Математические методы и информационные технологии в химии и химической технологии

УДК 681.3.06

ПРОГРАММА ДЛЯ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

Р.А. Осипов, студент, *А.Ю. Скрябина, студент, *Н.С. Рукк, доцент кафедра Высшей и прикладной математики *кафедра Неорганической химии им. А. Н. Реформатского МИТХТ им. М.В. Ломоносова

e-mail: <u>neorgan@mitht.ru</u>

а языке VBA Excel разработана программа СОРЕ для проведения статистической обработки результатов эксперимента, позволяющая значительно упростить и ускорить рутинную обработку экспериментальных данных.

Ключевые слова: программа, статистическая обработка, эксперимент, выборка, MS Excel, VBA.

С проблемой статистической обработки результатов эксперимента по измерению приходится физико-химических величин часто сталкиваться при организации как научного, так и учебного процесса. Для сокращения временных и трудозатрат можно использовать различные специализированные компьютерные программы, полные версии которых либо лицензионные и, следобесплатные, либо вательно, не Нелицензионные (со всеми вытекающими отсюда последствиями). В связи с этим на базе VBA Excel [1, 2] была разработана очень простая в использовании программа, апробированная на кафедре неорганической химии при результатов обработке измерений температуры плавления иодидов гексаантипиринлантаноидов $[Ln(A\Pi)_6]I_3$ (Ln = La, Ce-Nd, Sm-Lu, Sc, Y; АП – антипирин). Для конкретности, далее речь пойдет об обработке именно таких данных, хотя это нисколько не мешает использовать программу и для интерпретации результатов других физико-химических исследований.

В результате проведения измерения по температуры определению плавления данного вещества получается один элемент выборки. Всего для каждого вещества проводилось $5 \le n \le 20$ измерений, таким необходимо провести образом, статистическую обработку результатов эксперимента для малой выборки (важно отметить, что разработанная программа способна обрабатывать значительно большее число измерений, порядка 10⁴). Согласно следствию центральной предельной теоремы ИЗ считалось, что распределение случайной величины (результат произведенного измерения), Гауссово.

В соответствии с требованиями иссле-

дований подобного рода, статистическая обработка результатов содержит в себе следующие этапы [3–6]:

1) определение грубых ошибок (промахов) и их последующий отсев (в программе реализованы критерии Романовского $K_R(P,n)$ и 3σ);

2) определение основных статистик и дополнительных некоторых численных характеристик выборки (программа позволяет найти разброс R, выборочное среднее значение \overline{X} , выборочную дисперсию S_{y}^{2} , выборочное стандартное отклонение S_{y} , оценку точности Т, коэффициент вариации (относительное V, воспроизводимость стандартное отклонение) результатов S_и, доверительный интервал, относительную случайную ошибку Δ_r , относительную систематическую ошибку $\Delta_{0,r}$);

3) определение количества n_{ess} дополнительных измерений при заданном уровне точности и вариации, которое необходимо провести для того, чтобы с точностью p описать генеральную совокупность при данном значении коэффициента вариации V.

4) проверка различия средних значений с помощью критерия Стьюдента, если есть несколько выборок, которые представляется естественным объединить в одну (например, результаты измерений, проведенных в разные дни).

Математическая обработка результатов эксперимента, содержащая в себе эти четыре этапа, реализована в созданной программе СОРЕ, которая представляет собой группу макросов, написанных на языке Microsoft Visual Basic 6.3 на базе приложения MS Excel.

Продемонстрируем работу с ней на примере: конкретном пусть требуется обработать результаты опыта по определению температуры плавления комплекса [Sm(AП)₆]I₃. Числовые значения элементов выборки ИЗ восьми измерений ланной случайной величины, полученные на установке ПТП(М) таковы: 263.4; 263.5; 263.5; 263.4; 263.3; 263.6 (в градусах Цельсия).

Шаг 1

Открыв файл COPE.xls приложения MS Excel на листе «Расчет» начиная с ячейки A1 в столбик вносим данные, так, чтобы не было пустых ячеек между любыми двумя числовыми значениями. После этого нажимаем кнопку с названием «COPE» (рис. 1).



Рис. 1. Шаг 1: ввод данных и запуск программы СОРЕ.

Шаг 2

После нажатия кнопки «СОРЕ» появится окно «СОРЕ, вопрос о первичной обработке данных» (рис. 2.). Сейчас оно не нужно и поле остается пустым (о его применении речь пойдет ниже).



Рис. 2. Шаг 2: работа с окном: «СОРЕ, вопрос о первичной обработке данных».

Шаг З

После того, как работа на втором шаге завершена, появится окно «Обработка результатов эксперимента с малой выборкой» (рис. 3). В нем необходимо заполнить следующие поля:

1) название рабочего объекта — вручную вводится любая информация, не содержащая указанных в заголовке поля символов;

2) дата проведения эксперимента вручную или с помощью ниспадающего меню вводится дата проведения (обсчета или др.) эксперимента;

 рабочая вероятность — из ниспадающего меню выбирается одно из 4-х значений для доверительной вероятности: 0,9; 0.95; 0.98; 0.99;



Рис. 3. Шаг 3: работа с окном: «Обработка результатов эксперимента с малой выборкой».

4) информация о добавлении результатов в базу данных — если есть база данных,

созданная на основе MS Excel и совместимая с СОРЕ, то можно добавить в нее полученные результаты статистической обработки результатов эксперимента, выбрав «добавлять результаты в базу данных» (при этом появится стандартное окно, в котором можно выбрать файл базы данных), иначе выбирается «не добавлять результаты в базу данных»;

5) информация о выборе листа, на который будут записаны результаты в базе данных – если в поле из пункта 4) выбрано «да», то в ниспадающем меню можно увидеть все листы выбранной базы данных, остается (только) выбрать нужный. Если в поле 4) выбрано «нет», то это поле можно не заполнять, его содержание ни на что не повлияет;

6) поле с вопросом о закрытии базы данных после добавления в нее результатов статистической обработки;

7) поле, содержащее информацию о добавлении результатов в отдельный файл – если нужно создать отдельный файл с результатами работы программы над текущими данными, то в ниспадающем меню выбирается «добавлять результаты в файл» (при этом появляется стандартное окно Windows, в котором можно выбрать адрес, по которому создастся файл), иначе выбирается «не добавлять результаты в файл». При этом важно отметить, что файл будет иметь название, соответствующее записи в поле 1);

8) информация о закрытии созданного в соответствии с полем 7) файла после добавления в него результатов статистической обработки;

9) информация о закрытии программы СОРЕ (текущего файла MS Excel). Важно отметить, что если в полях 4) и 7) одновременно выбраны отрицательные ответы, то результаты не будут сохраняться в текущем файле, так как программа СОРЕ после завершения своей работы полностью очищает рабочие листы книги COPE.xls для последующей работы. После того, как заполнение полей завершено, окно закрывается (нажатием на крестик в его верхнем правом углу).

Шаг 4

Если какое-то из полей окна «Обработка результатов эксперимента с малой выборкой», (шаг 3) заполнено неверно, то программа выдаст сообщение об ошибке (рис. 4) и предложит методы решения создавшейся проблемы.

ы не корректно ввели значение рабочей ероятности. Введите одно из значений: 0,9;	ОК
95; 0,98; 0,99.	Cancel
РЕ, КРИТИЧЕСКАЯ ОШИБКА 5: нет данных о за	аписи
ы не назвали адрес, по которому находится аза данных, в которую нужно записать	ОК
езультаты. Введите его, или напишите 1, если обавлять результаты в базу данных не нужно.	Cancel
РЕ, КРИТИЧЕСКАЯ ОШИБКА 4: рабочий объек	т не н
ы не ввели название рабочего объекта. ведите его, с учетом того, что знаки: / \ : * ? <	ОК
запрещены!)	Cancel

Рис. 4. Примеры окон с критическими ошибками программы СОРЕ.

Шаг 5

После завершения шага 4 в файле и/или соответствующей базе данных будут представлены результаты статистической обработки, произведенные программой СОРЕ (рис. 5).

	A	В	C	D	E
1	263.5	Название рабочего объекта	(Sm(АП)6)(3 температура плавления	Количество измерений	4
2	263,5	Дата проведения эксперимента	Дата: 23.11.2008; время: 11:30:28	Количество дополнительных измерений	0
3	263,4	Рабочая вероятность, Р	0,95	Промахи с вероятностью Р=0,95	
4	263,4	Показатель точности, t	0,05	263,3	
5		Результаты статистической образотки резул	263,6		
6		Выборочное среднее значение Хср	263,4500		
7		Выборочная дисперсия S2Y	0,0033		
8		Выборочное стандартное отклонение	0,0577		
9		Максимум	263,5000		
10		Минимум	263,4000		
11		Разброс	0,1000		
12		Коэффициент вариации V, %	0,0219		
13		Точность	0,0500		
14		Воспроизводимость, Sr	0,0002		
15		95% доверительный интервал	263,45±9,18693977352016E-02		
16		Относительная ошибка, %	0,0349		
17					
18		COPF			
19		COFE			
20					_

Рис. 5. Результаты обработки данных, произведенных программой СОРЕ.

На этом статистическая обработка данных закончена. Остается только откорректировать результат (на рис. 5 это ячейки C6:C16).

Если статистическую обработку нужно произвести над данными, являющимися некоторыми производными исходных данных, то последние необходимо вводить на листе «Данные для предв. обработки». Поясним, что понимается под исходными данными на двух простейших примерах:

	Α	В	С	D	E	F			
1	Проверка различия средних значений								
~	Проверка различия средних значений с помощью								
2	критерия Стьюдента								
3	Вещество								
5									
6	- [Sm(AП) ₆]l ₃ первый день [Sm(AП) ₆]l ₃ второй де								
7									
8	Количество опытов								
9	6			8					
10		0		0					
11	Среднее значение определяемой величины								
12									
14	263,45			263,52					
15									
16	Значение дисперсии								
17	0.06		0.4						
18		-,		-,.					
19	Средневзвешанное значение дисперсии								
21									
22	0,307950212								
23	B		TL		0 00				
24	вероятность 0,99								
25	Экспериментальное значение критерия								
20	Стьюдента								
21	2,9160								
29	Табличное значение критерия Стьюдента								
30									
31	3 0545								
32	5,0345								
33	Различаются ли два средних с учетом ошибки								
34	опыта при уровне значимочти 5%?								
36	HET								

Рис. 6. Проверка различия двух средних.

1) пусть измеряется площадь квадрата, в этом случае нужно произвести над результатом измерения длины стороны квадрата операцию возведения во вторую степень, 2) если измеряется площадь прямоугольника, то тогда нужно произвести над результатами измерения длин двух непараллельных сторон прямоугольника операцию произведения.

После того, как исходные данные введены (аналогично шагу 1, каждому типу данных отведен собственный столбец), переходим на лист «Расчет» и нажимаем кнопку с названием «СОРЕ». На шаге 2 в поле окна «СОРЕ, вопрос о первичной обработке данных» вводится формула для обсчета введенных исходных данных для первой строки (например, =A1+B1 или =A1^2). После завершения работы с этим окном, остальные шаги совпадают с описанными выше.

В том случае, если нужно узнать, можно ли объединить две (или более) выборки в одну, то можно воспользоваться блоком программы СОРЕ, расположенным на листе «Проверка различия двух средних» (рис. 6). В два столбца вводятся данные:

1) название (в принципе, не обязательно);

2) количество опытов (равно количеству элементов в выборке);

3) среднее значение определяемой величины (полученное выборочное среднее X_{cn});

4) значение дисперсии (полученная выборочная средняя дисперсия S_v^2);

5) вероятность (с которой будет подтверждена положительная гипотеза об объединении двух выборок).

Внизу, в последней ячейке, выводится ответ о возможности объединения (с заданной вероятностью) двух выборок.

Важно отметить, что ошибка метода (с помощью которого были получены экспериментальные результаты) должна быть учтена при определении числа дополнительных измерений, среднего значения определяемой величины, а также значения дисперсии.

ЛИТЕРАТУРА:

1. Кузьменко, В. Г. Программирование на VBA 2002 / В. Г. Кузьменко. – М. : Бином, 2003. – 876 с.

2. Хальворсон, М. Microsoft Visual Basic 6.0 для профессионалов. Шаг за шагом : практ. пособ. / М. Хальворсон. – М. : ЭКОМ, 1999. – 720 с.

3. Справочник по вероятностным расчетам / Г. Г. Абезгауз [и др.]. – М. : Военное издательство Министерства обороны СССР, 1970. – 536 с.

4. Феллер, В. Введение в теорию вероятностей и ее приложения : в 2 т. / В. Фелер. – М. : Мир, 1967. – Т. 1. – 498 с.

5. Корн, Г. Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – М. : Наука, 1977. – 832 с.

6. Протасов, К. В. Статистический анализ экспериментальных данных / К. В. Протасов. – М. : Мир, 2005. – 142 с.