

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ

Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУРАЗИЯ ҰЛТТЫҚ УНИВЕРСИТЕТІ
КӨЛІК – ЭНЕРГЕТИКА ФАКУЛЬТЕТІ



***«КӨЛІК ЖӘНЕ ЭНЕРГЕТИКАНЫҢ ӨЗЕКТІ МӘСЕЛЕЛЕРІ:
ИННОВАЦИЯЛЫҚ ШЕШУ ТӘСІЛДЕРІ» XI ХАЛЫҚАРАЛЫҚ
ҒЫЛЫМИ-ТӘЖІРИБЕЛІК КОНФЕРЕНЦИЯСЫНЫҢ БАЯНДАМАЛАР
ЖИНАҒЫ***

***СБОРНИК МАТЕРИАЛОВ
XI МЕЖДУНАРОДНОЙ НАУЧНО – ПРАКТИЧЕСКОЙ
КОНФЕРЕНЦИИ: «АКТУАЛЬНЫЕ ПРОБЛЕМЫ ТРАНСПОРТА И
ЭНЕРГЕТИКИ: ПУТИ ИХ ИННОВАЦИОННОГО РЕШЕНИЯ»***

***PROCEEDINGS OF THE XI INTERNATIONAL SCIENTIFIC-PRACTICE
CONFERENCE «ACTUAL PROBLEMS OF TRANSPORT AND ENERGY:
THE WAYS OF ITS INNOVATIVE SOLUTIONS»***

Астана, 2023

УДК 656+620.9
ББК 39+31
А43

Редакционная коллегия:

Председатель – Курмангалиева Ж.Д. Член Правления – Проректор по науке, коммерциализации и интернационализации; Заместитель председателя – Кокаев У.Ш. декан транспортно-энергетического факультета, к.т.н., доцент; Султанов Т.Т. – заместитель декана по научной работе, к.т.н., доцент; Арпабеков М.И. – заведующий кафедрой «Организация перевозок, движения и эксплуатация транспорта», д.т.н., профессор; Тогизбаева Б.Б. – заведующий кафедрой «Транспорт, транспортная техника и технологии», д.т.н., профессор; Байхожаева Б.У. – заведующий кафедрой «Стандартизация, сертификация и метрология», д.т.н., профессор; Сакипов К.Е.– заведующий кафедрой «Теплоэнергетика», к.т.н., доцент; Жакишев Б.А.– заведующий кафедрой «Электроэнергетика», к.т.н., доцент.

А43 Актуальные проблемы транспорта и энергетики: пути их инновационного решения: XI Международная научно – практическая конференция, г. Астана, 16 марта 2023/Подгот. Ж.Д. Курмангалиева, У.Ш. Кокаев, Т.Т. Султанов – Астана, 2023. – 709с.

ISBN 978-601-337-844-2

В сборник включены материалы XI Международной научно – практической конференции на тему: «Актуальные проблемы транспорта и энергетики: пути их инновационного решения», проходившей в г. Астана 16 марта 2023 года.

Тематика статей и докладов участников конференции посвящена актуальным вопросам организации перевозок, движения и эксплуатации транспорта, стандартизации, метрологии и сертификации, транспорту, транспортной техники и технологии, теплоэнергетики и электроэнергетики.

Материалы конференции дают отражение научной деятельности ведущих ученых дальнего и ближнего зарубежья, Республики Казахстан и могут быть полезными для докторантов, магистрантов и студентов.



ДИСПЕРСИОННЫЙ АНАЛИЗ И РАСЧЕТ НЕОПРЕДЕЛЕННОСТИ "СВЕРХУ ВНИЗ"**Бектурганова Г.К.***gulmirabekt@yandex.kz*

Старший преподаватель Евразийского национального университета им. Л.Н. Гумилева

Джуманиязова С.Ж., Мерекеев А.Б., Каскабайулы Т.

Магистранты 1 курса Евразийского национального университета им. Л.Н. Гумилева

Обеспечение качества измерений является главной задачей метрологии. Статистические данные дают возможность исключить из расчета промахи и оставить только надежные показания. Дисперсионный анализ позволяет обеспечить как внутрилабораторный, так и внешний, межлабораторный контроль уровня измерений. Кроме того? Дисперсионный анализ позволяет оценить неопределенность измерений без рутинного подсчета вкладов всех влияющих величин.

Аббревиатура ANOVA (analysis of variation) обозначает дисперсионный анализ. Этот термин был предложен Рональдом Фишером в 1918 году. В 1925 году ученый написал книгу «Статистические методы для научных работников», где представил полное разложение общей суммы квадратов. Часто статистику, используемую в методе ANOVA, называют F-статистикой в его честь.

Дисперсионный анализ принимает допущение, что средние двух или более групп равны, если выборка осуществлена из совокупности, которая имеет нормальное распределение. Полная вариация в исследуемом наборе данных может быть разложена на несколько компонентов. Каждая из компонентов связана с определенным источником вариации. При этом сравниваются дисперсии (либо средние квадраты отклонений) внутри групп (или выборок) и между ними.

Сравнивают межгрупповую дисперсию и среднюю внутригрупповых дисперсий. Если первая больше, чем вторая, то средние таких выборок не будут равны. Если они примерно равны, то и между выборочными средними не будет существенной разницы.

Гипотезы в ANOVA и F-тест

Нулевая гипотеза допускает, что все средние взятых совокупностей или выборок или равны (альтернативной гипотезой является то, что хотя бы одна пара средних не равна). Для доказательства первой гипотезы для независимых выборок используется парный t-тест. Однако если необходимо сравнивать более 2 выборочных средних, то неоднократное использование парных t-тестов повысит вероятность появления ошибки I рода (при уровне значимости α), чем в случае парных сравнений. Поэтому чаще анализ ограничивается сравнением всех имеющихся выборочных средних путем проверки одной нулевой гипотезы, для этого используют F-тест. Использование F-критерия Фишера делает контроль за вероятностью появления ложных различий о эффективным.

ANOVA предполагает сравнение сумм квадратов отклонений:

- межгрупповых средних и общей средней (MSG);
- индивидуальных значений и внутригрупповых средних (MSE).

для каждого сочетания значений признаков. F-тест рассчитывается как соотношение средней суммы квадратов межгрупповых отклонений к средней сумме квадратов отклонений внутри групп:

$$F = \frac{MSG}{MSE}$$

Большое значение F-критерия (больше критического значения при заданном уровне значимости) рекомендует нам исключить нулевую гипотезу, а различия между группами признать более значимыми, чем внутригрупповые.

Виды процедур дисперсионного анализа

Однофакторный дисперсионный анализ исследует различия между средними по одному из факторов. Нулевая гипотеза состоит в том, что нет существенных различий между средними по группам, представляющим собой уровни одного и того же фактора. Альтернативная гипотеза заключается в том, что средние имеют существенные различия при данном уровне значимости.

Двухфакторный дисперсионный анализ представляет собой статистическую процедуру, в которой два фактора могут быть использованы для объяснения изменчивости зависимой переменной. Эти факторы закрепляются исследователем на разных уровнях. Возможно тестирование следующих нулевых гипотез:

1. Нет различий между групповыми средними по переменной А.
2. Нет различий между групповыми средними по переменной В. Альтернативная гипотеза для случаев 1 и 2: средние не равны.
3. Нет различий между групповыми средними по переменным А и В. Альтернативная гипотеза для случая 3: есть взаимодействие между А и В, средние существенно различаются.

N-факторный дисперсионный анализ рассматривает n-факторов одновременно, тестируя группы на различия между уровнями каждого фактора и эффект взаимодействия факторов.

Допущения в дисперсионном анализе

Метод предполагает использование трех статистических предпосылок:

1. Наблюдения из генеральных совокупностей отобраны методом случайного отбора.
2. Каждая выборка сформирована из нормально распределенной совокупности.
3. Генеральные совокупности, из которых организованы выборки, имеют равные дисперсии

ANOVA является эффективным инструментом даже в случае отклонений от нормальности в распределении генеральных совокупностей и наличия неравных дисперсий. Однако, когда неравенство дисперсий сопровождается разными объемами выборок, ситуация становится более сложной. В связи с этим, при планировании исследования предпочтение отдают методике, в которой выборки равны по числу наблюдений.

В межлабораторных сличениях достаточно большое количество лабораторий повторяет испытания на одном и том же образце.

Если образец испытывали разные лаборатории, дисперсия полученных значений характеристики отразит комбинацию дисперсии, которая может быть результатом внутрилабораторной и межлабораторной дисперсии.

Прецизионность (повторяемость и воспроизводимость) метода измерений может быть оценена на основе анализа данных группы лабораторий, которые используют один метод. Результаты лабораторий сравниваются между собой.

Вместо лабораторий могут быть группы аналитиков или операторов.

Статистическая обработка основана на следующих допущениях:

- несколько выборок результатов анализируются несколькими лабораториями;
- каждая лаборатория выполняет метод испытаний и фиксирует результаты повторений;
- на каждом уровне общее рассеивание результатов формируется за счет случайных изменений.

Этот процесс описан в ГОСТ ИСО 5725-2 «Точность (правильность и прецизионность) методов и результатов измерений» уравнением:

$$y = m + B + e,$$

где m - среднее (математическое ожидание) результатов;
 B - лабораторная составляющая смещения в условиях повторяемости;
 e - случайное отклонение, имеющее место при любом измерении в условиях повторяемости.

Наилучшая оценка истинного значения исследуемой характеристики - общее среднее всех результатов m . Единственное фиксируемое значение в общем случае не будет равно m .

Общая погрешность $y - m$ состоит из двух составляющих: e и B (e - отклонение от среднего из большого количества результатов, полученных в конкретной лаборатории, в которой значение y было измерено, а B - отклонение (или разность) между этим концептуальным, средним и общим средним). Это разбиение $y = m + B + e$ изображено на рисунке 1.

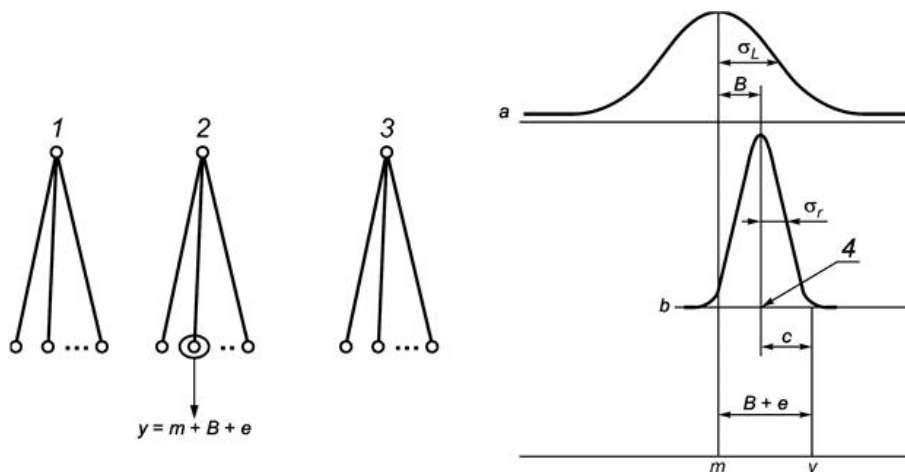


Рисунок 1 - Разбиение общей погрешности

- 1 - лаборатория № 1; 2 - лаборатория № 2; 3 - лаборатория № j;
- 4 - среднее лаборатории № 2;
- a - дисперсия лабораторной составляющей смещения;
- b - внутрилабораторная дисперсия результатов испытаний

Ожидается, что e и B близки к нулю. Целью ГОСТ ИСО 5725-2 является оценка их дисперсий. Дисперсия e - внутрилабораторная дисперсия σ_r^2 , а дисперсия B - межлабораторная дисперсия σ_L^2 . Внутрилабораторная дисперсия - дисперсия повторяемости.

Сумма межлабораторной дисперсии и внутрилабораторной дисперсии $\sigma_R^2 = \sigma_L^2 + \sigma_r^2$, представляет собой дисперсию результатов, полученных в условиях, которые значительно различаются (например, замена лабораторий), и является дисперсией воспроизводимости σ_R^2 . Структуру, изображенную на рисунке 1, называют вложенной или иерархической схемой. Для оценки дисперсий и необходимо оценить σ_L^2 и σ_r^2 .

Лабораторное воздействие с точки зрения статистики является случайным. Это означает, что выборка лабораторий, которые принимают участие в межлабораторной программе, также является случайной из большого числа лабораторий, которые могли бы теоретически использовать в работе данный метод испытаний. Эксперт оценивает, насколько эти лаборатории подобны друг другу.

Пример 1

Например возьмем 4 лаборатории, каждая из которых выполняет три параллельных определения содержания вещества в одном и том же образце. Потерянные данные и выбросы отсутствуют.

Таблица 1 – Данные измерений 4 лабораторий

№ измерения	1 лаборатория	2 лаборатория	3 лаборатория	4 лаборатория
1	15	16	13	15
2	16	13	15	14
3	17	15	15	16
дисперсия результатов σ_r^2	1,00	2,33	1,33	1,00
степени свободы	2	2	2	2

Имеется четыре выборочных средних (16,00, 14,67, 14,33 и 15,00) по одному для каждой лаборатории. Каждое среднее является оценкой измеряемого параметра. Общее среднее m равно 15,00, а дисперсия равна 0,52 и имеет три степени свободы. Поскольку в этом случае каждое лабораторное среднее является выборочным средним трех отдельных результатов, дисперсия этих средних не является оценкой только σ_L^2 , но содержит также часть дисперсии повторяемости (одна треть в этом примере).

Эту информацию представляют в виде таблицы 2.

Таблица 2 - Оценка компонентов дисперсии

Оцениваемая величина	Значение оценки
$\sigma_L^2 + \sigma_r^2/3$	0,52
σ_r^2	1,42

Приравнивая оценки к соответствующим оцениваемым величинам, можно получать оценки для σ_r^2 и σ_L^2 .

Здесь $s_r^2 = 1,42$;

$s_L^2 = 0,05$ (т.е. $0,52 - 1,42/3$).

Оценка s_R^2 дисперсии s_r^2 и s_L^2 повторяемости определяется как $s_r^2 + s_L^2$ и $s_R^2 = 1,47$ (т.е. $0,05 + 1,42$). Результаты представлены в виде таблицы 3.

Таблица 3 - Оценка компонентов дисперсии

Оцениваемая величина	Значение оценки
$3\sigma_L^2 + \sigma_r^2$	1,56 (=3x0,52)
σ_r^2	1,42

Вычисления усложняются, когда количество результатов, представленных лабораториями, неодинаково, в этом случае рекомендуется использовать программное обеспечение. Однако принцип остается тем же самым.

Пример 2

В таблице приведены необработанные данные, соответствующие средние и дисперсии для различных лабораторий.

Таблица 4 - Средние и дисперсии

Параметр	Лаборатория 1	Лаборатория 2	Лаборатория 3	Лаборатория 4
Результаты	63	44	50	53
	57	51	40	57
	54	43	42	46
Среднее	58	46	44	52
$s_{r,i}^2$	21	19	28	31
Число степеней свободы	2	2	2	2

Примечание – Индекс i указывает i -ю лабораторию.

Из этих результатов, следует:

- общее среднее m 50,00, т. е. $(58+46+44+52)/4$;
 - дисперсия средних 40,00;
 - дисперсия повторяемости s_r^2 24,75, т. е. $(21+19+28+31)/4$.
- Следовательно,
- межлабораторная дисперсия s_L^2 31,75, т.е. $40-24,75/3$;
 - дисперсия воспроизводимости $s_R^2 = s_L^2 + s_r^2$ 56,50, т. е. $31,75+24,75$.

Таким образом, дисперсионный анализ позволяет количественно оценить уровень измерений как в отдельно взятой лаборатории, так и в целом по данному виду измерений. Дальнейший анализ позволяет исключать выбросы из расчетов, находить лабораторное смещение и смещение метода, оценивать правильность используемых методик и т.д.

Расчет неопределенности «сверху вниз»

Дисперсионный анализ позволяет оценить неопределенность измерений методом «сверху вниз».

Пример 3

4 дня 4 оператора измеряли концентрацию стандартного образца (СО), имеющего аттестованное значение $12,6 \pm 0,9$ мг/кг (коэффициент охвата $k=2$).

Таблица 5 – Результаты измерений концентрации СО, мг/кг

	1-й день	2-й день	3-й день	4-й день
Оператор 1	12,27	11,97	12,02	12,71
Оператор 2	11,68	12,46	12,27	12,86
Оператор 3	12,41	12,27	11,92	12,56
Оператор 4	11,78	12,46	12,22	12,41

С помощью таблиц Excel вычисляем дисперсии:

Однофакторный дисперсионный анализ						
ИТОГИ						
Группы	Счет	Сумма	Среднее	Дисперсия		
Столбец 1	4	48,14	12,035	0,128966667		
Столбец 2	4	49,16	12,29	0,053533333		
Столбец 3	4	48,43	12,1075	0,027291667		
Столбец 4	4	50,54	12,635	0,0375		
Дисперсионный анализ						
Источник вариации	SS	df	MS	F	P-Значение	F критическое
Между группами	0,860869	3	0,286956	4,641583825	0,022378856	3,490294821
Внутри групп	0,741875	12	0,061823			
Итого	1,602744	15				

В таблице значение дисперсии обозначены символом MS.

Дисперсия внутри групп равна: $s_r^2=0,06$ мг/кг;

Межлабораторная дисперсия (между группами) равна: $s_L^2=0,29$ мг/кг.

Дисперсия воспроизводимости равна:

$$s_R^2 = \sqrt{s_L^2} = 0.25 \text{ мг/кг};$$

Дисперсия промежуточной прецизионности равна:

$$s_{ip} = \sqrt{\frac{s_L^2 - s_r^2}{n_{rep}}}$$

Здесь n_{rep} – количество независимых определений в 1 день;

Суммарная стандартная неопределенность измерений рассчитывается следующим образом:

$$u_m = \sqrt{\frac{s_R^2}{N_{rep}} + \frac{s_{ip}^2}{n_{days}} + u_{cal}} = 0,14 \text{ мг/кг}.$$

Здесь N_{rep} – общее количество независимых определений, n_{days} – количество дней.

Неопределенностью калибровки в данном случае пренебрегаем. Чтобы найти расширенную неопределенность, суммарную стандартную неопределенность умножаем на коэффициент охвата.

Таким образом, дисперсионный анализ позволяет оценить качество измерений, а также рассчитать неопределенность измерений методом «сверху вниз»

Список использованных источников

1. Ниязбекова Р.Қ., Михалченко В.Н., Бектурганова Г.К., Донбаева В.А. Тоғызбаев К.У., Байбосынова Л.А. Метрология. Алматы: Дәуір, 2012. 296 с.
2. ГОСТ ИСО 5725-2 «Точность (правильность и прецизионность) методов и результатов измерений».