

УДК 621.002

# ПРЕДСТАВЛЕНИЕ И ОБРАБОТКА СТАТИСТИЧЕСКИХ ДАННЫХ, НЕ ПОДЧИНЯЮЩИХСЯ УНИМОДАЛЬНЫМ ЗАКОНАМ РАСПРЕДЕЛЕНИЯ

**А.В. Горошко, В.П. Ройзман**

*Показана необходимость умения работы со статистическими данными, не подчиняющимися унимодальным законам распределения вероятностей. Предложен итерационный метод построения гистограмм, отличающийся обоснованным выбором шага построения гистограммы. Разработан параметрический метод построения эмпирических законов распределения с учетом их многомодальности, основанный на представлении смеси распределений суперпозицией унимодальных законов распределения. Предложен метод определения обоснованных допустимых значений исследуемых параметров, в общем случае не подчиняющихся унимодальным законам распределения вероятностей.*

**Ключевые слова:** закон распределения, многомодальное распределение, дисбаланс, допуски.

## **Введение**

Наиболее общей задачей математической статистики является выбор статистической модели распределения исследуемых признаков, содержащей оценку неизвестных законов распределения и их параметров, проверку статистических гипотез и т.д. Целью построения статистической модели является представление данных наблюдений путем подбора аппроксимирующего распределения.

Исторически сложилось так, что нормальное распределение считается практически всеобъемлющей статистической моделью из-за достаточно общих условий, поскольку результат измерения (наблюдения) складывается под действием многих причин, причем каждая из них вносит лишь малый вклад, а совокупный итог определяется аддитивно [1]. Поэтому подавляющее большинство статистических критериев, методов и оценок разработаны именно для нормального распределения.

Между тем такое положение вещей не всегда соответствует действительности. Так, например, в работах [2, 3] указывается, что

нормальный закон распределения погрешностей на самом деле может быть получен только при выполнении множества различных условий: если в выборке представлена одна партия изделий; нет доминирующих причин возникновения погрешностей; количество случайных факторов, которые обуславливают возникновение погрешностей, неизменно во времени; все случайные факторы являются взаимонезависимыми и действуют аддитивно и т.д.

Проведенное авторами изучение гистограмм, построенных по результатам измерений дисбалансов достаточно большого количества однотипных роторов компрессоров авиадвигателей АИ-20 после изготовления или после эксплуатации, показало, что *плотность вероятности* (ПВ) дисбалансов хорошо аппроксимируется одновершинной кривой, если изготовление и сборка роторов происходили в идентичных условиях, а также наблюдались одинаковые условия эксплуатации. Дисбалансы же однотипных роторов, вызванные, например, различными производственными или

эксплуатационными причинами, имеют четко выраженные многовершинные гистограммы. Примеры таких гистограмм относительных частот (нормализованных гистограмм), являющихся эмпирическими оценками плотности распределения, представлены на рис. 1 и рис. 2 [4] по результатам исследования выборки из 252 роторов. По оси абсцисс отложены значения дисбалансов с шагом (интервалом)  $h = 4 \text{ г}\cdot\text{см}$ , а по оси ординат – плотности относительных частот  $n_i/(nh)$ , где  $n$  – объем выборки;  $n_i$  – частота, т. е. число наблюдений значений дисбаланса в заданном интервале ( $\sum n_i = n$ ),  $n_i/n$  – относительная частота. Площадь  $i$ -го прямоугольника численно равна относительной частоте  $n_i/n$  в  $i$ -м интервале. Следовательно, площадь гистограммы относительных частот равна сумме всех относительных частот, т. е. единице.

Физическая сущность многомодального распределения объясняется, по-видимому, наличием различных типов доминирующих причин, вызывающих появление дисбалансов. Так, одни авиадвигатели из изучаемой партии работали на самолетах в условиях Крайнего Севера, другие – в условиях полевых, плохо оборудованных аэродромов Юга, третьи – в условиях больших перепадов температур при перелетах из Северного полушария в Южное, четвертые – в морских, коррозионных условиях и т.п. Поэтому появление эксплуатационных дисбалансов первого типа вызвано в основном попаданием в двигатель мелких частиц льда и снега, второго типа – попаданием мелких камней и других твердых частиц, третий тип дисбаланса вызван, как правило, температурными процессами, четвертый – коррозионными и т.п.

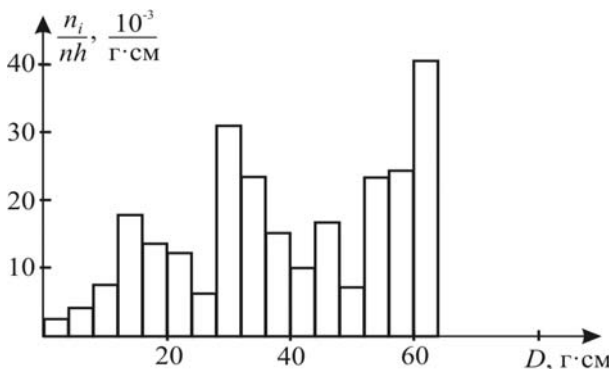


Рис. 1. Зависимость распределения дисбалансов  $D$  ротора компрессора от плотности относительных частот при низкочастотной балансировке для его левой опоры

Таким образом, однородная в исходном состоянии выборка роторов в процессе эксплуатации распадается на несколько подвыборок, каждая из которых объединена типом доминирующей причины, вызывающей появление эксплуатационного дисбаланса. Попадая на завод, например для межресурсного ремонта, эти роторы образуют партии с многомодальными законами распределения дисбалансов.

Такая же картина может иметь место и для любых других роторов, например, электрических машин, компрессоров, центрифуг и т. д., эксплуатируемых в различных условиях. Рост эксплуатационного дисбаланса определяется этими конкретными условиями и, следовательно, результирующий дисбаланс не будет описываться одновершинной гистограммой.

Аналогичная картина наблюдается для начальных дисбалансов роторов в случае изготовления или балансировки деталей на нескольких различных станках или поточных линиях, имеющих свои преобладающие погрешности.

### Постановка задачи

Поскольку далеко не всегда причины, вызывающие разброс значений исследуемого параметра, очевидны, возникает проблема поиска методов обработки многомодальных эмпирических законов распределения.

Математически многомодальные законы распределения, которые называют смесями функций распределения, можно описать следующим образом [4, 5]. Пусть в приведенном ранее примере распределения дисбалансов роторов компрессоров авиадвигателей количество условий эксплуатации, формирующих доминирующие причины появления подвыборок,

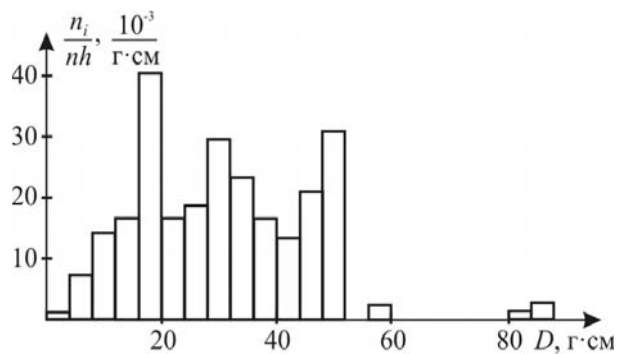


Рис. 2. Зависимость распределения дисбалансов  $D$  ротора компрессора от плотности относительных частот при низкочастотной балансировке для его правой опоры

равняется  $N$ , вероятность того, что авиадвигатель эксплуатировался в  $i$ -м эксплуатационном условии, равняется  $\rho_i$  ( $i = 1, 2, \dots, N$ ;  $\sum_{i=1}^N \rho_i = 1$ ,

$0 \leq \rho_i \leq 1$ ), дисбаланс  $X$  авиадвигателя есть случайная величина с ПВ  $f_i(x, M_i, S_i)$ , где  $M_i$  и  $S_i$  – математическое ожидание и среднее квадратическое отклонение  $i$ -й подвыборки. Полная («усредненная») функция распределения  $F(x)$  случайной величины  $X$  с учетом случайности ее закона распределения определяется выражением

$$F(x) = \sum_{i=1}^N \rho_i F_i(x), \quad (1)$$

где функции распределения  $F_i(x)$  называют компонентами смеси, а  $\rho_i$  – весами соответствующих компонент.

Дискретной смеси распределения  $F(x)$  соответствует дискретная ПВ

$$f(x) = \sum_{i=1}^N \rho_i f_i(x, M_i, S_i), \quad x \in R. \quad (2)$$

В опубликованных работах (см., например, [6]) представлены различные многомодальные распределения и решается задача определения вида и параметров результирующей кривой по заданным видам и параметрам составляющих законов распределения. Однако при построении статистических моделей распределения более важной является обратная задача – разделения (расщепления) смесей, поскольку структура полученных при выборочных наблюдениях данных, как правило, неизвестна. Это задача определения количества, доли и параметров каждой из подвыборок в общей выборке (смешанном распределении). Прделанный авторами работы [6] обзор известных подходов к решению таких задач показал, что, несмотря на высокую теоретические ценность, при решении практических задач они трудно применимы, поскольку смеси ПВ, как правило, задаются не графиками, а некоторыми конечными выборками реализаций случайной величины, которые представляются в удобном для обработки виде, например в виде гистограммы.

Сама по себе гистограмма дает возможность прогнозировать величину исследуемого параметра только на ограниченном интервале. Для построения статистической модели распределения и обоснованного прогнозирования

в области малых вероятностей необходимо аппроксимировать ее некоторой аналитической функцией. Таким образом, выбор статистической модели распределения определяется видом гистограммы, который, в свою очередь, зависит от способа ее построения и в особенности от шага интервала значений.

Рекомендации по выбору шага разбиения интервала значений исследуемой случайной величины, которые есть в литературе по теории вероятностей и математической статистике, носят чисто эмпирический характер, (например, правило Штюргеса) [3].

Проведенный авторами анализ показал, что существующие подходы и методы решения задачи обработки статистических данных, подчиняющихся многомодальным законам распределения, имеют существенные недостатки. Это ограничивает их применение для обработки эмпирических данных, в связи с чем требуется разработка более эффективных методов. Наряду с решением проблемы разделения смесей распределений случайных величин требуют решения и другие задачи, в частности обеспечение устойчивости решений и создание методов построения гистограмм.

### Обсуждение предложенного метода обработки статистических данных

Если вероятность того, что случайная величина  $X$  принимает значения, принадлежащие множеству  $x_i$  ( $i$ -й подвыборке), равняется  $\rho_i$ , то для последующей обработки статистических данных предлагается ПВ величины  $X$  представлять линейной комбинацией вида (2), в которой ПВ  $f_i$  – одномодальные (например, гауссовские), а  $\rho_i \geq 0$  ( $i = 1, 2, \dots, N$ ) и  $\sum_{i=1}^N \rho_i = 1$ . Следовательно, гистограмма аппроксимируется линейной комбинацией гауссовых функций ПВ (с весовыми коэффициентами  $\rho_i$ ) вида

$$f(x, M_1, M_2, \dots, M_N; S_1, S_2, \dots, S_N; \rho_1, \rho_2, \dots, \rho_N) = \sum_{i=1}^N \frac{\rho_i}{S_i \sqrt{2\pi}} \exp\left(-\frac{(x - M_i)^2}{2S_i^2}\right). \quad (3)$$

Для дальнейшей обработки результатов эксперимента прежде всего необходимо определить неизвестные параметры, применив, например, интерполяцию на некотором точечном множестве. Согласно такой интерполяции неизвестные параметры находят из условия

совпадения значений функции (3) в отдельных точках (например, вершинах и впадинах) со значениями аппроксимируемой функции, график которой в виде плавной кривой огибает построенную гистограмму. Понятно, что для однозначного определения  $3N$  неизвестных параметров ( $N$  – количество мод, для каждой моды неизвестными являются параметры  $M_i$ ,  $S_i$  и  $\rho_i$ ) количество точек в множестве должно быть не меньше, чем  $3N - 1$  (поскольку коэффициенты  $\rho_i$  всегда связаны уравнением  $\sum_{i=1}^N \rho_i = 1$ ).

Таким образом, для нахождения неизвестных  $M_i$ ,  $S_i$  и  $\rho_i$  необходимо составить и решить систему уравнений вида

$$\left\{ \begin{aligned} F(x)_j &= \int_{-\infty}^{\infty} \left( x_j, M_1, M_2, \dots, M_N; \right. \\ &\left. S_1, S_2, \dots, S_N; \rho_1, \rho_2, \dots, \rho_N \right) = \\ &= \sum_{i=1}^N \frac{\rho_i}{S_i \sqrt{2\pi}} \int_{-\infty}^x \exp \left( -\frac{(x_j - M_i)^2}{2S_i^2} \right) dx; \quad (4) \\ j &= 1, 2, \dots, 3N - 1; \\ \sum_{i=1}^N \rho_i &= 1, \end{aligned} \right.$$

где  $M_i$ ,  $S_i$ ,  $\rho_i$  – постоянные, но неизвестные параметры распределения  $i$ -й подвыборки и ее весовой коэффициент.

Очевидно, что решение системы (4) после подстановки в выражение (3) тем точнее приближается к реальным значениям неизвестных параметров, чем меньше участки разбиения при построении гистограммы, т. е. чем точнее гистограмма и огибающая ее плавная кривая.

Поиск описанных параметров можно осуществить и способом наименьших квадратов [7]. Для этого необходимо записать следующую функцию, равную сумме квадратов отклонений:

$$V = \sum_{j=1}^n \left[ g(x_j) - \sum_{i=1}^N \frac{\rho_i}{S_i \sqrt{2\pi}} \left\{ \exp \left( -\frac{(x_j - M_i)^2}{2S_i^2} \right) \right\} \right]^2; \quad (5)$$

$n = 1, 2, \dots; n \geq 3N - 1,$

где  $g(x_j)$  – значения огибающей функции в выбранных точках.

Затем необходимо приравнять к нулю частные производные функции (5) по каждому из параметров.

При решении аналогичной задачи можно использовать и метод моментов, но поскольку

ку подсчет эмпирических моментов высоких порядков приводит к значительным погрешностям, то такой способ предлагается применять для предварительных оценок искомых величин. Уточнение этих оценок следует проводить, максимизируя функцию максимального правдоподобия [7]:

$$W = \prod_{j=1}^{3N-1} \sum_{i=1}^N \frac{\rho_i}{S_i \sqrt{2\pi}} \exp \left( -\frac{(x_j - M_i)^2}{2S_i^2} \right), \quad (6)$$

т. е. приравнивая к нулю ее частные производные по искомым параметрам. В любом случае задача сводится к решению системы трансцендентных уравнений, которое должно выполняться на компьютере.

При обработке результатов эксперимента возникают определенные трудности. Прежде всего перед исследователями встает вопрос: каким должен быть шаг разбиения при построении гистограмм? Для выбора оптимального шага авторами предложены следующие рекомендации. Шаг должен быть минимальным, но не меньше, чем точность измерения параметра, а число  $N$  в выражении (2) должно равняться количеству полученных при построении вершин. Одним из предложенных ранее методов необходимо определить неизвестные параметры  $M_i$ ,  $S_i$  и  $\rho_i$ . Если в результате расчетов один или несколько весовых коэффициентов  $\rho_i$  окажутся меньше некоторой наперед заданной величины  $\beta$ , то соответствующими членами в линейной комбинации (3) можно пренебречь. Действительно, интегральная функция распределения с ПВ (3) имеет вид

$$F(x) = \sum_{i=1}^N \frac{\rho_i}{S_i \sqrt{2\pi}} \int_{-\infty}^x \exp \left( -\frac{(x - M_i)^2}{2S_i^2} \right) dx. \quad (7)$$

Пусть, например,  $\rho_1 < \beta$ . Тогда после отбрасывания первого слагаемого в линейной комбинации (3) новую интегральную функцию распределения можно записать в виде

$$\bar{F}(x) = \sum_{i=2}^N \frac{\rho_i}{S_i \sqrt{2\pi}} \int_{-\infty}^x \exp \left( -\frac{(x - M_i)^2}{2S_i^2} \right) dx.$$

Оценим разницу

$$\begin{aligned} |F(x) - \bar{F}(x)| &= \frac{\rho_1}{S_1 \sqrt{2\pi}} \int_{-\infty}^x \exp \left( -\frac{(x - M_1)^2}{2S_1^2} \right) dx \leq \\ &\leq \frac{\rho_1}{S_1 \sqrt{2\pi}} \int_{-\infty}^{\infty} \exp \left( -\frac{(x - M_1)^2}{2S_1^2} \right) dx \leq \beta, \end{aligned}$$

причем полученная оценка верна для любого  $x$ .

Например, если функция распределения вероятностей должна измеряться с точностью до 0,01, то значение  $\beta$  достаточно взять 0,005. Далее шаг гистограммы можно увеличивать до тех пор, пока число вершин не станет равным числу членов в линейной комбинации (3) после отбрасывания ее малых членов. Снова применяя тот же метод решения, но уже для меньшего числа неизвестных, можно определить их уточненные значения и отбросить малые члены. Такой процесс следует продолжать до тех пор, пока все  $\rho_i$  не станут сравнимы с выбранной точностью  $\beta$ . Полученный при этом шаг может быть принят в качестве оптимального.

Физически этот процесс означает, что подвыборки с малым  $\rho_i$  вносят весьма незначительный вклад в общую выборку и поэтому их объединяют с одной из подвыборок изделий с близкими величинами изучаемого параметра.

Следующая трудность, возникающая при решении подобных задач, носит чисто вычислительный характер. Дело в том, что сходимость решения системы трансцендентных уравнений на компьютере любым итерационным методом требует знания достаточно «хороших» начальных приближений, а они в большинстве случаев неизвестны. Применение метода моментов само по себе связано с вычислительными трудностями. Поэтому как один из путей решения такой задачи может быть предложен следующий способ. На начальном этапе необходимо применять градиентный метод [7], который при относительно невысокой точности может использоваться при «грубых» начальных приближениях. Полученное этим методом решение принимается за начальное приближение для применения более точного метода, например метода Ньютона [7].

Нахождение закона распределения вероятности исследуемого параметра в виде линейной комбинации (3) позволяет перейти к решению одной из важных практических задач – назначению допустимого значения этого параметра с определенной надежностью. Как известно, допустимое значение параметра, который характеризует свойства или качество работы новых изделий и материалов, не имеющих изученных аналогов, устанавливается путем испытаний одной партии. При этом для изделий (материалов) создаются критические, наиболее

неблагоприятные ситуации, при которых эти изделия (материалы) еще способны выполнять возложенные на них функции, и определяют значение исследуемого параметра.

Как известно, рассеяние значений исследуемого параметра зависит от принятого способа изготовления изделия. Границы интервалов рассеивания определяются законами распределения параметра, который рассматривается как случайная величина, представляющая собой сумму случайных величин, каждая из которых вызывается одним из непреодолимых факторов. Если число слагаемых в сумме достаточно велико, то может возникнуть два варианта при назначении функции распределения параметра.

В случае, когда значение каждой из составляющих в описанной ранее сумме мало по сравнению со значением суммы, по центральной предельной теореме [3, 5] распределение суммы близко к нормальному. Физически это условие малости каждого слагаемого означает, что ни один из факторов, обусловивших появление соответствующей случайной величины, не имеет преобладающего значения.

Если же среди указанных факторов появляются один или несколько доминирующих, то соответствующие слагаемые имеют преобладающее значение в сумме, и закон распределения суммы становится многомодальным.

В случае нормального закона распределения параметра его допустимое значение устанавливается на основе полученных ранее реальных значений из следующих представлений.

Точечной оценкой для математического ожидания в силу закона больших чисел является выборочное среднее арифметическое  $x_{cp}$ . Нижняя и верхняя доверительные границы для математического ожидания имеют вид  $x_{cp} - t_\gamma S/\sqrt{n}$  и  $x_{cp} + t_\gamma S/\sqrt{n}$  соответственно. Здесь  $x_{cp}$  – выборочное среднее арифметическое значение;  $t_\gamma$  – квантиль распределения Стьюдента, взятый из таблицы для заданной доверительной вероятности  $\beta$  и числа степеней свободы

$n - 1$ ;  $S = \sqrt{(n-1)^{-1} \sum_{j=1}^n (x_{cp} - x_j)^2}$  – выборочное среднее квадратическое отклонение;  $n$  – объем выборки.

Тогда с доверительной вероятностью  $\beta$  можно утверждать, что допустимое значение параметра находится в пределах

$$x_{cp} - t_\gamma S / \sqrt{n} \leq x \leq x_{cp} + t_\gamma S / \sqrt{n}. \quad (8)$$

Отметим, что, если результаты наблюдений не подчиняются нормальному закону распределения вероятностей, следует применять непараметрический подход, согласно которому в интервале (8) квантиль распределения Стьюдента заменяют на число  $U(p)$ , заданное равенством  $\Phi(U(p)) = (1+p)/2$ , где  $\Phi(x)$  – функция стандартного нормального распределения с математическим ожиданием 0 и дисперсией 1. При этом полученные доверительные интервалы будут несколько уже [12], и возможная ошибка пойдет в запас точности.

За допустимое значение параметра целесообразно взять значение на одном из концов доверительного интервала (8), например

$$[X] = x_{cp} - t_\gamma S / \sqrt{n}. \quad (9)$$

Если полученная гистограмма описывается многомодальным законом распределения, дальнейшие действия по назначению допустимого значения исследуемого параметра могут производиться одним из двух способов.

1. Рассматривается подвыборка с минимальным (максимальным) значением  $M_i$ . Очевидно, что характеристика этой подгруппы минимальна (максимальна). Следовательно, определенную характеристику для таких изделий возможно принять и для всей партии, так как получаемые при этом погрешности пойдут в запас точности. В этом случае дальнейшая обработка экспериментальных данных может происходить только для указанной нормально распределенной подвыборки значений с параметрами распределения  $M_i, S_i$ , как описано выше.

Если есть возможность разделить исходную выборку изделий на подвыборки, объединенные одной из доминирующих причин появления разброса значений, то аналогичные операции по обработке экспериментальных данных следует проводить для каждой подвыборки.

2. Определенные параметры позволяют записать интегральную функцию распределения (7), которую, как и гауссову случайную величину, с помощью компьютера можно задать таблично следующим образом. Для каждого значения величины  $X$ , которая изменяется с определенным числовым шагом, например 0,1, по таблице функции распределения нормированного нормального распределения

$$\Phi^x(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{x^2}{2}\right) dx$$

определяют вероятности

$$\begin{aligned} \gamma_i &= \frac{1}{S_i \sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{(x - M_i)^2}{2S_i^2}\right) dx = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x - M_i)/S_i} \exp\left(-\frac{x^2}{2}\right) dx, \end{aligned}$$

$$i = 1, 2, \dots, N$$

и далее – значения интегральной функции

$$F^x(x) = \sum_{i=1}^N \rho_i \gamma_i. \quad (10)$$

Это означает, что функция  $F^x(x)$  задана таблично. Полученная таблица позволяет не только по значениям  $x$  определять функцию  $F^x(x)$ , но и наоборот, – по заданным значениям функции определять аргумент. Таким образом, для заданной доверительной вероятности можно найти искомое допустимое значение параметра  $[X]$  из соотношения вида

$$\begin{aligned} \gamma &= P\{X < [X]\} = F^x([X]) = \\ &= \sum_{i=1}^N \frac{\rho_i}{S_i \sqrt{2\pi}} \int_{-\infty}^{[X]} \exp\left(-\frac{(x - M_i)^2}{2S_i^2}\right) dx, \end{aligned}$$

составленного на основании определения интегральной функции распределения случайной величины параметра с ПВ (3).

Необходимо отметить, что, во-первых, второй способ более точный, поскольку он учитывает функции распределения всех подвыборок, а во-вторых, более универсальный, так как с его помощью можно решать поставленную задачу в случае произвольного распределения, если предварительно составить для него таблицу зависимости доверительной вероятности и аргумента интегральной функции распределения исследуемой величины.

Кроме того, второй способ назначения допустимого значения при многомодальном распределении параметра распространяется, как на частный случай, и на одномодальный закон. Более того, назначение допуска с помощью интегральной функции распределения в этом частном случае может служить даже уточнением способа решения аналогичной задачи при одномодальном законе распределения параметра, описанного ранее в допущении, что истинное значение измеряемой величины совпадает с ее математическим ожиданием.

Далее приводятся результаты проверки эффективности изложенного метода расше-

пления смесей распределений Гаусса для распределения, показанного на рис. 1. Для поиска параметров слагаемых законов распределения дисбалансов по левой опоре был применен способ интерполяции, при этом для сверки результатов взяты три разные выборки узлов интерполяции.

Расчеты производились с использованием пакета MATLAB. Поскольку начальные приближения для решения системы уравнений (4) неизвестны, применялся градиентный метод, относительно быстро сходящийся по достаточно «грубым» начальным приближениям, а затем полученные решения принимались за начальные приближения, и расчет продолжался методом Ньютона. Результаты, полученные для разных выборок интерполяции, показали несущественное различие. В таблице и на рис. 3 приведены результаты указанных расчетов.

В соответствии с интервалом (8) по вышеизложенной методике анализировалась третья мода смеси распределений, поскольку для нее характерны большие (а значит, более опасные) дисбалансы. В результате с доверительной вероятностью 0,95 было установлено, что допустимое значение функционального дисбаланса находится в пределах  $51,1 \leq D \leq 64,9$  (г·см). Следовательно, для всей исследуемой партии роторов компрессоров авиадвигателей было принято максимальное возможное значение дисбаланса 64,9 г·см с вероятностью его появления 0,95.

Содержащиеся в данной таблице параметры позволили рассчитать таблицы интегральной функции распределения. Их анализ показывает, что несмотря на определенное различие параметров составляющих законов, полученных по разным наборам узлов интерполяции, значения интегральной функции отличаются несущественно, что важно в способе назначения допусков, где применяется именно интегральная функция распределения.

На основе предложенного метода интегральных характеристик с помощью пакета MATLAB была записана интегральная функция (10), и с доверительной вероятностью 0,95 установлено, что максимальное допустимое значение функционального дисбаланса не превышает 61,6 г·см, а вероятность появления в выборке дисбаланса, не превышающего 64,9 г·см, равна 0,99. В данной ситуации целесообразно выбрать в качестве допустимого значение

дисбаланса 64,9 г·см, тогда возникающая погрешность пойдет в запас прочности изделия, поскольку изделия с большим допустимым значением, а именно 64,9 г·см, будут обладать большей прочностью.

Был также проделан следующий эксперимент. Один из весовых коэффициентов, который оказался существенно меньше двух других, полагали равным нулю, и интегральный закон распределения считали двухмодальным. Для этих случаев снова определяли соответствующие пять параметров значений  $M_i$ ,  $S_i$ ,  $\rho_i$  и вновь рассчитывали таблицы интегральной функции распределения. (Как указывалось ранее, количество неизвестных параметров составляет  $3N-1$ , и для трех весовых коэффициентов равно 8. Полагая меньший весовой коэффициент равным 0, будем иметь пять неизвестных параметров.)

Сравнительный анализ таблиц, рассчитанных вышеуказанным способом, для трех- и двухмодальных законов распределения показал, что они отличаются несущественно. Это открывает возможность при определенном анализе на компьютере автоматически находить не только практически достаточное количество мод закона распределения для любого шага  $h$ , выбранного при построении гистограмм, но и уточнять значения самого шага.

Таблица

Параметры смеси распределений

$M_1$	$M_2$	$M_3$	$S_1$	$S_2$	$S_3$	$\rho_1$	$\rho_2$
15,13	33,77	58,49	5,12	6,98	3,54	0,18	0,51

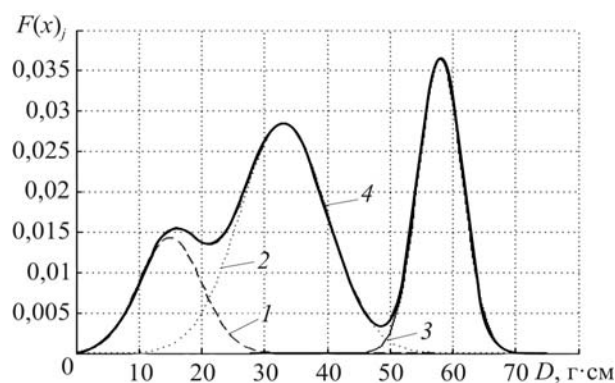


Рис. 3. Функции плотности распределения смеси (4) и ее искомых составляющих для различных значений дисбалансов:

1 – с параметрами  $M_1, S_1$ ; 2 – с параметрами  $M_2, S_2$ ; 3 – с параметрами  $M_3, S_3$

### Заключение

Предложенный метод статистической обработки данных экспериментов с измеренными параметрами технических объектов, позволяет, во-первых, раскрыть внутреннюю структуру данных с учетом возможной многомодальности закона их распределения, а, во-вторых, дает правила работы с такими статистическими данными, в частности, методы определения обоснованных допустимых значений исследуемых параметров.

Заметим, что исследуемый вопрос имеет и самостоятельное значение, поскольку статистические данные необходимо обрабатывать и при изучении производственных погрешностей изготовления объектов, и при идентификации технологических процессов, и при составлении нормативной документации, а также в целом ряде других важных случаев в практике проектирования и производства технических объектов [8].

### Список литературы

1. Плескунин В.И., Воронина Е.Д. Теоретические основы организации и анализа выборочных данных в эксперименте / Под. ред. А.В. Башарина. – Л.: Изд-во ЛГУ, 1979. – 232 с.
2. Расчет электрических допусков радиоэлектронной аппаратуры / Под ред. В.П. Гусева и А.В. Фомина. – М.: Сов. радио, 1963. – 367 с.
3. Орлов А.И. Прикладная статистика : учебник. – М.: Изд-во «Экзамен», 2004. – 656 с.
4. Горошко А.В., Ройзман В.П. Об одном параметрическом методе обработки экспериментальных данных // Современные достижения в науке и образовании : сб. тр. VIII Международной научно-методической конференции, 28 апреля – 5 мая 2013 г., Париж (Франция). Хмельницкий: ХНУ, 2013. С. 42–45.
5. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения : учеб. пособие для вузов. 2-е изд., стер. – М.: Высшая школа, 2000. – 480 с.
6. Горошко А.В., Ройзман В.П. Про задачу обробки статистичних матеріалів, що не підкоряються одномодальним законам розподілу // Современные достижения в науке и образовании : сб. трудов VII Международной научной конференции, посвящ. 50-летию ХНУ, 25 авг. – 1 сент. 2012 г., Опатия (Хорватия) Хмельницкий: ХНУ, 2012. с. 58–66.
7. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы : Анализ, алгебра, обыкновенные дифференциальные уравнения. – М.: Наука, 1973. – 631 с.
8. Royzman V., Goroshko A. Multiple inverse problem // Journal of vibroengineering. 2012, V. 14. P. 1417–1424.

Материал поступил в редакцию 30.04.13.

#### ГОРОШКО

Андрей

Владимирович

E-mail: iftomm@ukr.net

Тел.: (8038) 272-87-43

Кандидат технических наук, доцент кафедры физики и электротехники Хмельницкого национального университета (Украина). Сфера научных интересов – обратные задачи в механике, прикладная статистика, диагностика технического состояния. Автор 50 научных статей, трех изобретений.

#### РОЙЗМАН

Вилен Петрович

E-mail: royzman\_v@mail.ru

Тел.: (8038) 272-87-43

Доктор технических наук, профессор кафедры инженерной механики Хмельницкого национального университета (Украина). Сфера научных интересов – балансировка роторов, прочность в электронике, вибрационная прочность. Автор 485 научных публикаций, трех монографий, 20 изобретений.