

***Русаков Александр Александрович***

канд. физ.-мат. наук, д-р пед. наук, профессор  
ФГБОУ ВО «Московский технологический университет»

г. Москва

***Русакова Вера Николаевна***

канд. пед. наук, доцент  
ФГБОУ ВО «Орловский государственный  
университет им. И.С. Тургенева»

г. Орел, Орловская область

***Халатова Барият Шапиевна***

студентка

Институт естественных наук и биотехнологии

ФГБОУ ВО «Орловский государственный  
университет им. И.С. Тургенева»

г. Орел, Орловская область

## **ОБУЧЕНИЕ ПРИМЕНЕНИЮ ВОЗМОЖНОСТЕЙ MS EXCEL ДЛЯ ОПРЕДЕЛЕНИЯ ВЫБОРОЧНЫХ ХАРАКТЕРИСТИК НА ПРИМЕРЕ РАСПРЕДЕЛЕНИЯ КОЛИЧЕСТВА МЕДИ В РАСТВОРЕ**

***Аннотация:** в статье описываются возможности MS Excel для определения выборочных характеристик распределения на примере распределения количества меди в растворе и обосновывается важность ознакомления с соответствующим инструментарием студентов естественнонаучных направлений подготовки.*

***Ключевые слова:** методы математической статистики, обработка результатов эксперимента, выборочные характеристики распределения, функции MS Excel, инструменты MS Excel.*

Умение применять математико-статистические методы обработки информации при решении задач аналитической химии – одна из важнейших компетенций, которую студенты специальности 04.03.01 – Химия должны освоить в ходе

изучения курса математики. Однако значительная часть статистических методов требует существенных временных затрат на вычисления, производимые с большим объемом данных. Поэтому важно освоить современные инструменты, позволяющие несколько упростить работу исследователя, хотя бы в указанном плане. Речь идет о статистических программных пакетах и программах, позволяющих вычислять стандартные статистические функции.

К таким программам можно отнести

- пакет STATISTICA, который предоставляет мощные и удобные в использовании инструменты для статистического и графического анализа данных и прогнозирования;

- пакет MathCAD, который способен решать практически все задачи из различных областей применения математики, в том числе, задачи теории вероятностей и математической статистики. К аналогичным можно отнести пакеты Matlab, Mathematica, Derive и т. п. [2].

Однако перечисленные и подобные им программные продукты не являются общедоступными как в плане их приобретения и установки, так и специфики работы – необходимо изучить интерфейс программы, особенности работы необходимых функций, а для некоторых даже собственный язык для ввода формул. В связи с этим, не профессиональному математику не подготовленному должным образом к применению специализированных программ будет проще использовать в работе программу MS Excel, входящую в пакет MS Office. Конечно, набор статистических функций в нем достаточно ограничен, но для начинающего исследователя вполне приемлем, поэтому остановимся именно на нем.

Рассмотрим в качестве примера вычисление основных числовых характеристик распределения количества меди в растворе.

Группой из 20 студентов был проведен опыт по определению количества меди в выданном растворе (20 мл раствора соли меди (II) ( $\text{CuSO}_4$ ) с содержанием примерно 600 мг меди) [5]. Всем студентам было предложено сдать результаты пяти параллельных определений, округлив ответы до 0,5 мг. Были получены

следующие результаты в порядке увеличения найденного содержания меди ( $n_i$  - число совпадающих результатов,  $q_i$  - количество меди, мг), (табл. 1).

Таблица 1

$q_i$	$n_i$	$q_i$	$n_i$	$q_i$	$n_i$	$q_i$	$n_i$	$q_i$	$n_i$
600,5	1	605,5	3	609	3	612,5	5	616	2
601,5	1	606	1	609,5	6	613	4	616,5	1
602,5	1	606,5	3	610	5	613,5	3	617	2
603	1	607	4	610,5	5	614	3	617,5	1
604	2	607,5	4	611	6	614,5	3	618	1
604,5	2	608	5	611,5	4	615	2	618,5	1
605	4	608,5	4	612	4	615,5	2	621	1

Отдельные выборочные характеристики (среднее выборочное  $\bar{q}=610,0$  мг и стандартное отклонение  $s=4$  мг) найдены в [5]. Проведем расчеты этих и др. основных числовых характеристик представленной выборки при помощи MS Excel, воспользовавшись его вычислительными возможностями и встроенными функциями, и сравним результаты.

	A	B	C	D	E	F	G
1	$q_i$	$n_i$	$q_i n_i$	$q_i - \bar{q}$	$n_i (q_i - \bar{q})^2$	$n_i (q_i - \bar{q})^3$	$n_i (q_i - \bar{q})^4$
2	$q_1$	$n_1$					
3	$q_2$	$n_2$					
4	...	...	...	...	...		
5	$q_m$	$n_m$					
6		$n = \sum_{i=1}^m n_i$	$\bar{q} = \frac{\sum_{i=1}^m q_i n_i}{n}$		$s^2 = \frac{\sum_{i=1}^m n_i (q_i - \bar{q})^2}{n - 1}$	$A = \frac{\sum_{i=1}^m n_i (q_i - \bar{q})^3}{n s^3}$	$E = \frac{\sum_{i=1}^m n_i (q_i - \bar{q})^4}{n s^4} - 3$
7					$s = \sqrt{s^2}$		

Рис. 1

Составим в MS Excel расчетную таблицу (рис.1.) и вычислим среднее выбо-

рочное  $\bar{q} = \frac{\sum_{i=1}^m q_i n_i}{n}$ ; несмещенную оценку выборочной дисперсии

$s^2 = \frac{\sum_{i=1}^n (q_i - \bar{q})^2 n_i}{n - 1}$ , среднего квадратического отклонения  $s$ ; рассчитаем также

коэффициент асимметрии  $A = \frac{\mu_3}{s^3}$ , где  $\mu_k = \frac{\sum_{i=1}^m (q_i - \bar{q})^k n_i}{n}$  – центральный момент  $k$ -го порядка статистического ряда распределения (если  $A=0$ , то варианты, равноудаленные от среднего выборочного имеют одинаковую частоту; при  $A>0$  или  $A<0$  говорят о положительной (правосторонней) или отрицательной (левосторонней) асимметрии); эксцесс (или коэффициент эксцесса) статистического ряда распределения  $E = \frac{\mu_4}{s^4} - 3$  (эксцесс является показателем «крутости» статистического ряда распределения по сравнению с нормальным распределением, так как эксцесс нормально распределенной случайной величины равен нулю, если  $E>0$  ( $E<0$ ), то полигон статистического ряда распределения имеет более крутую (пологую) вершину по сравнению с нормальной кривой) [1; 4].


Для этого скопируем значения  $q_i$  в столбец А электронной таблицы, а значения  $n_i$  в столбец В. Подсчитаем объем выборки – в ячейке В37 найдем сумму всех элементов столбца по формуле =СУММ(В2:В36). Далее рассчитываем произведение  $q_i n_i$ . Для этого в ячейку С2 помещаем формулу =А2\*В2 и с помощью автозаполнения растягиваем на весь интервал С2:С36.

Вычисляем среднее выборочное  $\bar{q}$ : в ячейку С37 вводим формулу =СУММ(С2:С36)/В37. В столбце D находим разность  $q_i - \bar{q}$ , копируя формулу =А2-С\$37 из ячейки D2 на интервал D2:D36. В столбце E вычисляем произведение элементов столбца В на квадраты соответствующих элементов столбца D: =В2\*D2^2. В ячейке E37 найдем дисперсию  $s^2$  при помощи формулы =СУММ(E2:E36)/(В37–1). В ячейке E38 – среднее квадратическое отклонение  $s$  по формуле =КОРЕНЬ(E37). В столбцах F и G копируем соответственно формулы =В2\*D2^3 и =В2\*D2^4. В ячейке F37 находим коэффициент асимметрии А по формуле =СУММ(F2:F36)/(В37\*Е38^3), а в ячейке G37 – эксцесс Е по формуле =СУММ(G2:G36)/(В37\*Е38^4)-3 (результаты вычислений – рис. 2).

	A	B	C	D	E	F	G
1	$q_i$	$n_i$	$q_i n_i$	$q_i - \bar{q}$	$n_i (q_i - \bar{q})^2$	$n_i (q_i - \bar{q})^3$	$n_i (q_i - \bar{q})^4$
2	600,5	1	600,50	-9,74	94,87	-924,01	8999,86
3	601,5	1	601,50	-8,74	76,39	-667,63	5835,07
4	602,5	1	602,50	-7,74	59,91	-463,68	3588,92
5	603	1	603,00	-7,24	52,42	-379,50	2747,60
6	604	2	1208,00	-6,24	77,88	-485,94	3032,27
30	616	2	1232,00	5,76	66,36	382,21	2201,51
31	616,5	1	616,50	6,26	39,19	245,31	1535,67
32	617	2	1234,00	6,76	91,40	617,83	4176,54
33	617,5	1	617,50	7,26	52,71	382,66	2778,09
34	618	1	618,00	7,76	60,22	467,29	3626,16
35	618,5	1	618,50	8,26	68,23	563,56	4655,01
36	621	1	621,00	10,76	115,78	1245,77	13404,45
37		100	610,24		16,00	0,05	-0,27
38					4,00		

Рис. 2

Упростить вычисления позволит применение встроенных функций MS Excel. Среднее выборочное вычисляется при помощи функции =СРЗНАЧ(), «исправленная» выборочная дисперсия =ДИСП(), выборочное среднее квадратическое отклонение =СТАНДОТКЛОН(), асимметрия =СКОС(), эксцесс =ЭКССЕСС(), где в скобках указывается диапазон значений – столбец, в ячейках которого перечислены все значения признака, причем каждое столько раз, сколько оно встречается в выборке. Для приведенного примера получим следующие результаты (в соответствующих единицах измерения)  $\bar{q}=610,24$ ,  $s^2=16,00$ ,  $s=4,00$ ,  $A=0,06$ ,  $E=-0,16$ .

Те же значения можно найти, воспользовавшись инструментом «Описательная статистика» надстройки «Пакет анализа» (подробнее с возможности данного инструмента можно ознакомиться, например, в [3]). Отметим, что, например, в версии MS Excel 2007 средства, которые включены в надстройку *Пакет анализа*, доступны через команду *Анализ данных* вкладки *Данные*. (Если кнопки *Анализ данных* нет на вкладке *Данные*, нужно подключить эту надстройку: кнопка *Microsoft Office*  → кнопка *Параметры Excel* → на вкладке *Надстройки* выбрать *Пакет анализа* и в поле *Управление* → *Надстройки Excel* → кнопка *Перейти* →

в поле *Доступные надстройки* установить флажок *Пакет анализа* → ОК). Итог применения инструмента приведен на рис. 3.

Столбец1	
Среднее	610,24
Стандартная ошибка	0,40
Медиана	610,25
Мода	609,50
Стандартное отклонен	4,00
Дисперсия выборки	16,00
Эксцесс	-0,16
Асимметричность	0,06
Интервал	20,50
Минимум	600,50
Максимум	621,00
Сумма	61024,00
Счет	100,00

Рис. 3

Проанализируем полученные результаты. Небольшие расхождения в значениях коэффициентов асимметрии и эксцесса, найденных по расчетной таблице и при помощи встроенных функций и инструментов MS Excel объясняются большей точностью оценки данных коэффициентов в последних (точные формулы можно уточнить в справке MS Excel для этих функций). А вот вычисления, выполненные «вручную» (быть может, с использованием лишь калькулятора) дали неточность уже при вычислении выборочной средней, что может быть объяснено как потерей точности при округлении промежуточных вычислений, так и элементарными ошибками более чем допустимыми при работе с большим объемом данных, что еще раз подтверждает преимущество применения возможностей MS Excel для нахождения основных числовых характеристик распределения.

### ***Список литературы***

1. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. – М.: Юнити-Дана, 2004. – 573 с.

2. Математические пакеты для решения статистических задач. Ч. 2 [Электронный ресурс]. – Режим доступа: [https://www.distanz.ru/feed/lectures/matematicheskie-pakety-dlya-resheniya-statisticheskikh-zadach\\_3777](https://www.distanz.ru/feed/lectures/matematicheskie-pakety-dlya-resheniya-statisticheskikh-zadach_3777)
3. Минько А.А. Статистический анализ в MSExcel. – М.: Вильямс, 2004. – 448 с.
4. Русаков А.А. Учебный курс математики для преподавателей, студентов, учащихся старших классов и учителей математики / Русаков А.А., Русакова В.Н, Ильина Н.А. – Saarbrücken: LAP LAMBERT Academic Publishing GmbH & Co. KG, 2012. – 340 с.
5. Чарыков А.К. Математическая обработка результатов химического эксперимента: Учеб. пособие для вузов. – Л.: Химия, 1984. – 168 с.