ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2008 Управление, вычислительная техника и информатика № 1(2)

УДК 519.24

В.П. Шуленин, В.В. Табольжин

ИЗУЧЕНИЕ СВОЙСТВ РАНГОВЫХ АНАЛОГОВ F-КРИТЕРИЯ ФИШЕРА ПРИ ОТКЛОНЕНИЯХ ОТ ГАУССОВСКОЙ МОДЕЛИ ДИСПЕРСИОННОГО АНАЛИЗА

Проводится сравнение характеристик F-критерия Фишера, H-критерия Краскела – Уоллиса и L-критерия Пейджа в рамках различных супермоделей, описывающих отклонения от классической гауссовской модели дисперсионного анализа. Сравнение проводится как при конечных объемах выборки методом статистического моделирования, так и в асимптотике путем вычисления относительной эффективности Питмена.

Ключевые слова: *ранговые критерии, дисперсионный анализ, непараметрические модели.*

Пусть объекты изучаемой совокупности (или популяции) W характеризуются некоторым результирующим показателем X. В соответствии с факторным признаком A, который может принимать k значений $A_1,\ldots,A_k$, вся совокупность W разбивается на k групп $W_1,\ldots,W_k$ (или k подпопуляций $W_1,\ldots,W_k$ популяции W). Статистическими данными являются наблюденные реализации $x_{11},\ldots,x_{n_1 1},\ldots,x_{1k},\ldots,x_{n_k k}$ k выборок $X_{11},\ldots,X_{n_1 1},\ \ldots,\ X_{1k},\ldots,X_{n_k k}$ из совокупностей $W_1,\ldots,W_k$ с непрерывными распределениями изучаемого показателя X. Исходные данные кратко записываются в виде $\{X_{ij}\}$, $j = 1,\ldots,k$, $i=1,\ldots,n_j$, они получены в результате n_j наблюдений за результирующим показателем X при каждом фиксированном j-м уровне A_j, $j = 1,\ldots,k$, фактора A. Рассмотрим различные модели наблюдений.

1. Гауссовская модель

Предполагается, что исходные данные $\{X_{ij}\}$, $i = 1,\ldots,n_j$, $j = 1,\ldots,k$, представляют собой выборку, полученную в результате n независимых наблюдений над показателем X из k нормальных совокупностей $W_1,\ldots,W_k$ со средними значениями $\mu_1,\ldots,\mu_k$ и с равными, но неизвестными дисперсиями $\sigma_1^2=\sigma_2^2=\cdots=\sigma_k^2=\sigma^2$. Эту модель наблюдений называют *нормальной (или гауссовской) моделью* 1 *однофакторного дисперсионного анализа с фиксированными эффектами*. Для удобства дальнейших ссылок выделим в явном виде и пронумеруем все предположения этой модели наблюдений:

$$X_{ij}=\mu_j+\varepsilon_{ij}\,,\quad i=1,\ldots,n_j\,,\quad j=1,\ldots,k\,,\quad n=(n_1+\cdots+n_k)\,, \tag{1}$$

где

а) $\mu_j = M(X \mid A=A_j)$, $j=1,\ldots,k$, *постоянные* величины,

б) ε_{ij} – *независимые* случайные величины,

в) ε_{ij} – *нормальные* случайные величины, т.е. $L(\varepsilon_{ij})=N(0\,;\sigma^2)$,

г) дисперсии совокупностей $W_1,\ldots,W_k$ *равны неизвестному параметру* σ^2, то есть $\sigma_1^2=\sigma_2^2=\cdots=\sigma_k^2=\sigma^2$.

В рамках этой модели требуется убедиться в том, что изменение фактора A не влияет на итоговый показатель X. На статистическом языке эта задача сводится к проверке статистической однородности наблюдаемых данных $\{X_{ij}\}$, $i = 1,...,n_j$, $j = 1,...,k$, которая кратко записывается в виде проверки гипотез:

$$H_0 : \mu_1 = \mu_2 = ... = \mu_k = \mu\,,\ H_1 : \textit{не все}\ \mu_j\ \ \textit{равны}\,,\ j = 1,...,k\,. \tag{2}$$

Эти гипотезы проверяются с помощью F-критерия Фишера (см., например, [1]), основанного на статистике $F = S_B^2 / S_W^2$, где S_B^2 и S_W^2 средние квадраты соответственно между и внутри групп $W_1,\ldots,W_k$, вычисляемые по формулам

$$S_B^2 = SS_B /(k-1) = \frac{1}{k-1}\sum_{j=1}^{k} n_j (\overline{X}_{\bullet j} - \overline{X}_{\bullet\bullet})^2\,,$$

$$S_W^2 = SS_W /(n-k) = \frac{1}{n-k}\sum_{j=1}^{k}\sum_{i=1}^{n_j} (X_{ij} - \overline{X}_{\bullet j})^2\,.$$

Статистика $F = S_B^2 / S_W^2$ при гипотезе H_0 имеет F-распределение Фишера с числами степеней свободы $(k-1)$ и $(n-k)$, то есть справедливо выражение

$$L\{F = S_B^2 / S_W^2 \mid H_0\} = F(k-1\,, n-k)\,. \tag{3}$$

Критическая область размера α находится справа от квантиля $F_{1-\alpha}\,(k-1, n-k)$ уровня $(1-\alpha)$ для F-распределения с числами степеней свободы $(k-1)$ и $(n-k)$.

2. Непараметрическая модель с произвольными альтернативами

На практике предположения нормальности наблюдений не всегда могут быть обоснованы. В таких случаях рассматривают более общие модели наблюдений и предполагают, что $\{X_{ij}\}$, $i = 1,...,n_j$, $j = 1,...,k$, являются независимыми случайными величинами, которые одинаково распределены лишь при фиксированном j-м уровне A_j, $j = 1,...,k$, фактора A, то есть $X_{ij},...,X_{n_j j}$ является выборкой из условной функции распределения $F_j(x) = P\{X_{ij} \le x \mid A = A_j\}$, $j = 1,...,k$, $\forall i \in (1,...,n_j)$. Отметим, что $F_j(x)$ является произвольным непрерывным распределением, функциональный характер которого не конкретизируется и изучение влияния фактора A на итоговый показатель X в условиях этой непараметрической модели сводится к проверке гипотез

$$H_0^* : F_1 = F_2 = ... = F_k\,,\ \ H_1^* : \textit{не все}\ F_j\ \ \textit{равны}\,,\ \ j = 1,...,k\,. \tag{4}$$

Эти гипотезы проверяются с помощью H-критерия Краскела – Уоллиса (см., например, [2, 3]), статистика которого вычисляется не по исходным наблюдениям $\{X_{ij}\}$, а по их рангам $\{R_{ij}\}$, $i = 1,...,n_j$, $j = 1,...,k$, по формуле

$$H = \frac{12}{n(n+1)}\sum_{j=1}^{k} n_j \{\overline{R}_{\bullet j} - (n+1)/2\}^2\,, \tag{5}$$

где $\overline{R}_{\bullet j}$ – средний ранг наблюдений j-й группы, $j = 1,...,k$. При больших объемах выборки H-критерий определяется асимптотической критической областью размера α в виде неравенства $H \geq \chi^2_{1-\alpha}(k-1)$, где $\chi^2_{1-\alpha}(k-1)$ обозначает квантиль уровня (1 – α) для хи-квадрат распределения с числом степеней свободы k – 1.

3. Непараметрическая модель с упорядоченными альтернативами сдвига

Часто на практике уровни $A_1,\ldots,A_k$ фактора A отражают эффективность воздействия на показатель X в определенном направлении, например по мере увеличения интенсивности воздействия. В таких случаях рассматривают упорядоченные альтернативы. Предполагается, что $X_{ij},\ldots,X_{n_j j}$ – н.о.р. случайные величины с произвольной непрерывной функцией распределения $F(x-\theta_j)$, $j=1,\ldots,k$, $\forall i \in (1,\ldots,n_j)$. Для изучения влияния фактора A на итоговый показатель X в условиях этой непараметрической модели проверяются гипотезы

$$H_0^{**}: \theta_1 = \theta_2 = \ldots = \theta_k, \quad H_1^{**}: \theta_1 \leq \theta_2 \leq \ldots \leq \theta_k, \tag{6}$$

где хотя бы одно из неравенств строгое. Эти гипотезы также непараметрические, так как $F(x-\theta_j)$ – *произвольная непрерывная* функция распределения, и они проверяются с помощью L-критерия Пейджа (см., например, [2, 3]), статистика которого вычисляется также не по исходным наблюдениям $\{X_{ij}\}$, а по их рангам $\{R_{ij}\}$, $i=1,\ldots,n_j$, $j=1,\ldots,k$, по формуле

$$L = \frac{1}{nk}\sum_{j=1}^{k}\{j-(k+1)/2\}\{\overline{R}_{\bullet j}-(nk+1)/2\}. \tag{7}$$

При больших объемах выборки L-критерий Пейджа определяется асимптотической критической областью размера α в виде неравенства

$$L \geq \lambda_{1-\alpha}\{(k^2-1)(nk+1)/144n\}^{1/2},$$

где $\lambda_{1-\alpha} = \Phi^{-1}(1-\alpha)$ и Φ^{-1} обозначает квантильную функцию стандартного нормального распределения $\Phi(x)$.

4. Рассматриваемые типы супермоделей

Понятие «супермодель» (см., например, [4]) используют при изучении свойств робастности статистических процедур. Существуют различные подходы к заданию супермоделей. При изучении робастности процедур по распределению, один из вариантов задания супермодели состоит в конкретизации семейств распределений, включающих «идеальное» распределение, в которое мы верим и выбираем его в качестве основы, а также распределения, которыми могут характеризоваться наблюдения в условиях реального эксперимента. Мы рассмотрим два типа супермоделей, предложенных Тьюки [6].

Первый тип содержит λ-аппроксимацию стандартных симметричных распределений и задается в виде семейства распределений путем конкретизации их квантильных функций, то есть в виде

$$\Im_\lambda(F) = \{F : F^{-1}(u) = \lambda_1 + [u^{\lambda_3} - (1-u)^{\lambda_3}]/\lambda_2\}, \quad 0 \leq u \leq 1, \tag{8}$$

где λ_1 характеризует параметр положения, λ_2 является масштабным параметром и λ_3 – параметром формы распределения. Подходы к определению этих параметров описаны в [7]. В семействе распределений $\Im_\lambda(F)$ мы выделим супермодель $\Im_\lambda(\gamma_2)$, которая описывает отклонения от нормального распределения по эксцессу γ_2 при следующих значениях эксцесса: 1,75, 3, 4, 5, 9. Отметим, что для нормального распределения эксцесс γ_2=3. Второе семейство $\Im_\lambda(r)$ содержит λ-аппроксимацию распределений Стьюдента с числом степеней свободы r, принимающим следующие значения: 1, 5, 6, 7, 8, 9, 10, 25, 50, ∞. Отметим, что семейство распределений Стьюдента включает нормальное распределение ($r \to \infty$) и распределение Коши (r=1). Это семейство является удобным для описания широкого класса распределений, упорядоченных по степени «тяжести их хвостов» (см., например, [4]).

Второй тип супермоделей содержит гауссовские распределения с масштабным засорением и определяется в виде

$$\Im_{\varepsilon,\tau}(\Phi) = \{F : F_{\varepsilon,\tau}(x) = (1-\varepsilon)\Phi(x) + \varepsilon\Phi(x/\tau)\}, \quad 0 \le \varepsilon < 1/2, \quad \tau \ge 1. \tag{9}$$

Отметим, что при $\varepsilon = 0$, или при $\tau = 1$, имеем нормальное распределение $\Phi(x)$, $x \in R^1$.

5. Сравнение критериев при конечных объемах выборки

В рамках описанных типов супермоделей приведем результаты сравнения характеристик F-критерия Фишера, H-критерия Краскела – Уоллиса и L-критерия Пейджа. В качестве сравниваемых характеристик критериев используются их вероятности ошибок первого и второго рода. Изучение робастности F-критерия Фишера по уровню значимости при конечных объемах выборки проводится методом статистического моделирования, при этом исходные наблюдения $\{X_{ij}\}$ вычисляются по формуле

$$X_{ij} = \lambda_1 + [U_{ij}^{\lambda_3} - (1-U_{ij})^{\lambda_3}]/\lambda_2, \quad i = 1,...,n_j, \quad j = 1,...,k, \tag{10}$$

где U_{ij} случайные величины с равномерным распределением в интервале [0,1]. Отметим, что ранговые статистики H- и L-критериев имеют дискретные распределения, поэтому при сравнении критериев, которое проводилось при фиксированном уровне значимости α = 0,05, использовались асимптотические непрерывные аппроксимации их распределений при нулевой гипотезе. При этом в процессе моделирования проверялось качество этих аппроксимаций при различных объемах выборки путем построения оценок уровней значимости критериев по числу опытов M = 10 000. Отметим, что при моделировании использовались равные объемы выборок в группах $W_1,\ldots,W_k$, то есть $n_1 = n_2 = \ldots = n_k = n$. Мощности критериев сравнивались при альтернативах сдвига вида (6), при этом параметр положения λ_1 в (10) зависел от номера группы j и вычислялся по формуле $\lambda_1(j) = (j-1)\Delta$, $j = 1,...,k$, где $\Delta > 0$ – заданный параметр, характеризующий сдвиг распределений по группам $W_1,\ldots,W_k$. Результаты моделирования в виде оценок уровней значимости α критериев (при $\Delta = 0$) и оценок мощностей критериев $W(\Delta)$ при различных значениях параметра Δ, полученные по числу опытов M = 10 000, при числе групп k = 5, приведены в табл. 1 для $F \in \Im_\lambda(\gamma_2)$ и в табл. 2 для $F \in \Im_\lambda(r)$. Результаты эксперимента для $F \in \Im_{\varepsilon,\tau}(\Phi)$ приведены в табл. 3.

Т а б л и ц а 1. **Оценки уровня значимости и мощности F-, H- и L-критериев для модели Тьюки – отклонение от гаусса по эксцессу**

Объем выборки	Парам.	$\gamma_2 = 3$ $\lambda_2 = 0{,}1975$ $\lambda_3 = 0{,}1350$					$\gamma_2 = 4$ $\lambda_2 = 0{,}0262$ $\lambda_3 = 0{,}0148$					$\gamma_2 = 5$ $\lambda_2 = -0{,}0870$ $\lambda_3 = -0{,}0443$					$\gamma_2 = 9$ $\lambda_2 = -0{,}3203$ $\lambda_3 = -0{,}1359$					$\gamma_2 = 1{,}75$ $\lambda_2 = 0{,}5943$ $\lambda_3 = 1{,}4501$				
	Δ	0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60
$n = 5$	F	0,051	0,093	0,283	0,601	0,872	0,051	0,107	0,300	0,622	0,870	0,046	0,102	0,300	0,622	0,871	0,045	0,103	0,323	0,642	0,862	0,051	0,093	0,273	0,585	0,879
	H	0,038	0,069	0,221	0,517	0,813	0,038	0,081	0,257	0,564	0,828	0,036	0,083	0,269	0,587	0,847	0,036	0,091	0,317	0,651	0,864	0,033	0,068	0,194	0,459	0,771
	L	0,053	0,265	0,644	0,912	0,991	0,051	0,288	0,680	0,930	0,993	0,053	0,302	0,708	0,935	0,992	0,051	0,328	0,757	0,955	0,995	0,052	0,256	0,610	0,889	0,989
$n = 10$	F	0,051	0,169	0,614	0,949	0,998	0,049	0,179	0,613	0,943	0,998	0,051	0,177	0,621	0,943	0,996	0,048	0,184	0,639	0,941	0,992	0,053	0,165	0,603	0,950	0,999
	H	0,045	0,152	0,569	0,930	0,997	0,041	0,170	0,608	0,942	0,999	0,045	0,179	0,649	0,955	0,998	0,046	0,209	0,720	0,973	0,999	0,044	0,144	0,517	0,897	0,996
	L	0,053	0,424	0,895	0,996	1,000	0,052	0,457	0,922	0,998	1,000	0,051	0,479	0,935	0,999	1,000	0,050	0,544	0,954	0,999	1,000	0,053	0,410	0,871	0,995	1,000
$n = 20$	F	0,046	0,350	0,928	1,000	1,000	0,055	0,357	0,929	1,000	1,000	0,047	0,350	0,925	1,000	1,000	0,050	0,356	0,926	0,998	1,000	0,049	0,333	0,935	1,000	1,000
	H	0,045	0,324	0,914	0,999	1,000	0,051	0,366	0,941	1,000	1,000	0,044	0,386	0,954	1,000	1,000	0,048	0,443	0,977	1,000	1,000	0,045	0,306	0,892	0,999	1,000
	L	0,053	0,658	0,993	1,000	1,000	0,054	0,701	0,996	1,000	1,000	0,047	0,734	0,998	1,000	1,000	0,050	0,788	0,999	1,000	1,000	0,049	0,650	0,990	1,000	1,000

Т а б л и ц а 2. **Оценки уровня значимости и мощности F-, H- и L-критериев в условиях модели Тьюки – семейство распределений Стьюдента**

Объем выборки	Парам.	$r = 1$ $\lambda_2 = -3{,}0674$ $\lambda_3 = -1{,}000$					$r = 5$ $\lambda_2 = -0{,}2480$ $\lambda_3 = -0{,}1358$					$r = 9$ $\lambda_2 = -0{,}0003$ $\lambda_3 = -0{,}0002$					$r = 25$ $\lambda_2 = 0{,}1342$ $\lambda_3 = 0{,}0892$					$r = \infty$ $\lambda_2 = 0{,}1975$ $\lambda_3 = 0{,}1350$				
	Δ	0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60
$n = 5$	F	0,016	0,017	0,027	0,051	0,078	0,042	0,075	0,200	0,436	0,675	0,050	0,089	0,215	0,453	0,714	0,049	0,097	0,271	0,563	0,838	0,051	0,102	0,296	0,609	0,868
	H	0,034	0,050	0,102	0,188	0,285	0,036	0,066	0,191	0,418	0,478	0,039	0,072	0,184	0,404	0,669	0,035	0,071	0,216	0,487	0,782	0,036	0,076	0,233	0,521	0,808
	L	0,056	0,162	0,347	0,523	0,679	0,056	0,247	0,583	0,840	0,962	0,053	0,230	0,550	0,836	0,961	0,054	0,262	0,633	0,898	0,987	0,053	0,269	0,651	0,910	0,992
$n = 10$	F	0,015	0,024	0,032	0,048	0,082	0,044	0,126	0,412	0,767	0,845	0,049	0,130	0,450	0,818	0,975	0,054	0,168	0,564	0,922	0,996	0,053	0,175	0,615	0,946	0,999
	H	0,046	0,090	0,225	0,417	0,620	0,041	0,137	0,479	0,842	0,978	0,045	0,124	0,447	0,829	0,975	0,047	0,154	0,533	0,906	0,994	0,047	0,157	0,567	0,929	0,998
	L	0,051	0,242	0,539	0,787	0,914	0,052	0,377	0,842	0,984	1,000	0,052	0,366	0,807	0,980	1,000	0,054	0,417	0,886	0,993	1,000	0,052	0,427	0,898	0,996	1,000
$n = 20$	F	0,015	0,018	0,032	0,052	0,087	0,049	0,219	0,736	0,976	0,999	0,050	0,242	0,792	0,990	1,000	0,055	0,316	0,905	0,999	1,000	0,052	0,342	0,935	1,000	1,000
	H	0,046	0,137	0,476	0,784	0,937	0,046	0,272	0,847	0,997	1,000	0,047	0,255	0,815	0,993	1,000	0,051	0,302	0,895	0,999	1,000	0,047	0,318	0,916	1,000	1,000
	L	0,049	0,368	0,790	0,962	0,995	0,05	0,604	0,982	1,000	1,000	0,051	0,569	0,976	1,000	1,000	0,052	0,645	0,991	1,000	1,000	0,048	0,659	0,993	1,000	1,000

Таблица 3

Оценки уровня значимости и мощности F- и H-критериев для $F \in \Im_{\varepsilon,\tau}(\Phi)$, число групп $k = 5$, число опытов $M = 10000$

Объем выборки	Δ	ε = 0 , τ = 1					ε = 0,1 , τ = 3				
		0,00	0,15	0,30	0,45	0,60	0,00	0,15	0,30	0,45	0,60
$n = 20$	F	0,046	0,220	0,803	0,995	1,000	0,048	0,135	0,495	0,892	0,992
	H	0,045	0,203	0,775	0,993	1,000	0,047	0,169	0,639	0,971	1,000

Анализируя данные этих таблиц, можно сделать следующие выводы.

1. Эмпирический уровень значимости F-критерия обладает стабильностью при отклонениях от гауссовской модели по эксцессу в рамках супермодели $\Im_\lambda(\gamma_2)$ (см. табл.1). Однако F-критерий не обладает свойством робастности по уровню значимости в рамках супермодели $\Im_\lambda(r)$. В частности, для распределений с «тяжелыми хвостами» (см. табл.2 при $r = 1$), вместо заданного уровня $\alpha = 0{,}005$, эмпирический уровень значимости равен $\approx 0{,}016$. При увеличении числа степеней свободы r «затянутость хвостов» распределений начинает приближаться к гауссовской и эмпирические уровни начинают проявлять стабильность в окрестности заданного уровня.

2. Асимптотическая аппроксимация точного распределения ранговой статистики H-критерия Краскела – Уоллиса при нулевой гипотезе с помощью выражения $L(H \mid H_0^*) = \chi^2(k-1)$, является неудовлетворительной при малых объемах выборки. См., например, табл. 2 при $n = 5$ и любом числе степеней свободы, начиная с $r = 1$ и до $r \to \infty$. Вместо заданного уровня значимости α=0,005, эмпирический уровень значимости равен $\approx 0{,}03$. При увеличении объемов выборки качество аппроксимации улучшается, и при $n \geq 10$ она уже является удовлетворительной для целей практики. Этот вывод сохраняется и для супермодели, описывающей отклонения от гауссовской модели по эксцессу, то есть для $F \in \Im_\lambda(\gamma_2)$.

3. Для рассмотренных в эксперименте альтернатив и для гауссовской модели наблюдений вида (1), F-критерий имеет незначительное преимущество в мощности перед H-критерием. Однако при отклонениях от гауссовской модели, то есть в рамках супермоделей $\Im_\lambda(\lambda_2), \Im_\lambda(r)$ и $\Im_{\varepsilon,\tau}(\Phi)$, ситуация меняется. H-критерий имеет преимущество в мощности по сравнению с F-критерием, причем оно проявляется в большей степени при «утяжелении хвостов распределений» и при увеличении объемов выборки. Для рассмотренных в эксперименте упорядоченных альтернатив, L-критерий Пейджа, как и ожидалось, имеет существенно большую мощность по сравнению с F и H-критериями. Причем качество нормальной аппроксимации распределения ранговой статистики L при нулевой гипотезе вполне удовлетворительное и для малых объемов выборки, начиная с n = 5.

4. Проведенные эксперименты при числе групп $k = 10$, качественно не меняют эти выводы.

Отметим, что рассмотренные в предыдущих экспериментах супермодели $\Im_\lambda(\lambda_2)$, $\Im_\lambda(r)$ и $\Im_{\varepsilon,\tau}(\Phi)$, были использованы, в частности, для изучения робастности по распределению уровня значимости F-критерия. Эти супермодели описывают различные варианты отклонения от *предположения нормальности* (1в) гауссовской модели (1). Изучим теперь робастность уровня значимости F-критерия при отклонениях от предположения (1г) *о равенстве дисперсий в группах* $W_1,\ldots,W_k$, оставив все остальные предположения гауссовской модели (1) верными. Для этого исходные наблюдения $\{X_{ij}\}$ будем вычислять по формуле (10), в которой $\lambda_1 = 0$, что обеспечивает справедливость предположения нулевой гипотезы (2), то есть $H_0 : \mu_1 = \mu_2 = \ldots = \mu_k = \mu$. Далее, коэффициенты λ_2 и λ_3 соответственно будут равны $\lambda_2 = 0{,}1975$ и $\lambda_3 = 0{,}1350$, что обеспечивает выполнение предположения нормальности модели (1). Затем для нарушения предположения (1г) о равенстве дисперсий в группах $W_1,\ldots,W_k$, сделаем масштабный параметр λ_2 зависящим от номера группы j, то есть $\lambda_2(j) = j\lambda_2$, $j=1,\ldots,k$. В результате исходные наблюдения $\{X_{ij}\}$ вычисляются по формуле

$$X_{ij} = \lambda_1 + [U_{ij}^{\lambda_3} - (1-U_{ij})^{\lambda_3}]/\lambda_2(j)\,,\quad i=1,\ldots,n_j\,,\quad j=1,\ldots,k\,. \tag{11}$$

Результаты эксперимента приведены в табл. 4.

Таблица 4

Оценки уровня значимости F- и H-критериев в случае нарушения предположения о равенстве дисперсий

Тесты	Количество уровней и объемы выборок			
	$k = 5$, $n = 10$	$k = 10$, $n = 10$	$k = 5$, $n = 20$	$k = 10$, $n = 20$
F	0,102	0,137	0,095	0,136
H	0,063	0,067	0,067	0,072

Из табл. 4 видно, что при невыполнении предположения (1г) о равенстве дисперсий в гауссовской модели вида (1) уровень значимости F-критерия превышает заданный уровень $\alpha = 0{,}05$ больше, чем в два раза. Причем уровень значимости F-критерия значительно возрастает с увеличением количества уровней факторного признака A. Отметим, что условия рассматриваемого эксперимента для H-критерия соответствуют альтернативе H_1^*, так как дисперсии распределений в группах разные и, следовательно, не все F_j, $j = 1,\ldots,k$, равны. Приведенные данные для H-критерия превышают заданный уровень значимости $\alpha = 0{,}05$, что является проявлением свойства «несмещенности» H-критерия, так как эти данные характеризуют его мощность при рассмотренных альтернативах.

6. Асимптотическое сравнение критериев

В литературе разработаны различные подходы к асимптотическому сравнению критериев. Наиболее часто используют асимптотическую относительную эффективность Питмена (см. [2, 5]), которая вычисляется не для фиксированной альтернативы, а для последовательности контигуальных альтернатив, сходящихся к нулевой гипотезе при неограниченном увеличении объема выборки. Для многих непараметрических критериев получены общие выражения для эффективности Питмена по отношению к их «конкурентам» из нормальной теории. В частности, в [2] показано, что эффективность Питмена для H-критерия Краскела – Уоллиса относительно F-критерия Фишера вычисляется по формуле

$$ARE_F(H:F) = 12\sigma_f^2 \left[\int_{-\infty}^{\infty} f^2(x)dx \right]^2 = 12\sigma_f^2 \left[\int_0^1 f(F^{-1}(u)du) \right]^2, \tag{12}$$

где $\sigma_f^2 = D(X)$ и $f(x)$ – плотность функции распределения $F(x)$ наблюдений над показателем *X*. Отметим, что формула (12) имеет достаточно общий характер. По формуле (12) вычисляется также асимптотическая относительная эффективность Питмена для критерия знаковых рангов Уилкоксона и t-критерия Стьюдента в одновыборочном варианте и в двухвыборочном варианте для рангового критерия Уилкоксона и двухвыборочного t-критерия Стьюдента [5]. Это замечание распространяется на относительную эффективность Питмена многих непараметрических критериев по отношению к их «конкурентам» нормальной теории (см., например, [3]).

Отметим, что плотность распределения вероятностей, выражаемая через квантильную функцию $F^{-1}(u) = \lambda_1 + [u^{\lambda_3} - (1-u)^{\lambda_3}]/\lambda_2$, $0 \le u \le 1$, которая определяет элементы множества $\Im_\lambda(F)$ вида (8), записывается в виде

$$f(F^{-1}(u)) = 1/(F^{-1}(u))' = [\lambda_3\{u^{\lambda_3-1} + (1-u)^{\lambda_3-1}\}/\lambda_2]^{-1}, \; 0 \le u \le 1. \tag{13}$$

Далее, можно убедиться, что для $F \in \Im_\lambda$ все центральные моменты $\mu_k = M(X-\alpha)^k$ нечетного порядка равны нулю и, следовательно, коэффициент асимметрии $\gamma_1 = \mu_3/\mu_2^{3/2} = 0$, а коэффициент эксцесса $\gamma_2 = \mu_4/\mu_2^2$, вычисляется по формуле

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} = \frac{\{1/(4\lambda_3+1) - 4B(\lambda_3+1, 3\lambda_3+1)\}}{2[1/(2\lambda_3+1) - B(\lambda_3+1,\lambda_3+1)]^2} +$$

$$+ \frac{3B(2\lambda_3+1, 2\lambda_3+1)}{2[1/(2\lambda_3+1) - B(\lambda_3+1,\lambda_3+1)]^2}, \tag{14}$$

где $B(x,y)$ обозначает бета-функцию. Кроме того, выражение для дисперсии имеет вид

$$\sigma_f^2 = 2[1/(2\lambda_3+1) - B(\lambda_3+1,\lambda_3+1)]/\lambda_2^2. \tag{15}$$

С учетом формул (13) и (15), выражение (12) для $F \in \Im_\lambda$ запишется в виде

$$ARE_F(H:F) = 24[1/(2\lambda_3+1) - B(\lambda_3+1,\lambda_3+1)] \times$$

$$\times \left(\int_0^1 [\lambda_3\{u^{\lambda_3-1} + (1-u)^{\lambda_3-1}\}]^{-1} du \right)^2. \tag{16}$$

Численные расчеты показывают, что для $F \in \Im_\lambda(\lambda_2)$ асимптотическая относительная эффективность Питмена H-критерия относительно F-критерия при значениях эксцесса γ_2: 3, 4, 5, 9, 1,75 соответственно равна 0,954, 1,067, 1,167, 1,379, 1,066, а для семейства распределений Стьюдента $F \in \Im_\lambda(r)$ при числе степеней свободы *r*: 5, 7, 25 и $r \to \infty$ она соответственно равна 1,382, 1,162, 0,993, 0,954. Эти результаты на качественном уровне хорошо согласуются с результатами моделирования.

Рассмотрим теперь гауссовскую модель с масштабным засорением вида (9), то есть предполагаем, что $F \in \Im_{\varepsilon,\tau}(\Phi)$. Отметим, что распределения $F_{\varepsilon,\tau}(x)$ этого семейства характеризуются симметричными относительно нуля плотностями распределения вероятностей вида $f_{\varepsilon,\tau}(x) = (1-\varepsilon)\phi(x) + (\varepsilon/\tau)\phi(\varepsilon/\tau)$, где $\phi(x) = (2\pi)^{-1/2}\exp\{-x^2/2\}$, $-\infty < x < \infty$. Для $F \in \Im_{\varepsilon,\tau}(\Phi)$ имеем

$$\sigma_f^2 = \int_{-\infty}^{\infty} x^2 f_{\varepsilon,\tau}(x)dx = 1 + \varepsilon(\tau^2 - 1),$$

$$\int_{-\infty}^{\infty} f_{\varepsilon,\tau}^2(x)dx = \frac{(1-\varepsilon)^2}{2\sqrt{\pi}} + \frac{\sqrt{2}\varepsilon(1-\varepsilon)}{\sqrt{\pi(\tau^2+1)}} + \frac{\varepsilon^2}{2\tau\sqrt{\pi}}.$$

С учетом этих выражений, из (12) получаем, что асимптотическая относительная эффективность Питмена H-критерия относительно F-критерия для $F \in \Im_{\varepsilon,\tau}(\Phi)$ вычисляется по формуле

$$ARE_{F_{\varepsilon,\tau}}(H:F) = \{3[1+\varepsilon(\tau^2-1)]/\pi\}\times$$

$$\times\{(1-\varepsilon)^2 + 2\sqrt{2}\,\varepsilon(1-\varepsilon)/\sqrt{\tau^2+1} + \varepsilon^2/\tau\}^2. \qquad (17)$$

Численные значения асимптотической относительной эффективности Питмена H-критерия относительно F-критерия для гауссовской модели с масштабным засорением приведены в табл. 5.

Таблица 5

Эффективность Питмена $ARE_F(H:F)$ для $F \in \Im_{\varepsilon,\tau}(\Phi)$

τ	ε							
	0,00	0,01	0,03	0,05	0,08	0,10	0,15	0,20
3	0,955	1,009	1,108	1,196	1,309	1,373	1,497	1,575
5	0,955	1,150	1,505	1,814	2,201	2,412	2,795	3,006
7	0,955	1,369	2,115	2,759	3,553	3,977	4,724	5,099

Из приведенной таблицы следует, что H-критерий Краскела – Уоллиса, проигрывая лишь 5% в эффективности оптимальному при гауссовском распределении F-критерию Фишера, обладает существенными преимуществами даже при небольших, трудно обнаруживаемых, отклонениях от гауссовской модели. Или, другими словами, можно сказать, что F-критерий Фишера теряет оптимальность очень быстро при переходе от нормальной модели к модели из ее окрестности, содержащей распределения с «более тяжелыми хвостами».

Таким образом, подводя итог, можно сказать, что при возможных отклонениях от предположений гауссовской модели наблюдений вида (1) в условиях реального эксперимента, предпочтение в выборе критерия следует отдать ранговому критерию Краскела – Уоллиса (или критерию Пейджа при упорядоченных альтернативах), а не классическому F-критерию Фишера.

ЛИТЕРАТУРА

1. *Афифи А., Эйзен С.* Статистический анализ: Подход с использованием ЭВМ: Пер. с англ. М.: Мир, 1982.
2. *Хеттсманспергер Т.* Статистические выводы, основанные на рангах. М.: Финансы и статистика, 1987.
3. *Холлендер М., Вулф Д.* Непараметрические методы статистики. М.: Финансы и статистика, 1983.
4. *Шуленин В.П.* Введение в робастную статистику. Томск: Изд-во Том. ун-та, 1993.
5. *Кендэлл М., Стьюарт А.* Статистические выводы и связи. М.: Наука, 1973.
6. *Randles R.H., Wolf P.H.* Introduction to the Theory of Nonparametric Statistics. N.Y.: Wiley, 1979.
7. *Ramberg J.S.* An approximation method for generation symmetric random variables // Commun. ACM. 1972. V. 15. P. 987 – 990.

Статья представлена кафедрой теоретической кибернетики факультета прикладной математики и кибернетики Томского государственного университета, поступила в научную редакцию 17 сентября 2007 г.