		0,643			
10	13,0		0,714			$n = 7$
11	22,4		0,785			$D^*_{m,n} = 0,57$
12	87,4		0,857			
13	93,2		0,929			
14	113,5		1,000			

мендуется ориентироваться на критерий, обеспечивающий наименьшую ошибку второго рода. Достаточно четкие методические рекомендации при этом отсутствуют.

Разработанные авторами методы такого сравнения на основе имитационного моделирования выходят за рамки настоящей статьи. Однако следует заметить, что критерий непосредственного КПВ Δ_m имеет значительно меньшую ошибку второго рода, чем критерий косвенного контроля $D_{m,n}$.

The estimation of parameters of individual reliability at mechanical classification of the statistical data is connected to a possible essential deviation (rejection) of this estimation from the valid meanings. The new criterion of the control of sampling representative is proposed, which allows to estimate the importance of attributes at classification of the data.

1. *Мостеллер Ф., Тьюки Дж.* Анализ данных и регрессия. — М. : Финансы и статистика, 1982. — 240 с.
2. *Фархадзаде Э. М., Мурадалиев А. З., Рафиева Т. К.* Имитационное моделирование состояний энергоблоков//Проблемы энергетики. — 2005. — № 4.
3. *Фархадзаде Э. М., Сафарова Т. Х., Мурадалиев А. З. и др.* Автоматизированная система анализа индивидуальной надежности и эффективности энергоблоков ГРЭС. //: Электрические станции. — 2005. — № 11.
4. *Рябинин И. А.* Основы теории и расчета надежности судовых электроэнергетических систем. 2-е изд. — Л. : Судостроение, 1971.
5. *Гнеденко Б. В., Беляев Ю. К., Соловьев А. Д.* Математические методы в теории надежности. — М. : Наука, 1965. — 524 с.

Поступила 06.09.06;
после доработки 30.01.07

ФАРХАДЗАДЕ Эльмар Мехти оглу, д-р техн. наук, профессор, руководитель лаборатории «Надежность энергетического оборудования» АзНИПИИ энергетики (Баку). В 1961 г. окончил Азербайджанский ин-т нефти и химии (АзИНЕФТЕХИМ). Область научных исследований — надежность и эффективность электроэнергетических систем.

МУРАДАЛИЕВ Айдын Зураб оглу, канд. техн. наук, ст. науч. сотр. лаборатории «Надежность энергетического оборудования» АзНИПИИ энергетики (Баку). В 1982 г. окончил Азербайджанский ин-т нефти и химии (АзИНЕФТЕХИМ). Область научных исследований — количественная оценка индивидуальной надежности оборудования и устройств электроэнергетических систем.

ФАРЗАЛИЕВ Юсиф Зейни оглу, аспирант лаборатории «Надежность энергетического оборудования» АзНИПИИ энергетики (Баку). В 1985 г. окончил Азербайджанский госуниверситет. Область научных исследований — точность и достоверность оценок показателей индивидуальной надежности.