

УДК 536.7+ 519.2

МЕТОДИКА ДИСПЕРСИОННОГО АНАЛИЗА НЕРАВНОТОЧНЫХ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

Е. А. Солдатов

(кафедра радиохимии)

Разработана новая методика и выведена соответствующая формула дисперсионного анализа неравноточных экспериментальных данных для выявления влияния качественных факторов на результаты эксперимента. Методика не требует ни одинаковой точности, ни одинакового числа данных по каждому фактору. Методика апробирована на литературных данных по стандартной энтальпии испарения меди.

Цель настоящей работы – расширение возможностей дисперсионного анализа для выявления влияния качественных факторов эксперимента в том случае, когда проводится анализ значений, погрешности которых значимо отличаются друг от друга.

Напомним общепринятую методику дисперсионного анализа [1]. Все n данных x_{ij} делятся на группы по выбранному фактору с последующим сопоставлением дисперсий между группами s_1^2 и внутри групп s_2^2 . Группы имеют номера i от 1 до m . В каждой группе номера данных j меняются от 1 до k_i (пределы суммирования). Фактор считается влияющим, если табличное значение критерия Фишера при 5%-м уровне значимости меньше рассчитанного по использованным данным

$$F = s_1^2 / s_2^2 \quad (1)$$

с числом степеней свободы соответственно $m - 1$ и $n - m$. Из таблицы, приведенной в [1], следует, что если x_i – среднее в i -й группе, а x – общее среднее, то для дисперсий получаются формулы:

$$s_1^2 = \sum_i k_i (x_i - x)^2 / (m - 1), \quad (2)$$

$$s_2^2 = \sum_i \sum_j (x_{ij} - x_i)^2 / (n - m). \quad (3)$$

Здесь и далее, следуя авторам [1, 2], будем использовать упрощенные обозначения сумм, тем более, что пределы суммирования указаны выше. Раскрывая равенства (2) и (3), получаем расчетные формулы:

$$s_1^2 = [\sum_i (\sum_j x_{ij})^2 / k_i - (\sum_i \sum_j x_{ij})^2 / n] / (m - 1), \quad (4)$$

$$s_2^2 = \sum_i [\sum_j x_{ij}^2 - (\sum_j x_{ij})^2 / k_i] / (n - m). \quad (5)$$

Общепринятая методика, как подчеркивает Дерфель [1], требует: «внутри этих групп случайные ошибки должны быть одинаковой величины», иначе «результаты надо объединять в группы с одинаковой воспроизводимостью», т.е. фактически разбивать группы, и так содержащие небольшое число данных, на еще более мелкие подгруппы.

Однако нашлась возможность снять требование равноточности данных. В [2] доказано, что если p_i – статистический вес данного x_i , а p – сумма весов, то для взвешенного среднего значения

$$x = \sum_i p_i x_i / p \quad (6)$$

получается необходимое нормальное распределение. При этом дисперсия взвешенного измерения, в соответствии с правилом Бесселя, имеет вид

$$s^2 = \sum_i p_i (x_i - x)^2 / f, \quad (7)$$

где f – число степеней свободы.

Применим этот подход к общепринятой методике. Для этого достаточно ввести веса данных, проще всего – обратные погрешностям

$$p_{ij} = 1 / s_{ij}. \quad (8)$$

Тогда формулы (2) и (3) приобретают другой вид. Пусть p_i – сумма весов данных в i -й группе и p – сумма всех весов. Тогда в группах взвешенное среднее

$$x_i = \sum_j p_{ij} x_{ij} / p_i \quad (9)$$

и общее взвешенное среднее

$$x = \sum_i \sum_j p_{ij} x_{ij} / p. \quad (10)$$

Теперь можно преобразовать формулы (2) и (3) для дисперсий между группами и внутри групп:

$$s_1^2 = \sum_j p_i (x_i - x)^2 / p (m - 1), \quad (11)$$

$$s_2^2 = \sum_i \sum_j p_{ij} (x_{ij} - x_i)^2 / p_i (n - m). \quad (12)$$

Раскрывая значения x_i и x в этих формулах, после алгебраических преобразований получаем необходимые расчетные формулы для дисперсий:

$$s_i^2 = [\sum_i (\sum_j p_{ij} x_{ij})^2 / p_i - (\sum_i \sum_j p_{ij} x_{ij})^2 / p] / (m - 1), \quad (13)$$

$$s_i^2 = \sum_i [\sum_j p_{ij} x_{ij}^2 - (\sum_j p_{ij} x_{ij})^2 / p_i] / (n - m). \quad (14)$$

По этим уравнениям не составляет труда рассчитать критерий Фишера (1). Кстати, эти уравнения легко превращаются в (4) и (5), если веса приравнять единице. Составленная для персонального компьютера программа рассчитывала не только экспериментальные и табличные величины критерия Фишера, но и средние взвешенные значения с 95%-й погрешностью для всех возможных комбинаций групп данных внутри фактора.

Апробацию методики провели на 49 литературных данных по стандартной энтальпии испарения меди при 0 К,

Экспериментальные стандартные энтальпии испарения меди

Фактор	Группа	k_i	$F_{рас}$	$\Delta H^0_{\text{в}}$, Дж/моль
Метод измерения	мас-спектр.	5	1,14	338,9±2,2
	Лангмюр	11		335,6±2,1
	прочие	6		334,6±0,7
	Кнудсен	27		333,2±1,8
Материал камеры	Мо или W	12	8,94	338,9 ± 1,5
	другие металлы	13		335,2 ± 0,2
	алунд	13		333,9 ± 1,4
	графит	11		332,2 ± 2,6
Годы публикации	1923–1961	11	9,82	338,1 ± 1,3
	1964–1971	11		335,2 ± 1,9
	1974–1980	12		335,8 ± 1,0
	1982–1987	15		332,2 ± 1,0

в свое время любезно предложенных В.И. Севериным [3]. Проверка критерием Бартлета показала значимую неоднородность дисперсий, что обуславливает необходимость применения разработанной методики. В качестве весов измерений использовали обратные 95%-е погреш-

ности. Итоги расчетов приведены в таблице при числах степеней свободы $f_1 = 3$, $f_2 = 45$ и табличном значении критерия Фишера 2,81. Указаны факторы, по которым проводили деление на группы, числа данных в группах k_i , расчетные значения критерия Фишера и значения стандартной энтальпии в кДж/моль с 95%-м доверительным интервалом.

Наблюдается тенденция уменьшения значений энтальпии с течением времени, т.е. с совершенствованием методик испарения. Наименьшие значения получаются также при испарении из графитовых камер Кнудсена. По критерию Фишера влияние метода не очевидно, но при исключении метода Кнудсена значение расчетного критерия составляет 4,85, что больше табличного (3,52). Эти данные количественно подтверждают качественные заключения работы [4] о наибольшей достоверности данных, полученных методом Кнудсена при испарении меди из сверхчистого графита, и показывают большие возможности новой методики. Например, можно экстраполировать на бесконечность зависимость от даты публикации. За ноль отсчета можно принять год первой публикации, а затем точками обозначать обратную величину разности последнего года публикации в группе и года первой публикации. Тогда значение стандартной энтальпии испарения, полученное с помощью линейной экстраполяции, составит 225 кДж/моль (предполагаемая погрешность 3 кДж/моль).

СПИСОК ЛИТЕРАТУРЫ

1. Дёрфель К. Статистика в аналитической химии. М., 1994. С. 139.
2. Агекян Т.А. Теория ошибок для астрономов и физиков. М., 1968. С. 134.
3. Северин В.И. // Дис. ... канд. хим. наук. М., 1982. С. 86.
4. Северин В.И. и др. // Теплофизика высоких температур. 1986. 24. С. 492.

Поступила в редакцию 20.05.00