УДК 519.281:621.391.2

В. В. Пусь, д-р техн. наук
Санкт-Петербургский университет ГПС МЧС России
(Россия, 196105, Санкт-Петербург, Московский пр-т, 149,
тел. (812)3881047, E-mail: pous@list.ru)

Нецентральное распределение статистики Фишера в гармоническом анализе

(Статью представил д-р техн. наук В.П. Симоненко)

Использован тест Фишера для проверки гипотезы H_0 относительно взаимозаменяемой альтернативы H_j. Получено распределение статистики Фишера при гипотезе H_j и вычислены вероятности правильного выбора этой гипотезы.

Використано тест Фішера для перевірки гіпотези H_0 відносно взаємозамінної альтернативи H_j. Отримано розподіл статистики Фішера при гіпотезі H_j та обчислено імовірності правильного вибору цієї гіпотези.

К л ю ч е в ы е с л о в а: анализ периодограмм, аномальное наблюдение, экспоненциальное распределение, множественные решения, вероятность правильного выбора.

Для выявления скрытых периодичностей при анализе временных рядов статистические данные аппроксимируют тригонометрическим полиномом

$$\sum_{k=1}^{m}(A_k \cos w_k t + B_k \sin w_k t) = \sum_{k=1}^{m} R_k \cos(w_k t + \varphi_k),$$

где $R_k^2 = A_k^2 + B_k^2$, $tg\,\varphi_k = B_k / A_k$. При этом полагают, что если периодичность отсутствует, то коэффициенты A_k и B_k взаимно независимы и нормально распределены с нулевым средним и дисперсией σ^2. Тогда квадраты амплитуд R_k также взаимно независимы и имеют $\sigma^2\chi_2^2$ распределение с двумя степенями свободы (экспоненциальное).

При значимом отклонении величины R_k^2 от нуля приходят к заключению о наличии периодичности. Такое статистическое обоснование задачи впервые дал А. Шустер (метод периодограмм) [1], а решающее правило, инвариантное к σ^2, предложил Р. Фишер [2]. В современной трактовке указанная задача обсуждается в рамках критериев для аномальных наблюдений (выбросов) [3, 4], или, в технических приложениях, как задача обнаружения с различением (классификации) сигналов [5, 6].

Постановка задачи. Предварительные замечания. Сформулируем тест Фишера в терминах проверки гипотез. Пусть

$$X_1, ..., X_j, ..., X_m \tag{1}$$

есть выборка из m популяций (каждое X_k может быть достаточной статистикой, основанной на нескольких наблюдениях), элементы которой независимы, подчиняются $\sigma^2\chi_2^2$ распределению с двумя степенями свободы, за исключением, возможно, одного элемента, имеющего нецентральное $\sigma^2\chi_2^2(\lambda)$ распределение с двумя степенями свободы, и параметром нецентральности λ.

Проверяем гипотезу H_0:

$$X_k \sim \sigma^2\chi_2^2, \ k=1, ..., m, \tag{2}$$

при взаимозаменяемой альтернативе $\overline{H}$ — верна одна из равновероятных гипотез H_j ($j = 1, ..., m$):

$$X_j \sim \sigma^2\chi_2^2(\lambda), \ X_k \sim \sigma^2\chi_2^2, \ j \neq k, \ k=1, ..., m, \tag{3}$$

где символ $\sim$ означает принадлежность, или подчиненность; параметры j, λ, σ^2 неизвестны.

Замечание. Случайная величина (СВ) X_j в (3), подчиняющаяся $\sigma^2\chi_2^2(\lambda)$ распределению, описывается плотностью вероятностей:

$$p(x_j) = (1/(2\sigma^2))\exp(-(x_j/\sigma^2 + \lambda^2)/2))\,\mathrm{I}_0[\lambda x_j^{1/2}/\sigma], \ 0 \le x_j < \infty, \tag{4}$$

где $\mathrm{I}_0[\cdot]$ — модифицированная функция Бесселя первого рода нулевого порядка. Кроме того, X_j является суммой квадратов двух независимых нормально распределенных СВ, U_j и V_j, со средними μ_{1j}, μ_{2j} и дисперсией σ^2. Учитывая это, параметр λ (λ^2) можно записать в виде

$$\lambda = (\mu_{1j}^2 + \mu_{2j}^2)^{1/2}/\sigma \quad (\lambda^2 = (\mu_{1j}^2 + \mu_{2j}^2)/\sigma^2). \tag{5}$$

В литературе не сложилось устоявшегося мнения о том, как обозначать параметр нецентральности. Например в [7, 8] так называют конструкцию, подобную λ^2из (5) (называя при этом «параметр нецентральности λ^2»); в [9, 10] выражение, похожее на правую часть λ^2, обозначают $\lambda(m)$ и называют параметром нецентральности $\lambda(m)$.

Здесь и далее под нецентральным $\sigma^2\chi_2^2(\lambda)$ распределением с двумя степенями свободы и параметром нецентральности λ будем понимать распределение с плотностью (4) и параметром λ (5), т.е. следуя логике [7, 8], для краткости параметр λ будем записывать без степени. Такое написание λ оправдано тем, что в радиотехнике параметру λ (λ^2) соответствует первичное (вторичное) понятие «отношение сигнал—шум по напряжению (по мощности)». Следует заметить, что $\sigma^2\chi_2^2(\lambda)$ распределение подстановкой $w = x_j / \sigma^2$ можно привести к стандартной нецентральной плотности $p(w) = (1/2)\exp(-(w+\lambda^2)/2))\,\mathrm{I}_0[\lambda w^{1/2}]$), которая при $\lambda = 0$ становится плотностью стандартного χ_2^2 распределения.

Процедура выбора из гипотез (2), (3) основана на критерии Фишера, согласно которому гипотеза H_0 принимается, если

$$\left(Y_k = X_k \Big/ \sum_{r=1}^{m} X_r \right) < C, \; k = 1, \ldots, m, \tag{6}$$

и отвергается в пользу одной из гипотез H_j, $j = 1, \ldots, m$, если

$$Y_j > C, \; Y_j > Y_k, \; k \neq j, \; k = 1, \ldots, m. \tag{7}$$

Здесь C — верхняя α-значимая точка, определяемая решением уравнения

$$\alpha = \mathbf{P}\left\{ \max_{1 \le k \le m} Y_k > C | H_0 \right\} = \sum_{r=1}^{t} (-1)^{r-1} C_m^r (1 - rC)^{m-1}, \tag{8}$$

где $\mathbf{P}\{\cdot\}$ — вероятность события $\{\cdot\}$; $t = [1/C]$ ($[x]$ — целая часть x); $C_m^r = m!((m-r)!r!)^{-1}$ — биномиальные коэффициенты.

Процедура выбора (6), (7) при условии (8) — равномерно наиболее мощная в классе симметричных инвариантных процедур [7]. В радиолокации процедура (6), (7) интерпретируется как задача обнаружения с различением m ненулевых сигналов в гауссовом шуме с неизвестной интенсивностью. Приемное устройство содержит m ветвей приема, объединенных на входе, и подключенных на выходе через квадратичные детекторы к блоку выбора максимального сигнала. Прием некогерентных сигналов осуществляется с помощью согласованной фильтрации или (взаимно) корреляционной обработки в каждой ветви (синфазной и квадратурной подветвях) приема, выбора максимального сигнала и сравнения его с выходным напряжением перемножителя, на входы которого подаются пороговое напряжение и сумма напряжений с выходов всех ветвей приема (процедура перемножения технически проще, чем операция деления). Выбору гипотезы H_0 соответствует принятие решения об отсутствии какого-либо сигнала на входе приемника, выбору гипотезы H_j — принятие решения о наличии сигнала j-й позиции ($j = 1, \ldots, m$).

Обозначим через G наибольшее из отношений Y_k в (6), (7), т.е. положим $G = \max_{1 \le k \le m} X_k / (X_1 + \ldots + X_m)$. Распределение G при нулевой гипотезе ($\lambda = 0$) известно [2, 8] и описывается плотностью

$$f(g) = (m-1) \sum_{r=1}^{m} (-1)^{r-1} r C_m^r (1 - rg)^{m-2} \tag{9}$$

для $0 < g < 1/r$ и $f(g) = 0$ в противном случае.

Несмотря на давность работы [2] и продолжающийся интерес к тесту Фишера в различных областях знаний [11—13], о распределении статистики G при ненулевой гипотезе не имеется новых данных — рабочие характеристики теста по-прежнему оценивают моделированием [14—16].

Найдем распределение статистики G при гипотезе H_j и на его основе вычислим вероятность правильного выбора (ПВ) этой гипотезы, т.е. вероятность принятия правильного решения.

Вспомогательные утверждения. Сформулируем условия, при которых осуществляется ПВ гипотезы H_j .

1. Повторная выборка (1) включает аномальный элемент X_j, имеющий нецентральное $\sigma^2\chi_2^2(\lambda)$ распределение с параметром нецентральности λ; при этом индексы элемента X_j и гипотезы H_j совпадают.

2. Элемент X_j является максимальным в выборке (1), т.е. $X_j > X_k$ для всех $k \neq j$, $k = 1, ..., m$.

Расположим СВ (1) в порядке возрастания их значений:

$$Z_1 \le Z_2 \le ... \le Z_m. \tag{10}$$

Будем считать, что условия ПВ выполнены. Тогда $Z_m = X_j$.

Для нахождения распределения статистики G при гипотезе H_j необходимо знать совместное распределение элементов вариационного ряда (10). Такое распределение получено в [17] на основе следующего вспомогательного утверждения.

Лемма 1. Пусть выполнены условия ПВ и (1) — независимые СВ, из которых $m-1$ подчиняются $(1/2a)\chi_2^2$ распределению с двумя степенями свободы и плотностью $p(x_k) = a\exp(-ax_k)$, $(a = 1/(2\sigma^2))$, $0 \le x_k < \infty$, $k = 1, ...$ $..., j-1,\ j+1, ..., m$, а j-я СВ ($j \neq k$, $j = 1, ..., m$, $k = 1, ..., m$) имеет нецентральное $(1/2a)\chi_2^2(\lambda)$ распределение с параметром нецентральности λ и плотностью

$$p(x_j) = a\exp(-ax_j - \lambda^2/2)\,\mathrm{I}_0[\lambda(2ax_j)^{1/2}],\ 0 \le x_j < \infty, \tag{11}$$

совпадающей (с учетом более удобного обозначения $a = 1/(2\sigma^2)$ с (4)). Тогда вероятность $\mathbf{P}$ того, что m СВ расположатся в (1) в виде неубывающей последовательности (10) с наибольшим элементом Z_m, равным X_j, определяется выражением

$$\mathbf{P}\{Z_1 \le ... \le Z_{m-1} \le Z_m\} = (S_\lambda / m!)\exp(-\lambda^2/2), \tag{12}$$

где

$$S_\lambda = \sum_{r=1}^{m}(-1)^{r-1} C_m^r \exp(\lambda^2/(2r)). \tag{13}$$

Лемма 1 доказывается прямым вычислением вероятности

$$\mathbf{P}\{Z_1 \le \ldots \le Z_{m-1} \le Z_m\} =$$
$$= a^m \exp(-\lambda^2/2) \int_0^\infty \exp(-az_m)\, \mathrm{I}_0[\lambda\,(2az_m)^{1/2}]\, dz_m \times$$
$$\times \int_0^{z_m} \exp(-az_{m-1})\, dz_{m-1} \ldots \int_0^{z_2} \exp(-az_1)\, dz_1. \tag{14}$$

Элемент вероятности порядковых статистик $Z_1, \ldots, Z_{m-1}, Z_m$ в области их изменения $0 \le Z_1 \le \ldots \le Z_{m-1} \le Z_m < \infty$ находим по известному правилу $\mathbf{P}\{Z_1, \ldots, Z_m | Z_1 \le, \ldots, \le Z_m\} = \mathbf{P}\{X_1, \ldots, X_m\} / \mathbf{P}\{Z_1 \le Z_2, \ldots, \le Z_m\}$ делением элемента вероятности неупорядоченных компонент $X_1, \ldots, X_j, \ldots, X_m$

$$a^m \exp(-\lambda^2/2 - a(x_1 + \ldots + x_j + \ldots + x_m))\, \mathrm{I}_0[\lambda\,(2ax_j)^{1/2}]\, dx_1 \ldots dx_j \ldots dx_m \tag{15}$$

на элемент вероятности конкретного упорядочения этих компонент в области изменения, задаваемой неравенствами $0 \le Z_1 \le \ldots \le Z_{m-1} \le Z_m$. Этот элемент представляет собой интеграл от (15) по указанной области, т.е. правую часть соотношения (14), значение которого, согласно лемме 1, определяется выражением (12). Отсюда элемент вероятности порядковых статистик $Z_1, \ldots, Z_{m-1}, Z_m$ при гипотезе H_j имеет вид

$$(m!/S_\lambda)\, a^m \exp(-a(z_1 + \ldots + z_{m-1} + z_m))\, \mathrm{I}_0[\lambda\,(2az_m)^{1/2}]\, dz_1 \ldots dz_{m-1} dz_m.$$

Задача нахождения совместной плотности порядковых статистик $Z_1, \ldots, Z_{m-1}, Z_m$ при гипотезе H_j может быть решена и другим способом.

Пусть по-прежнему выполняются условия ПВ. Совместное распределение статистик Z_k, $k = 1, \ldots, m-1$, образующих нижнюю часть вариационного ряда (10), известно (см., например, [18, (75)]) и может быть записано в виде $(m-1)!a^{m-1} \exp(-az_1 + \ldots + z_{m-1})$, $0 \le z_1 \le \ldots \le z_{m-1} < \infty$. Для того чтобы получить совместное распределение всех порядковых статистик Z_k, $k = 1, \ldots$ $\ldots, m$, необходимо последнее выражение умножить на плотность (11) нецентрального $(1/2a)\,\chi_2^2(\lambda)$ распределения

$$K\,(m-1)!a^m \exp(-a(z_1, + \ldots + z_{m-1} + z_m) - \lambda^2/2)\, \mathrm{I}_0[\lambda\,(2az_m)^{1/2}] \tag{16}$$

и нормирующую константу, например K, которую можно найти исходя из условия нормировки и способа задания вариационного ряда (10), а именно условия о том, что Z_m является его максимальным элементом. Эти ограничения формулируются как следующее требование:

$$1 = K\,(m-1)!a^m \exp(-\lambda^2/2) \int_0^\infty \exp(-a\, z_m)\, \mathrm{I}_0[\lambda\,(2az_m)^{1/2}]\, dz_m \times$$

$$\times \int_0^{z_m} \exp(-az_{m-1})\, dz_{m-1} \ldots \int_0^{z_2} \exp(-az_1)\, dz_1. \tag{17}$$

Выполнив в (17) интегрирование, получим $K = m\exp(\lambda^2/2)/S_\lambda$. Подставив значение K в (16), окончательно получим

$$(m!/S_\lambda)\, a^m \exp(-a(z_1 + \ldots + z_{m-1} + z_m))\, \mathrm{I}_0[\lambda(2az_m)^{1/2}],$$

$$0 \le z_1 \le \ldots \le z_{m-1} \le z_m < \infty. \tag{18}$$

Таким образом, справедливо следующее вспомогательное утверждение.

Лемма 2. Пусть выполнены условия ПВ. Тогда совместная плотность элементов вариационного ряда (10) определяется выражением (18).

Вывод нецентрального распределения. Для получения распределения $G = Z_m/(Z_1 + \ldots + Z_m)$ при гипотезе H_j вначале, с учетом (18), найдем моменты распределения λ, записав h-й момент в форме легко проверяемого соотношения [8, (17.6.24)]

$$\mathbf{E}(\lambda^h) = \int_{-\infty}^{0} \ldots \int_{-\infty}^{0} \mathbf{E}(\exp(a(z_1 + \ldots + z_m)\varphi)(az_m)^h)\, dt_1 \ldots dt_h, \tag{19}$$

где $\mathbf{E}$ — символ математического ожидания; $\varphi = t_1 + \ldots + t_h$. Интегрируя $\mathbf{E}(\exp(a(z_1 + \ldots + z_{m-1} + z_m)\varphi)(az_m)^h)$ по $z_1, \ldots, z_{m-1}$, находим

$$\mathbf{E}(\exp(a(z_1 + \ldots + z_{m-1} + z_m)\varphi)(az_m)^h) =$$

$$= (m!/S_\lambda)\, a^{m+h} \int_0^\infty \exp(-a(1-\varphi)\, z_m)\, z_m^h\, \mathrm{I}_0[\lambda(2az_m)^{1/2}]\, dz_m \times$$

$$\times \int_0^{z_m} \exp(-a(1-\varphi) z_{m-1})\, dz_{m-1} \ldots \int_0^{z_2} \exp(-a(1-\varphi)\, z_1)\, dz_1 = ma^{h+1}[S_\lambda (1-\varphi)^{m-1}]^{-1} \times$$

$$\times \int_0^\infty \mathrm{I}_0[\lambda(2az_m)^{1/2}] \exp(-a(1-\varphi)\, z_m)(z_m)^h (1 - \exp(-a(1-\varphi)\, z_m))^{m-1}\, dz_m.$$

Применив в последнем интеграле подстановку $u = z_m^{1/2}$ и записав его в форме

$$2\sum_{r=1}^{m} (-1)^{r-1} r C_m^r \int_0^\infty \mathrm{I}_0[\lambda/(2a)^{1/2} u]\, u^{2h+1} \exp(-ra(1-\varphi)\, u^2)\, du,$$

после интегрирования [19, (6.631.1)] получим

$$\mathbf{E}(\exp(a(z_1+\ldots+z_{m-1}+z_m)\varphi)(az_m)^h)=$$

$$=(1/S_\lambda)\sum_{r=1}^{m}(-1)^{r-1}C_m^r[\Gamma(h+1)/(r^h(1-\varphi)^{m+h})]_1F_1(h+1;1;\lambda^2/(2r(1-\varphi))), \quad (20)$$

где ${}_1F_1(\cdot;\cdot;\cdot)$ — вырожденная гипергеометрическая функция [20, с. 237]. Подставив правую часть (20) в (19) и разложив функцию ${}_1F_1(\cdot;\cdot;\cdot)$ в ряд, после интегрирования по t_1, ..., t_h найдем

$$\mathbf{E}(g^h)=((m-1)/S_\lambda)\times$$

$$\times\sum_{r=1}^{m}(-1)^{r-1}C_m^r[\Gamma(m-1)\Gamma(h+1)/(r^h\Gamma(m+h))]_2F_2(m,h+1;m+h,1;\lambda^2/(2r)), \quad (21)$$

где ${}_2F_2(\cdot;\cdot;\cdot)$ — обобщенный гипергеометрический ряд [20, с.183].

Из интегральных соотношений для обобщенных гипергеометрических рядов [19, (7.512.12)]

$$\int_0^1(1-x)^{\mu-1}x^{\nu-1}{}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;ax)\,dx=$$

$$=[\Gamma(\mu)\Gamma(\nu)/\Gamma(\mu+\nu)]_{p+1}F_{q+1}(\nu,a_1,\ldots,a_p;\mu+\nu,b_1,\ldots,b_q;a),$$

или

$$\int_0^1(1-x)^{\mu-1}x^{\nu-1}{}_1F_1(a_1;b_1;ax)\,dx=$$

$$=[\Gamma(\mu)\Gamma(\nu)/\Gamma(\mu+\nu)]_2F_2(\nu,a_1;\mu+\nu,b_1;a),$$

заключаем, что

$$[\Gamma(h+1)\Gamma(m-1)/\Gamma(h+m)]_2F_2(m,h+1;h+m,1;\lambda^2/(2r))=$$

$$=\int_0^1 x^h(1-x)^{m-2}{}_1F_1(m;1;\lambda^2x/(2r))\,dx. \quad (22)$$

Таким образом, с учетом (22) выражение (21) можно записать в виде

$$\mathbf{E}(g^h)=((m-1)/S_\lambda)\sum_{r=1}^{m}(-1)^{r-1}r^{-h}C_m^r\int_0^1 x^h(1-x)^{m-2}{}_1F_1(m;1;\lambda^2/(2r))\,dx. \quad (23)$$

Выполнив в (23) подстановку $y = x/r$, получим

$$\mathbf{E}(g^h) = ((m-1)/S_\lambda)\sum_{r=1}^{m}(-1)^{r-1} r\, C_m^r \int_0^{1/r} y^h (1-ry)^{m-2}\, {}_1F_1(m;1;\lambda^2 y/2)\, dy. \quad (24)$$

Введем, по аналогии с [8, 17.6.1], функцию

$$\theta_r(y) = \begin{cases} (1-ry)^{m-2}\, {}_1F_1(m;1;\lambda^2 y/2), 0 < y < 1/r, \\ 0 \text{ в остальных точках,} \end{cases}$$

и запишем (24) в виде

$$\mathbf{E}(g^h) = ((m-1)/S_\lambda)\int_0^1 \sum_{r=1}^{m}(-1)^{r-1} r\, C_m^r y^h \theta_r(y)\, dy,$$

где h = 1, 2,… Из теоремы о единственности представления функции распределения ее моментами [8, 5.5.1] заключаем, что G и Y распределены одинаково. Следовательно, искомая функция плотности вероятности G при гипотезе H_j определяется выражением

$$f(g) = ((m-1)/S_\lambda)\sum_{r=1}^{m}(-1)^{r-1} r\, C_m^r \theta_r(g). \quad (25)$$

Сформулируем полученный результат. Пусть выполнены условия ПВ. Тогда функция плотности вероятностей статистики $G = X_j/(X_1 + ... + X_m)$ определяется формулой (25), или в развернутом виде

$$f(g) = \begin{cases} \dfrac{m-1}{S_\lambda}\displaystyle\sum_{r=1}^{m}(-1)^{r-1} r\, C_m^r (1-rg)^{m-2}\, {}_1F_1(m;1;\dfrac{\lambda^2 g}{2}, \ 0 < g < 1/r, \\ 0 \text{ в остальных точках,} \end{cases} \quad (26)$$

где ${}_1F_1(\cdot;\cdot;\cdot)$ — вырожденная гипергеометрическая функция. Очевидно, что при $\lambda = 0$ плотность (26) совпадает с (9). Проверим условие нормировки

$$\int_0^1 f(g)\, dg = \frac{(m-1)}{S_\lambda}\sum_{r=1}^{m}(-1)^{r-1} r\, C_m^r \int_0^{1/r} [1-rg]^{m-2}\, {}_1F_1\left(m;\, 1;\, \frac{\lambda^2 g}{2}\right) dg.$$

Из интегрального соотношения [21, 13.1, (95)]

$$[\Gamma(\mu)^{-1}\int_0^y x^{\nu-1}\, {}_1F_1(a_1;b_1;ax)(y-x)^{\mu-1}\, dx =$$

$$= [y^{\mu+\nu-1}\Gamma(\nu)/\Gamma(\mu+\nu)]\ {}_2F_2(\nu, a_1;\, \mu+\nu, b_1;\, ay)$$

следует

$$\int_0^{1/r} r^{m-2}(1/r-g)^{m-2}\,{}_1F_1(m;1;\lambda^2 g/2)\,dg =$$

$$=[(m-1)\,r]^{-1}\,{}_2F_2(1,m;m,1;\lambda^2/2r))=[(m-1)\,r]^{-1}\,e^{\,\lambda^2/(2r)}.$$

Отсюда

$$\int_0^1 f(g)\,dg=\frac{1}{S_\lambda}\sum_{r=1}^{m}(-1)^{r-1}C_m^r e^{\,\lambda^2/(2r)}=1.$$

Следовательно, вероятность ПВ гипотезы H_j, т.е. правильной идентификации j-го аномального наблюдения ($P_{\text{пр}}=\mathbf{P}\{G>C\}$), определяется так:

$$P_{\text{пр}}=\int_C^1 f(g)\,dg=\frac{(m-1)}{S_\lambda}\sum_{r=1}^{t}(-1)^{r-1}r\,C_m^r\left\{\int_0^{1/r}(1-rg)^{m-2}\,{}_1F_1(m;1;\frac{\lambda^2 g}{2})\,dg-\right.$$

$$\left.-\int_0^C(1-rg)^{m-2}\,{}_1F_1\left(m;1;\frac{\lambda^2 g}{2}\right)dg\right\}=\frac{(m-1)}{S_\lambda}\sum_{r=1}^{t}(-1)^{r-1}r\,C_m^r\{I_1-I_2\}. \quad (27)$$

Здесь t — верхний предел суммы, т.е. наибольшее целое, не превосходящее значения m, для которого $1-tC\ge 0$, $t\le[1/C]$, что определяется выбором C (см. (8)). Из проверки условия нормировки

$$I_1=\int_0^{1/r}(1-rg)^{m-2}\,{}_1F_1\left(m;1;\frac{\lambda^2 g}{2}\right)dg=[r(m-1)]^{-1}\,e^{\,\lambda^2/(2r)}, \quad (28)$$

$$I_2=\int_0^{C}(1-rg)^{m-2}\,{}_1F_1\left(m;1;\frac{\lambda^2 g}{2}\right)dg,\;\; 0<g<r^{-1}\,. \quad (29)$$

Применив в (29) подстановку $w=rg$ и записав функцию ${}_1F_1(\cdot;\cdot;\cdot)$ в виде ряда, получим

$$I_2=[r\,(m-1)]^{-1}\sum_{j=0}^{\infty}\frac{\Gamma\,(m+j)(\lambda^2/2r))^j}{\Gamma\,(m-1)\,\Gamma\,(1+j)}\int_0^{rC}w^j(1-w)^{m-2}dw,$$

или, учитывая соотношения для неполной бета-функции [20, 2.5.3]

$$B_x(p,q)=\int_0^x t^{p-1}(1-t)^{q-1}\,dt,\;\; B_1(p,q)=\frac{\Gamma\,(p)\,\Gamma\,(q)}{\Gamma\,(p+q)},\;\; I_x(p,q)=\frac{B_x(p,q)}{B_1(p,q)},$$

m	$P_{\text{пр}}$ при $\alpha = 0{,}01$				C	$P_{\text{пр}}$ при $\alpha = 0{,}05$				C
	$\lambda = 4$	$\lambda = 5$	$\lambda = 6$	$\lambda = 7$		$\lambda = 4$	$\lambda = 5$	$\lambda = 6$	$\lambda = 7$	
4	0,1910	0,3524	0,5444	0,7222	0,8643	0,4847	0,7052	0,8691	0,9552	0,7679
8	0,3661	0,6403	0,8599	0,9640	0,6152	0,6378	0,8617	0,9683	0,9957	0,5157
16	0,4629	0,7592	0,9384	0,9916	0,3885	0,6951	0,9044	0,9845	0,9988	0,3192
32	0,4975	0,7950	0,9566	0,9957	0,2292	0,7087	0,9032	0,9875	0,9992	0,1880

запишем

$$I_2 = [r(m-1)]^{-1} \sum_{j=0}^{\infty} I_{rC}(1+j, m-1)(\lambda^2/(2r))^j / j!\,. \tag{30}$$

Подставляя в (27) выражения (28) и (30), находим

$$P_{\text{пв}} = \frac{1}{S_\lambda} \sum_{r=1}^{[1/C]} (-1)^{r-1} C_m^r \left\{ e^{\lambda^2/(2r)} - \sum_{j=0}^{\infty} (\lambda^2/(2r))^j / j!)\, I_{rC}(1+j, m-1) \right\}, \tag{31}$$

где $I_x(p,q)$ — неполная бета-функция.

Таким образом, получен следующий результат. При выполнении условий ПВ вероятность ПВ гипотезы H_j, т.е. правильной идентификации статистики $G = X_j / (X_1 + ... + X_m)$, определяется по формуле (31).

Внутреннюю сумму в (31), по аналогии с [22], можно записать в удобной для вычислений рекурентной форме:

$$\sum_{j=0}^{\infty} (\lambda^2/(2r))^j / j!)\, I_{rC}(1+j, m-1) = (rC)^{m-1} \exp(C\lambda^2/2) \sum_{k=0}^{m-2} S_k, \tag{32}$$

где

$$S_0 = 1;\; S_1 = (1-rC)(m-1+C\lambda^2/2)/(rC);$$

$$S_k = (1-rC)((m-k+C\lambda^2/2)S_{k-1} + (1-rC)(\lambda^2/(2r))\,S_{k-2})/(krC).$$

В таблице приведены значения вероятности ПВ с использованием теста Фишера (6), (7), вычисленные непосредственно по формуле (31) и с учетом (32) для различных значений α, m и λ, а также значения констант C, полученные из (8).

Выводы. В работе [23] тест Фишера был обобщен на χ^2 распределенные СВ с произвольным числом степеней свободы, после чего подобные конструкции теста стали использовать для выявления сдвига дисперсии [24—26].

В рассмотренной постановке задачи сохранен подход [1, 2], при котором выявляется наличие сдвига средних значений $((\mu_{1j}^2+\mu_{2j}^2)>0)$ у одной из пар независимых нормально распределенных величин, в чем нетрудно убедится, вычислив среднее значение нецентрального $\sigma^2\chi_2^2(\lambda)$ распределения. Действительно, используя (4) с учетом (5), получаем

$$\int_0^\infty x_j\, p(x_j)\, dx_j = 2\sigma^2(1+\lambda^2/2) = 2\sigma^2 + (\mu_{1j}^2+\mu_{2j}^2),$$

где $(\mu_{1j}^2+\mu_{2j}^2)$ — искомая нецентральная добавка, или сдвиг к среднему значению $(2\sigma^2)$ $\sigma^2\chi_2^2$ распределения.

Таким образом, формула (31) для вычисления вероятности ПВ ненулевой гипотезы H_j $(j=1, ..., m)$ позволяет аналитически оценить эффективность решающего правила Фишера как по отношению к упрощенным решающим правилам (см., например, [17, 27]), так и по отношению к устройствам обнаружения с различением, реализующим это правило.

Fischer's test was used for checking the hypothesis H_0 in respect of interchangeable alternative H_j. The distribution of Fischer's statistics at a hypothesis H_j has been obtained and probabilities of correct choice of this hypothesis have been calculated.

1. *Schuster A.* On the Investigation of Hidden Periodicities with Application to a supposed Twenty-six-day Period of Meteorological Phenomena // Terr. Mag. — 1898. — **3**, № 1. — P. 13—41.
2. *Fisher R. A.* Tests of Significance in Harmonic Analysis // Proc. Roy. Soc. — 1929. — Ser. A. — **125**. — P. 54—59; Fisher R. A. Contribution to Mathematical Statistics. N. Y., John Wiley & Sons, Chapman & Hall, London, 1950. Tests of Significance…/Reprinted from Proc. Roy. Soc. — 1929. — Ser. A, **125**. — P. 54—59; P. 16.53a—16.59a.
3. *Дэйвид Г.* Порядковые статистики. — М. : Наука, 1979. — 336 с.
4. *Barnett V., Lewis T.* Outliers in Statistical Data. 2ed. — Norwich : Page Bros., 1984. — 464 p.
5. *Pfanzagl J.* Ein kombiniertes Test & Klassifications — Problem // Metrika. — 1959. — **2**, № 1. — S. 11—45.
6. *Теория* обнаружения сигналов / П.С. Акимов, П.А. Бакут, В.А. Богданович и др. Под ред. П.А. Бакута. — М. : Радио и связь, 1984. — 440 с.
7. *Андерсон Т.* Статистический анализ временных рядов. — М. : Мир, 1974. — 576 с.
8. *Уилкс С.* Математическая статистика. — М. : Наука, 1967. — 632 с.
9. *Рао С. Р.* Линейные статистические методы и их применения. — М. : Наука, 1968. — 548 с.
10. *Справочник* по теории вероятностей и математической статистике / В.С.Королюк, Н.И.Портенко, А. В. Скороход, А. Ф. Турбин; под ред. В. С. Королюка. — Киев : Наук. думка, 1978. — 584 с.
11. *Dargahi-Noubary G. R.* A Test of the Cyclicity of Earthquakes // Natural Hazards. — 1997. — **16**, № 2—3. — P. 127—134.

12. *Wichert, S., Fokianos K., Strimmer, K.* Identifying Periodically Expressed Transcripts in Microarray Time Series Data // Bioinformatics. — 2004. — **20**. — P. 5—20.
13. *Wang Z., Atchley W.R.* Spectral Analysis of Sequence Variability in Bacic-helix-loop-helix (bHLH) Protein Domains // Evol Bioinform Online. — 2006. — **2.** — P. 187—196.
14. *Darling D. A.* On the Test for Homogeneity and Extreme Value // Ann. Math. Statist. — 1952. — **23**. — P. 450—456.
15. *Ahdesmaki M., Lohdesmaki H., Gracey A. et al.* Robust Regression for periodicity Detection in Non-uniformly Sampled Time-course Gene Expression Data // BMC Bioinformatics. — 2007. — **8**:233.
16. *Liew A. W., Law N. F., Cao X. Q., Yan H.* Statistical Power of Fisher Test for the Detection of Short Periodic Gene Expression Profiles // Patt. Recogn. — 2009. — **42**, № 4. — P. 549—556.
17. *Пусь В. В.* Вероятность правильной классификации для быстрого аналога теста Фишера // Электрон. моделирование. — 2005. — **27**, № 5. — С. 99—105.
18. *Галамбош Я.* Асимптотическая теория экстремальных порядковых статистик. — М. : Наука, 1984. — 304 с.
19. *Градштейн И. С., Рыжик И. М.* Таблицы интегралов, сумм, рядов и произведений. — М. : Физматгиз, 1963. — 1100 с.
20. *Бейтмен Г., Эрдейи А*. Высшие трансцендентные функции. Т.1. Гипергеометрическая функция. Функция Лежандра. — М. : Наука, 1973. — 296 с.
21. *Бейтмен Г., Эрдейи А.* Таблицы интегральных преобразований. Т.2. Преобразования Бесселя. Интегралы от специальных функций. — М. : Наука, 1970. — 328 с.
22. *Tang P. C.* The power Function of the Analysis of Variance Tests with Tables and Illustrations of Their Use // Stat. Res. Mem. — 1938. — P. 126—157.
23. *Cochren W. G.* The Distribution of the Largest of a Set of Estimated Variances as a Fraction of Their Total // Ann. Eugenics. — 1941—1942. — **11**. — P. 47—52.
24. *Truak D.R.* An Optimal Slippage Test for the Variances of *k* Normal Distributions // Ann. Math. Statist. — 1953. — **24**. — № 4. — P. 669—674.
25. *Doornbos R., Prins H. J.* Slippage Test for a Set of Gamma-variates // Indag. Math. — 1956. — **18**. — P. 329—337.
26. *Себер Дж.* Линейный регрессионный анализ. — М. : Мир, 1980. — 456 с.
27. *Пусь В. В.* Тест классификации многопозиционного сигнала, основанный на порядковых статистиках // Изв. вузов Радиоэлектроника.— 1995. — **24**, № 4. — С. 65—69.

Поступила 07.04.10;
после доработки 17.05.10

ПУСЬ Вячеслав Васильевич, д-р техн. наук, профессор кафедры экономики и менеджмента Санкт-Петербургского университета ГПС МЧС России. В 1962 г. окончил Ин-т авиационного приборостроения, в 1970 г. — Ленинградский госуниверситет. Область научных исследований — обработка информации, статистическое моделирование, математическая статистика.