УДК 681.518.54

**В.Ф. МИРГОРОД, Г.С. РАНЧЕНКО**

***ОАО «Элемент», Одесса, Украина***

# ВЕРОЯТНОСТНЫЕ ХАРАКТЕРИСТИКИ ТРЕНДОВОЙ СТАТИСТИКИ ХАЛЬДА-АББЕ ПРИ ОБРАБОТКЕ ПАРАМЕТРОВ РЕГИСТРАЦИИ ГТД

Выполнена оценка вероятностных характеристик критерия Хальда-Аббе для обнаружения тренда при статистической обработке данных в автоматизированных системах контроля и диагностики состояния ГТД.

**системы контроля и диагностики, трендовый анализ, техническое состояние, газотурбинный двигатель**

## Введение

Необходимость перехода к эксплуатации ГТД и его агрегатов по техническому состоянию требует решения общей проблемы повышения эффективности функционирования автоматизированных систем контроля и диагностики (АСКД), в частности, подсистем трендового контроля изменения основных параметров двигательной установки. Важной практической задачей при создании АСКД является обоснованный выбор и оценка характеристик трендовых статистик, в наиболее полной мере отвечающих статистическим свойствам анализируемых временных рядов, образованных данными регистрации в процессе эксплуатации ГТД.

**Постановка проблемы и цель исследования.** Известные в прикладной статистике [1 – 3] критерии наличия тренда в исследуемой выборке широко используются в проектируемых и реализованных АСКД. При контроле состояния ГТД наибольшее распространение получили трендовые статистики Хальда-Аббе, коммулятивных сумм и их модификации [4]. Представляется перспективным применение в АСКД Ф-критерия, основанного на статистике Фишера [5]. Наиболее широко используется трендовый критерий Хальда-Аббе, как для авиационных двигателей [6], так и для газоперекачивающих агрегатов [7, 8]. В литературе [8] приводятся убедительные примеры успешного применения данного критерия для предотвращения развития аварийных ситуаций. В то же время недостаточное внимание уделяется анализу соответствия статистической модели порождения данных, на которой базируется тот или иной трендовый критерий, реальным статистическим свойствам временных рядов, образованных параметрами регистрации ГТД. При таком возможном несоответствии установленные пороговые уровни для трендовых статистик не отвечают реализуемым уровням значимости, что сопровождается ложными срабатываниями трендовых статистик. Поскольку по результатам трендового анализа должны приниматься ответственные решения о продлении ресурса двигателя либо о возможности его дальнейшей эксплуатации, то существенную значимость приобретает оценка вероятностных характеристик трендовых статистик, в частности Хальда-Аббе, применительно к условиям обработки данных в АСКД. Целью настоящего исследования является оценка вероятностных характеристик одного из наиболее широко применяемых для обнаружения тренда критерия Хальда-Аббе при анализе временных рядов в АСКД.

## Основные результаты

Применительно к системам контроля и диагностики ГТД критерий Хальда-Аббе формируется в виде следующей статистики [4]:

$$r = \left[2\sum_{k=1}^{N}(y_k - \bar{y})^2\right]^{-1}\sum_{k=1}^{N}(y_{k+1} - y_k)^2 > r_0, \qquad (1)$$

где $y_k$ – исследуемый временной ряд; $r_0$ – пороговое значение критерия; $\bar{y} = N^{-1}\sum_{k=1}^{N} y_k$ – выборочное среднее.

Для данного критерия опорной гипотезой $H_0$ является предположение об отсутствии тренда в данных регистрации, при выполнении которой статистика (1) обычно [4] формируется в виде модифицированного критерия

$$r' = 0{,}5 \cdot 1n[(2-r)/r] > r'_0\ , \qquad (2)$$

и при $N > 10$ нормализуется с дисперсией $Dr' = 1/(N-3)$. Поэтому уровни принятия гипотезы $H_0$ легко устанавливаются по их значимости для обычно применяемых в АСКД $\alpha = 0{,}01;\ 0{,}05.$ В частности, для применяемых [6] значений $N = 20$ эти уровни составляют: $r'_{0.01} = 0{,}6306;\ r'_{0.05} = 0{,}4729.$

Выполним детальный анализ функционала решающей статистики (1), преобразуя его к эквивалентному виду

$$r = \left[2\sum_{k=1}^{N}(y_k - \bar{y})^2\right]^{-1}\sum_{k=1}^{N}\left[(y_{k+1} - \bar{y}) - (y_k - \bar{y})\right]^2 =$$

$$= \left[2\sum_{k=1}^{N}(y_k - \bar{y})^2\right]^{-1} \times$$

$$\times \left[\begin{matrix}\sum_{k=1}^{N}(y_{k+1} - \bar{y})^2 + \sum_{k=1}^{N}(y_k - \bar{y})^2 - \\ 2\sum_{k=1}^{N}(y_{k+1} - \bar{y})(y_k - \bar{y})\end{matrix}\right] = \qquad (3)$$

$$= 1 - \left[\sum_{k=1}^{N}(y_k - \bar{y})^2\right]^{-1} \cdot \sum_{k=1}^{N}(y_{k+1} - \bar{y})(y_k - \bar{y}) +$$

$$+ \left[2\sum_{k=1}^{N}(y_k - \bar{y})^2\right]^{-1} \cdot \left(y_{N+1}^2 - y_1^2\right) =$$

$$= 1 - \hat{\rho} + \Sigma_N,$$

где $\hat{\rho} = \rho_{k,k+1}$ – выборочный коэффициент корреляции между соседними отсчетами данных, имеющий распределение [9]:

$$f(\rho, \hat{\rho}) =$$

$$= [(N-2)/\pi]\left(1-\rho^2\right)^{\frac{N-1}{2}}\left(1-\hat{\rho}^2\right)^{\frac{N-4}{2}} \times \qquad (4)$$

$$\times \int_0^1 \frac{\lambda^{N-2} d\lambda}{(1-\rho\hat{\rho}\lambda)^{N-1}\sqrt{1-\lambda^2}},$$

со средним $M\hat{\rho} = \rho + o(1/N)$ и дисперсией

$$D\hat{\rho} = \left(1-\rho^2\right)^2 \Big/ N + o\left(1/N^{3/2}\right).$$

Таким образом, если случайной добавкой $\sum_N$ в (3) можно пренебречь для некоторого $N$, то критерий Хальда-Аббе сводится к известному [9, 10] критерию некоррелированности выборки. Соответственно нормализуемый критерий (2) согласно (3) переходит в известный критерий оценки значимости выборочного коэффициента корреляции

$$\begin{gathered} r' = 0{,}5 \cdot 1n[(2-r)/r] \approx \\ \approx -0{,}5 \cdot 1n[(1-\hat{\rho})/(1+\hat{\rho})]\ , \end{gathered} \qquad (5)$$

который при $N > 10$ распределен приблизительно нормально [9] с центром

$$M_{r'} = -0{,}5 \cdot 1n[(1-\rho)/(1+\rho)]$$

и дисперсией $D_{r'} = 1/(N-3)$.

Собственно с $r$-критерием в прикладной статистике связывается важный частный случай распределения (4) при $\rho = 0$ (гипотеза некоррелированности), при котором указанное распределение приобретает вид [11]:

$$f(\hat{\rho}) = \frac{\Gamma\left(\frac{N-1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{N-2}{2}\right)}\left(1-\hat{\rho}^2\right)^{\frac{N-4}{4}}, \qquad (6)$$

где $\Gamma(\chi)$ – гамма-функция.

Как следует из (3), соотношение $r = 1 - \hat{\rho}$ справедливо с точностью до случайной добавки

$$\Sigma_N = \left(y_{N+1}^2 - y_1^2\right) \Big/ 2\sum_{k=1}^{N}(y_k - \bar{y})^2, \qquad (7)$$

влияние которой необходимо установить для различных значений объема выборки $N$.

Можно показать, что для стандартизованых нормально распределенных случайных величин (СВ) распределение указанной случайной добавки имеет вид

$$f_{\Sigma}(\Sigma)=\frac{1}{\pi\Gamma\left(\frac{N}{2}\right)}\int_0^{\infty}u^{\frac{N}{2}}e^{-u}\int_0^{\infty}\frac{1}{\sqrt{1-\left(\frac{u}{2}\Sigma\right)^2}}\frac{e^{-t}}{t}dtdu. \quad (8)$$

Выражение (8) получено по формулам [12] определения плотности вероятности СВ в виде отношения разности квадратов гауссовских СВ к сумме квадратов таких же СВ, распределенной по закону $\chi^2(N)$. Моменты (8) определяются с помощью табличных интегралов [13]:

$$m_{\kappa}=\frac{1}{2^{\kappa}\sqrt{\pi}}\frac{\Gamma\left(\frac{\kappa+1}{2}\right)\Gamma(\kappa+1)}{\Gamma\left(\frac{\kappa+1}{2}\right)\Gamma\left(\frac{N}{2}\right)}\Gamma\left(\frac{N}{2}-\kappa\right) \quad (9)$$

при $\kappa = 2n$, и $m_{\kappa} = 0$ при $\kappa = 2n + 1$, $n = 0, 1, 2,\ldots$ .

Таким образом, СВ $\Sigma_N$ является центрированной, с модой и медианой, равными нулю, а согласно (9) ее дисперсия при $\kappa = 2$ определяется выражением $D_{\Sigma_N}=1/(N-2)(N-4)$. Поскольку дисперсия (6) имеет вид [13]:

$$D\hat{\rho}=\frac{\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N}{4}\right)}{2\Gamma\left(\frac{N-2}{2}\right)\Gamma\left(\frac{N}{4}+\frac{3}{2}\right)},$$

то уже при $N > 10$ дисперсии отличаются по крайней мере на порядок: $D_{\Sigma_{N=10}}/D\hat{\rho}=0{,}097$.

По отношению к распределению (6) исследуемое распределение (8) является $\delta$-образным, поэтому плотность суммы выборочного коэффициента корреляции и СВ $\Sigma_N$ в виде свертки указанных плотностей совпадает с (6) по крайней мере до тех пор, пока отношение их дисперсий является малой величиной. Более точное представление о виде свертки двух функций, одна из которых $\delta$-образная, можно получить на основе методики [14], согласно которой

$$f_1(x)*f_2(x)\approx f_1(x)\cdot\int_{-\infty}^{\infty}f_2(x)dx$$

при выполнении условия

$$f_1(x)\cdot\left[\frac{d^2f_1(x+\lambda)}{d\lambda^2}\Big|_{\lambda=0}\right]^{-1}>>$$

$$>>\int_{-\infty}^{\infty}x^2f_2(x)dx\cdot\left[\int_{-\infty}^{\infty}f_2(x)dx\right]^{-1}.$$

Применение этого условия к плотности в виде (6) и (8) дает следующее неравенство

$$\frac{4\left(1-\hat{\rho}^2\right)^2}{(N-4)\left[\rho^2(N-6)-23\right]}>>D_{\Sigma_N}=\frac{1}{(N-2)(N-4)}. \quad (10)$$

При уровне значимости $\alpha = 0{,}01$ и объеме выборки $N = 20$ пороговое значение статистики $\hat{\rho}$ составляет 0,56. Для таких исходных данных правая и левая части (10) отличаются во всей области $|\hat{\rho}|\le\rho_0$ более чем на порядок (в 14 раз на границе $\hat{\rho}=\rho_0$).

Наконец, с точки зрения влияния $\Sigma_N$ на вычисляемый по уровню значимости квантиль распределения (6), приведем результаты численного интегрирования свертки (6) и (8) при $N = 20$ (табл. 1).

Таблица 1

Результаты численного интегрирования

| $P$, % | $\Delta P$, % | $P_{ср}$, % |
|---|---|---|
| 95 | 0,2 | 94,8 |
| 99 | 0,29 | 98,71 |
| 99,9 | 0,39 | 99,51 |

Как следует из табл. 1, случайная добавка $\Sigma_N$ уменьшает вероятность непревышения порога $P_ф$ против заданной доверительной $P$, однако величина этого смещения при $N = 20$ не является значимой.

Таким образом, трендовый критерий Хальда-Аббе по крайней мере при $N \ge 20$ не имеет статических различий по сравнению с известным критерием некоррелированности выборки.

Следовательно, отвергая опорную гипотезу об отсутствии тренда согласно классической формулировке критерия Хальда-Аббе, тем самым отвергается гипотеза о некорреллированности выборки из данных регистрации. Однако при обработке данных в АСКД указанные гипотезы не являются равносильными. Действительно, коррелированность выборки может быть обусловлена как физическими явлениями в объекте диагностирования – ГТД, так и фильтрацией данных в измерительном канале, что вовсе не предполагает наличие тренда. Поскольку модели изменения параметров ГТД детально описаны [4], заметим, что необходимая операция фильтрации данных закономерно превращает исходную выборку в коррелированную: $\rho = \omega(\Delta t)$, где $\omega(\Delta t)$ – значение импульсной переходной функции фильтра в момент времени $t_{\kappa+1} = t_\kappa + \Delta t$. При учете свойств объекта и/или высокочастотной цифровой фильтрации в измерительном канале априорное значение коэффициента корреляции положительно: $\rho > 0$. Согласно (5) происходит смещение среднего уровня решающей статистики в сторону установленного порога и, соответственно, рост уровня ложных тревог о наличии тренда. Для выборки с $N = 20$ и установленным уровнем значимости $\alpha = 0{,}05$ при априорном $\rho = 0{,}3$, признаваемым согласно [15] малозначимым, происходит рост уровня ложных тревог в три раза до $\alpha = 0{,}15$.

Применяемая предварительная обработка данных с целью исключения скрытых периодичностей временного ряда [3] сопровождается режекторной цифровой фильтрацией, для которой характерно $\omega(\Delta t) < 0$. В этом случае согласно (5) средний уровень решающей статистики смещается в противоположную от порога сторону, что сопровождается повышением вероятности пропуска трендового участка выборки. Коррекция порогового уровня возможна только в том случае, если известен априорный коэффициент корреляции, так как оценка выборочного коэффициента корреляции на короткой выборке имеет низкую достоверность и сводится к той же статистике (5). Априорный коэффициент корреляции, таким образом, должен быть определен до применения процедуры трендового контроля путем построения адекватной диагностической модели процесса порождения данных и тщательной оценке метрологических характеристик измерительных каналов.

Не менее важным фактором, влияющим на эффективность трендовых статистик, в том числе исследуемой Хальда-Аббе, является возможное отличие исходной выборки от Гаусовской. Анализ типовых измерительных каналов [15] показывает, что распределения ошибок измерения зачастую имеют существенно негаусовский характер, а при использовании методов косвенных измерений негауссовость выборки является скорее правилом, чем исключением.

Если исследуемая выборка имеет плотность распределения, отличную от нормального, то соответственно нет основания полагать, что статистика (1) совпадает с *r*-статистикой, а статистики (2), (5) близки к нормальному распределению, и уровни значимости критериев выбраны правильно. Разнообразие распределений ошибок различных измерительных каналов [15] не позволяет получить аналитические соотношения для искомых распределений трендовых статистик, поэтому для оценки влияния этого фактора использован метод статистического моделирования в интерактивной среде MATLAB. При таком моделировании использовались выборки из нормальной генеральной совокупности, которая подвергалась нелинейному преобразованию, и на полученной негауссовской выборке тестировались статистики (1), (2). Совместному анализу подвергались гистограммы исходной выборки и значений решающих статистик, а также значения выборочных моментов до четвертого порядка включительно. Как это установлено в результате численных экспериментов, наиболее существенное влияние на уровень значимости критериев тренда оказывает асимметрия распределения выборки, что выражается в росте вероятностей ошибок первого и второго рода.

## Заключение

Применяемый в АСКД трендовый критерий Хальда-Аббе при объеме выборки $N \geq 20$ не имеет статистических различий с критерием коррелированности выборки. Априорная коррелированность временного ряда, обусловленная свойствами объекта либо фильтрацией данных в измерительном канале, приводит к смещению уровня решающей статистики, что сопровождается возрастанием вероятностей ошибок обнаружения либо пропуск трендового участка.

Отличие распределения выборки от нормального также приводит к снижению эффективности трендовых статистик. Корректный выбор и определение параметров трендовых статистик могут быть выполнены путем предварительного построения адекватных моделей порождения данных (диагностических моделей) применительно к параметрам регистрации в АСКД, что и определяет перспективы дальнейших исследований.

## Литература

1. Жигалевский А.А., Красновский А.Е. Обнаружение разладки случайных процессов в задачах радиотехники. – Л.: Изд. Ленинград. ун-та, 1988. – 224 с.

2. Айвазян С.А., Енюков И.С. Мешалкин Л.Д. Прикладная статистика. Исследование зависимостей. – М.: Финансы и статистика, 1985. – 488 с.

3. Отнес Р., Эноксон Л. Прикладной анализ временных рядов. Основные методы. – М.: Мир, 1982. – 388 с.

4. Синтез систем управления и диагностирования газотурбинных двигателей / С.В. Епифанов, В.И. Кузнецов, И.И. Богаенко и др. – К.: Техника, 1998. – 312 с.

5. Кармалита В.А., Лобанов В.Э. Точность результатов автоматизированного эксперимента. – М.: Машиностроение, 1992. – 208 с.

6. Кулик Н.С., Тамаргазин А.А., Линник И.И. Показатели качества функционирования авиационных ГТД // Авиационно-космическая техника и технология. – Х.: Нац. аэрокосм. ун-т «ХАИ». – 2002. – Вып. 31. – С. 4 – 7.

7. Симкин Э.Л., Гагай В.С., Семенова Т.А. и др. Автоматизированный параметрический контроль технического состояния ТРДД НК-8-2У в эксплуатации по полетной информации // Авиационно-космическая техника и технология. – Х.: Нац. аэрокосм. ун-т «ХАИ». – 2001. – Вып. 26. – С. 222 – 227.

8. Опыт эксплуатации автоматизированных систем диагностирования газотурбинных приводов семейства НК на газоперекачивающих станциях / В.Б. Коротков, В.Н. Михнович, В.А. Оболенский и др. // Вестник двигателестроения. – 2004. – № 2. – С. 165 – 168.

9. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. – М: Наука, 1973. – 832 с.

10. Математическая статистика / Под ред. В.С. Зарубина и А.П. Крищенко. – М.: МВТУ им Н.Э. Баумана, 2002. – 452 с.

11. Тейлор Дж. Введение в теорию ошибок. – М.: Мир, 1985. – 272 с.

12. Горяинов В.Т., Журавлев А.Г., Тихонов В.И. Статическая радиотехника: Примеры и задачи / Под ред. В.И. Тихонова. – М.: Сов. радио, 1980. – 544 с.

13. Двайт Г.Б. Таблицы интегралов. – М.: Наука, 1978. – 228 с.

14. Теория когерентных изображений / Под ред. Н.Д. Устинова. – М.: Радио и связь, 1987. – 264 с.

15. Новицкий П.В., Заграф И.А. Оценка погрешностей результатов измерений. – Л.: Энергоатомиздат, 1985. – 248 с.

*Поступила в редакцию 4.07.2005*

**Рецензент:** д-р техн. наук, проф. С.А. Нестеренко, Одесский национальный политехнический университет, Одесса.