УДК 519.233.32

Е.Т. Володарский, Е.В. Козыр

*Национальный технический университет Украины*
*«Киевский политехнический институт», Киев, Украина*

# ПРОВЕРКА ОДНОРОДНОСТИ ПО КРИТЕРИЮ ФИШЕРА КАК ДОПОЛНИТЕЛЬНЫЙ ИНСТРУМЕНТ КОНТРОЛЯ ПРИ ПРОВЕРКЕ УСЛОВИЙ ПОВТОРЯЕМОСТИ В ИССЛЕДУЕМЫХ ЛАБОРАТОРИЯХ

*Критерий Фишера показывает изучение при условии небольшого размера выборок. Проверка условий воспроизводимости, совместный с критерием Фишера использует как дополнительный инструмент к полученной надежной информации об однородности разниц. Условия воспроизводимости проверяют особенности в случае образцов с 2 элементами. Руководящие принципы относительно использования эти два инструмента контроля даются.*

***Ключевые слова:*** *критерий Фишера, условия воспроизводимости, маленькие образцы, однородность разниц, воспроизводимость среднего квадратичного отклонения, предел воспроизводимости.*

## Введение

Необходимость рассмотрения ***"прецизионности"*** возникает из-за того, что измерения, выполняемые на предположительно идентичных материалах при предположительно идентичных обстоятельствах, не дают, как правило, идентичных результатов. Это объясняется неизбежными случайными погрешностями, присущими каждой измерительной процедуре. При практической интерпретации данных измерений эта изменчивость должна учитываться. Например, нельзя установить фактическое отклонение полученного результата измерений от некоторого определенного значения измеряемой величины, если он лежит в области неизбежных случайных погрешностей измерительной процедуры. Аналогичным образом, сопоставление результатов измерений измеряемых характеристик двух партий материала не выявит какого-либо существенного различия в качестве, если расхождение между результатами лежит в вышеупомянутой области[1].

## Основной материал

Два условия прецизионности, называемые условиями ***повторяемости – repeatability*** (основная характеристика при проведении внутрилабораторных испытаний) и ***воспроизводимости – reproducibility*** (основная характеристика при проведении межлабораторных испытаний), были признаны необходимыми и, во многих практических случаях, полезными для представления изменчивости метода измерений.

Выполнение измерений должно быть организовано с соблюдением следующих требований:

– любая предварительная проверка оборудования должна соответствовать требованиям стандарта на метод измерений;

– каждая группа из $n$ измерений, относящихся к одному уровню, должна осуществляться при соблюдении ***условий повторяемости***, т.е. условия, при которых независимые результаты измерений (или испытаний) получаются одним и тем же методом на идентичных объектах испытаний, в одной и той же лаборатории, одним и тем же оператором, с использованием одного и того же оборудования, в пределах короткого промежутка времени [2];

– необходимо, чтобы группа из $n$ измерений в условиях повторяемости выполнялась независимым образом так, как если бы это были $n$ измерений на различных материалах.

Если испытания проводятся только одной лабораторией, то понятия повторяемость и прецизионность идентичны.

Стандартное (среднеквадратическое) отклонение повторяемости (сходимости) (***repeatability standard deviation***): стандартное (среднеквадратическое) отклонение результатов измерений (или испытаний), полученных в условиях повторяемости (сходимости). Эта норма является мерой рассеяния результатов измерений в условиях повторяемости. ***Предел повторяемости*** (сходимости) (***repeatability limit***): значение, которое с доверительной вероятностью 95% не превышается абсолютной величиной разности между результатами двух измерений (или испытаний), полученными в условиях повторяемости (сходимости). Используемое условное обозначение – $r$ .

С целью оценки точности (правильности и прецизионности) метода измерений целесообразно предположить, что каждый результат измерений, $y$ представляет собой сумму трех составляющих

$$y = m + B + e , \qquad (1)$$

где (для конкретного исследуемого материала) $m$ – общее среднее значение (математическое ожидание); $B$ – лабораторная составляющая систематической погрешности в условиях повторяемости; e – случайная составляющая погрешности каждого результата измерений в условиях повторяемости.

В пределах одной лаборатории дисперсия в условиях повторяемости носит название внутрилабораторной дисперсии и выражается следующим образом:

$$\mathrm{var}(e) = \sigma^2_W . \qquad (2)$$

Можно ожидать, что $\sigma^2_W$ будет иметь различные значения в разных лабораториях вследствие различий, например в квалификации операторов, однако в [2] подразумевается, что для стандартизованного соответствующим образом метода измерений такие различия между лабораториями будут невелики и что оправдано установление общего значения внутрилабораторной дисперсии для всех лабораторий, использующих данный метод. Это общее значение, которое оценивают средним арифметическим внутрилабораторных дисперсий, носит название ***дисперсии повторяемости*** и его обозначают следующим образом:

$$\sigma^2_r = \overline{\mathrm{var}(e)} = \overline{\sigma^2_W} . \qquad (3)$$

Данное среднее арифметическое берут по всем лабораториям, принимающим участие в эксперименте по оценке точности, которые остаются после исключения выбросов из числа всех дисперсий.

В технических условиях на продукцию может содержаться требование повторения измерений в условиях повторяемости. В этих обстоятельствах для проверки приемлемости результатов измерений и для того, чтобы решить, какое действие необходимо предпринять в том случае, если они неприемлемы, может быть использовано стандартное отклонение повторяемости.

Выполняя регулярные измерения на стандартных образцах, лаборатория может проверить стабильность своих результатов и получить, таким образом, доказательство для подтверждения своей компетентности в отношении, как систематической погрешности, так и повторяемости результатов своих измерений.

Квалификация лаборатории (профессиональный уровень) при отсутствии стандартного образца оценивается путем сравнения результатов с опорной лабораторией. Для проверки результатов на однородность и подтверждения возможности совместного оценивания в статистике применяется критерий Фишера. В то же время в [2] вводится один из показателей точности – ***повторяемость***, подтверждающий в данном случае одинаковую организацию и условия проведения исследований в лаборатории. При проведении испытаний с разрушением образцов или же испытаний дорогостоящих образцов часто приходится ограничиваться двумя результатами. Проведенный анализ показал особенности использования статистических критериев в этом случае.

Для этого проводился моделирующий эксперимент, отображающий поведение исследуемых критериев при сравнении двух серий измерений полученных в одной лаборатории, по два измерения в каждой серии.

Для результатов, полученных в одной и той же лаборатории, должны выполняться условия повторяемости, при обработке результатов измерений был введен дополнительный этап, а именно, проверка дисперсий исследуемых выборок по критерию Фишера (критерий Фишера использовался как наиболее простой инструмент проверки на однородность):

$$F = \frac{s_1^2}{s_2^2}; \quad s_1^2 > s_2^2; \qquad (4)$$

$$s_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(a_{ij} - \overline{a}_j)^2 , \qquad (5)$$

где $s_1^2$ и $s_2^2$ – расчетные значения среднеквадратических отклонений для сравниваемых выборок. По таблице критических значений Фишера определяют $F_{табл}$ для уровня значимости $\alpha$ и степеней свободы $m_1 = n_1 - 1$, $m_2 = n_2 - 1$, где $n_1$ – это объем выборки, обладающей большей дисперсией, а $n_2$ – объем выборки, обладающей меньшей дисперсией.

Приведем полученные аналитические зависимости. Есть результаты, полученные в одной лаборатории после проведения двух серий экспериментов по два измерения в каждом, соответственно - две выборки объемом по два элемента каждая:

$$A_1 = \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} \text{ и } A_2 = \begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix}.$$

Если базовый элемент ij содержит лишь два результата измерений, то внутриэлементное расхождение (аналог стандартного отклонения) равно:

$$s_{ij} = \frac{|y_{i1} - y_{i2}|}{\sqrt{2}}, \qquad (6)$$

где $y_{i1}, y_{i2}$ – результаты, полученные в ходе одного эксперимента.

Таким образом, если во всех базовых элементах содержится по два результата измерений, для простоты вместо стандартных отклонений могут быть использованы абсолютные расхождения.

После несложных математических преобразований получаем, что стандартные отклонения для этих выборок равны:

$$s_1 = \frac{|a_{11} - a_{12}|}{\sqrt{2}} \text{ и } s_2 = \frac{|a_{21} - a_{22}|}{\sqrt{2}} \text{ соответственно.}$$

Тогда $F = \frac{s_1^2}{s_2^2}$, при условии что $s_1^2 > s_2^2$. (7)

Перепишем это выражение в виде: $\sqrt{F} = \frac{s_1}{s_2}$ и рассмотрим 4 возможные варианта:

1. $\begin{cases} |a_{11} - a_{12}| > 1; \\ |a_{21} - a_{22}| > 1. \end{cases}$

При этом $(1 < |a_{21} - a_{22}| < |a_{11} - a_{12}|)$, тогда $\Rightarrow F \in (1; |a_{11} - a_{12}|)$.

2. $\begin{cases} |a_{11} - a_{12}| < 1; \\ |a_{21} - a_{22}| > 1. \end{cases}$

Так как не выполняется заданное нами условие (7) мы не рассматриваем это случай.

3. $\begin{cases} |a_{11} - a_{12}| > 1 \\ |a_{21} - a_{22}| < 1 \end{cases} \Rightarrow F \in (|a_{11} - a_{12}|; \infty)$.

4. $\begin{cases} |a_{11} - a_{12}| < 1 \\ |a_{21} - a_{22}| < 1 \end{cases} \Rightarrow F \in (1; \infty)$.

Как видим, в последних двух случаях есть возможность возникновение аномально больших значение критерия Фишера, если не ввести ограничения на порядок малости сравниваемых величин.

При $P = 0,95$ табличное значение критерия Фишера для объемов выборок $n = 2$ $F = 161,4476...$ Будем исходить от обратного и определим, при каких значениях дисперсий расчетное значения равняется критическому табличному значению. Рассмотрим последний случай (№ 4), когда обе разницы меньше единицы.

$$\left(F_{\text{критич.}} = 161,4476 < (a_{11} - a_{12})^2 / (a_{21} - a_{22})^2\right) \Rightarrow$$

$$\Rightarrow \begin{cases} |a_{11} - a_{12}| > \sqrt{161} \cdot |a_{21} - a_{22}|; \\ |a_{11} - a_{12}| < 1; \\ |a_{21} - a_{22}| < 1, \end{cases}$$

тогда если $|a_{11} - a_{12}| \in (0,1)$, то при $|a_{21} - a_{22}| \in (0; 0,0788)$ критерий Фишера будет считать эти значения выбросами, т.е. чем меньше расхождение между измеренными значениями в одной из лабораторий, тем больше шансов у этих результатов быть принятыми за выброс.

Рассмотрим случай № 3, когда:

$$\begin{cases} |a_{11} - a_{12}| > 1; \\ |a_{21} - a_{22}| < 1 \end{cases} \Rightarrow F \in (|a_{11} - a_{12}|; \infty)$$

опять же для $P = 0,95$, $F = 161,4476...$, ситуация будет аналогичной:

$$\begin{cases} |a_{11} - a_{12}| > \sqrt{161} \cdot |a_{11} - a_{12};| \\ |a_{11} - a_{12}| < 1; \\ |a_{21} - a_{22}| < 1, \end{cases}$$

только теперь если $|a_{11} - a_{12}| \in [1, \infty)$, то $|a_{21} - a_{22}| \in [0.0788, 1)$

Случаи 3 и 4 отображают ситуацию, когда при проверке на однородность двух сравниваемых выборок по критерию Фишера получаемый результат о неоднородности сравниваемых дисперсий не является правдивым.

Теперь рассмотрим два выше рассмотренных случая с точки зрения проверки данных полученных в одной и той же лаборатории на выполнение условий повторяемости.

Дисперсия повторяемости равна:

$$s^2_{rj} = \frac{\sum_{i=1}^{p} (n_{ij} - 1) s^2_{ij}}{\sum_{i=1}^{p} (n_{ij} - 1)}, \tag{8}$$

где $s_{ij}$ было определено выше по формуле (6); $n_{ij}$ – количество результатов измерений в базовом элементе (ячейке) для лаборатории $i$ на уровне $j$.

Напомним, что $n_{ij} = n = 2$. Подставив в выражение (8) выражения для $s_1$ и $s_2$ и, преобразовав его, получим следующее выражение для дисперсии повторяемости:

$$s^2_{rj} = \frac{s^2_1 + s^2_2}{2}. \tag{9}$$

Перепишем выражение (9), подставив вместо $s_1$ и $s_2$ соответствующие значения согласно (6). Получим:

$$s^2_{rj} = \frac{(a_{11} - a_{12})^2 + (a_{21} - a_{22})^2}{4}$$

или $s_{rj} = \frac{1}{2}\sqrt{(a_{11} - a_{12})^2 + (a_{21} - a_{22})^2}$.

Теперь рассчитаем предел повторяемости $r$, согласно [3] предел повторяемости рассчитывается как

$$r = f\sigma\sqrt{2}. \tag{10}$$

Величина $f$ (называемая коэффициентом критического диапазона) зависит от доверительного уровня вероятности и закона распределения случайной величины. Для пределов воспроизводимости и повторяемости доверительный уровень вероятности составляет 95%, и в ГОСТ Р ИСО 5725 делается допущение, что лежащее в основе распределение является приближенно нормальным. Для нормального распределения на уровне вероятности 95% коэффициент $f$ равен 1,96, и $f\sqrt{2}$ тогда равен 2,77. Получим следующее выражение для $r$:

$$r = 2,77\sigma = 1,39\sqrt{(a_{11} - a_{12})^2 + (a_{21} - a_{22})^2}.$$

Исходя из полученного выражения видно, что ни одна абсолютная разница вида $|a_{11} - a_{12}|$ или $|a_{21} - a_{22}|$ не превышает значения предела повторяемости $r$. Следовательно, условие повторяемости выполняется.

Однако следует заметить, что в данной ситуации условие повторяемости будет выполняться при любых значениях $a_{11} \dots a_{22}$ и следовательно даже при наличии выбросов или неоднородности результатов измерений. Поэтому возможно возникновение следующих двух ситуаций (рис. 1).

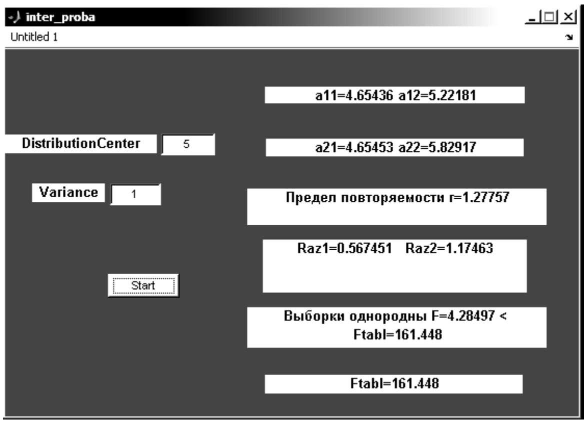


Рис. 1.Условия повторяемости выполняются, выборки однородны

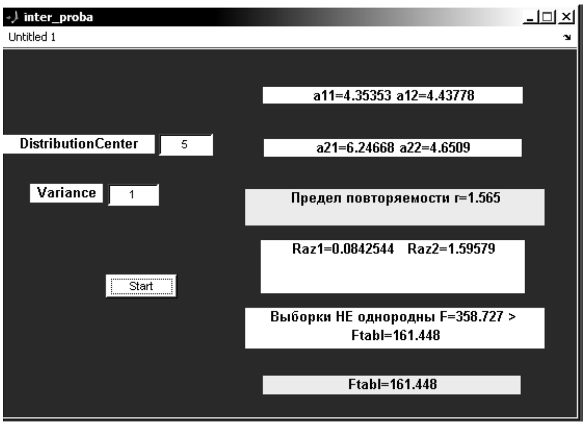


Рис. 2. Условия повторяемости выполняются, но выборки НЕ однородны

Моделирование проводилось для двух выборок подчиняющихся нормальному закону распределения с задаваемыми параметрами. В качестве основных параметров задаются центры распределения (***DistributionCenter***) и дисперсия (**Variance**), для конкретного случая были взяты параметры: центр – 5, дисперсия – 1.

$\mathrm{Raz1} = |a_{11} - a_{12}|$ и соответственно

$$\mathrm{Raz2} = |a_{21} - a_{22}|. \tag{11}$$

$F_{tabl}$ – табличное значение критерий Фишера при доверительном уровне вероятности $P = 0,95$. Также возможен и такой случай (рис. 2).

Как видно из рис. 2 значение $a_{21}$ отличается от всех остальных полученных значений и может являться выбросом, однако разность $|a_{21} - a_{22}|$ лежит в пределах $r$.

## Выводы

Использование критерия Фишера в совокупности с проверкой на выполнение условия повторяемости помогает установить неоднородность сравниваемых выборок, и дает основания для более глубокого изучения полученных результатов на предмет наличия выбросов в одной из выборок, например с помощью критерия Граббса.

## Список литературы

*1. ISO 3534-1:1993 Statistics-Vocabulary and symbols - Part 1: Statistical methods. Terms and definitions.*

*2. ГОСТ ИСО 5725-1-2002 Точность (прецизионность и правильность) результатов измерений.*

*3. ГОСТ ИСО 5725-6-2002 Точность (прецизионность и правильность) результатов измерений.*

*3. Володарский Е.Т., Малиновский Б.Н., Туз Ю.М. Планирование и организация измерительного эксперимента. – К.: Вища школа, 1987. – С. 79-85.*

*4. Иглин С.П. Математические расчеты на базе MATLAB. – С.-Пб.: БХВ-Петербург, 2005. – 640 с..*

*Поступила в редколлегию 7.05.2008*

**Рецензент:** д-р техн. наук, проф. И.В. Руженцев, Харьковский национальный университет радиоэлектроники, Харьков.

### ПЕРЕВІРКА ОДНОРІДНОСТІ ЗА КРИТЕРІЄМ ФІШЕРА ЯК ДОДАТКОВИЙ ІНСТРУМЕНТ КОНТРОЛЮ ПРИ ПЕРЕВІРЦІ УМОВ ПОВТОРЮВАНОСТІ В ДОСЛІДЖУВАНИХ ВИБІРКАХ

Володарський Є.Т., Козир О.В.

*Для оцінки повторюваності використовується границя повторюваності, значення, котре з довірчою вірогідністю 95% не перевищує абсолютної різниці між результатами двох вимірювань (чи випробувань), отриманих в умовах повторюваності (збіжності). Однак, при роботі з мінімальним об'ємом досліджуваних вибірок – дві порівнювані вибірки по дів вимірюванні у кожній рекомендується також використання критерію Фішера для перевірки однорідності дисперсій порівнюваних вибірок.*

***Ключові слова:*** *критерій Фішера, умови відтворюваності, маленькі зразки, однорідність різниць, відтворюваність середнього квадратичного відхилення, межа відтворюваності.*

### FISHER CRITERION HOMOGENEITY CHECK AS AN ADDITIONAL INSTRUMENT OF CONTROL IN TIME OF REPEATABILITY CONDITIONS CHECK

Volodarskyy E.T., Kozyr E.V.

*Repeatability limit use for repeatability estimation, a value which with confidence probability 95% is less than absolute value of difference between result of two measurements (or tests) which were received under repeatability conditions. However, when the small numbers of samples were used – two samples comparison with two results within criterion Fisher use are recommended to be able to check homogeneity of variances in compares samples.*

***Keywords:*** *fisher criterion, repeatability conditions, small samples, homogeneity of variances, repeatability standard deviation, repeatability limit.*