The misuse of 'standard error of mean' for the significant figure of the mean value

El uso inadecuado del error estándar de la media para determinar sus cifras significativas

SHYH-JEN WANG

Department: Division of Experimental Surgery, Department of Surgery Veterans General Hospital -Taipei and Institute of Biomedical Engineering, National Yang-Ming University, Taiwan Republic of China
wangsj@vghtpe.gov.tw

Abstract

There are two rules of thumb in determining the number of significant figures in a calculated. When multiplying or division several quantities, the numbers of significant figures in the final answer is the same as the number of significant figures in the least accurate of the quantities entering the calculation. When numbers are added or subtracted, the last significant figure is in the same column as the last significant figure in the last accurate number. Furthermore, exercising the rules to propagate significant figures of a mean value might still have problems. Other than those rules, the standard error of mean was treated as the gold standard to propagate significant figures of the mean value. However, this article reviews the definition of the standard error of mean in statistics and finds that maybe the common expression causes the misunderstanding.

Key words: *significant figure, mean value, standard error of mean.*

Resumen

Hay dos reglas para determinar el número de cifras significativas en un cálculo. Al multiplicar o dividir varios números, el número de cifras significativas en la respuesta final es el mismo que el número de cifras significativas en la menos precisa que entra en cómputo. Cuando los números son sumados o restados, la última cifra significativa está en la misma columna que la última cifra en el número menos preciso. Además, en las reglas existentes que propagan las cifras significativas de un valor medio se tienen aun problemas. En otras reglas se ha tratado el error estándar de la media, como estándar dorado en la propagación de las cifras significativas del valor de la media. En este artículo se revisa la definición del error estándar de la media en estadística y se muestra que tal vez la expresión habitual de este error es la causa de la interpretación inadecuada.

Palabras clave: *cifras significativas, valor medio, error estándar de la media.*

INTRODUCTION

When certain quantities are measured, the measured values are known only to with the limits of the experimental uncertainty. The value of the uncertainty depends on various factors, such as the quality of the apparatus, the skill of the experimenter, and the number of measurements performed. There is no such thing as an exact measurement. An approximate method to keep track of the accuracy of numbers is to write only those figures that are significant. The last digit in a number is considered to have some significance, but it may not be exact.

According to the method of significant figures, a calculated number, involving measured quantities, should have a limited number of significant figures. Even if a pocket calculator gives an answer to nine digits, not all of them are necessarily significant since the numbers that went into the calculation had limited accuracy. So the calculated number has uncertainty. There are two good rules of thumb (SCHWART LOWELL 1985, SERWAY 1998) in determining the number of significant figures in a calculated. When multiplying or division several quantities, the numbers of significant figures in the final answer is the same as the number of significant figures in the least accurate of the quantities entering the calculation. When numbers are added or subtracted, the last significant figure is in the same column as the last significant figure in the last accurate number.

Regarding the issue of significant figures in calculations, almost two decades ago, professor SCHWARTZ from University of Massachusetts wrote an interesting article with the title of 'propagation of significant figures' (SCHWART LOWELL 1985). Schwartz's article drew our attention to the fact that the se two conventional rules of thumb for propagating significant figures might have problems. However, others seem not to follow with Professor Schwartz's effort – and we also take this review. The very recently published textbooks including physics (URONE 2001, CUTNELL, JOHNSON 1998, EUGENE 2000) and chemistry (JONES 2000, MARTIN 2000, MARTIN 2000) still cite these conventional rules for propagating significant figures.

The issue of significant figure was also debated in my physics class. The test problem listed a group of eight measurements (1.28, 1.27, 1.28, 1.28, 1.28, 1.27, 1.28 and 1.27 seconds) and asked to calculate the mean value. A group of students comes out with the answer for the mean value of eight measurements was 1.276 (seconds), which had one more decimals place than those of the testing data. The supporting comment lied in that the 'true' value must be somewhere between 1.27 and 1.28 thus giving increased confidence that the mean value is 1.27 and a bit. However, some students would hold to the opposite view that the last digit of experimental data has some uncertainty in its value. For example, the result of a measurement may be 1.28 second with an uncertainty of 1%. Since 1% of 1.28 is approximately 0.01, the result is 1.28±0.01 second. The true value is likely lie between 1.27 and 1.29 second. Instead of an explicit statement of uncertainty, the number of digits retained often indicates the precision of a result. The value of 1.28 has three significant figures, with the understanding that the last figure may not be certain. Moreover, the above-mentioned problem did not specify the uncertainty. Therefore, according to the method of the significant figure, to add another decimal on the mean value is nothing but to increase the implied accuracy of experiment just simply by mathematical manipulation.

To access the debate, several rules for propagating significant figures have been proposed (SCHWART LOWELL 1985) for the calculation of mean value. Among these propagation procedures, the standard error of the mean seems to be treated as the gold standard, as indicated by Professor SCHWARTZ "…later calculated the standard error of the mean as 0.012g. This confirmed that the statistical uncertainty of the mean as expressed as the standard error estimate was within the hundredths decimal place", PARRATT (PARRATT 1961) also pointed out that "as a guide in determining the proper number of significant figures with which to express the precision of a mean determined from seven or more equally weighted measurement, the mean should have one more significant figure than has each measurement. In general, justification for the rule, and indeed the proper number of significant figures for the mean in any case, is indicated by the magnitude of, say, the standard deviation or, better, of the standard deviation in the mean". If the standard error of the mean or the standard deviation in the mean is the best way to propagate significant figures, how come Professor SCHWARTZ (SCHWART LOWELL 1985) summarized that "…hope this paper will stimulated others to offer contribution toward these unsolved problems"? Therefore, the more detail discussion of the standard error of the mean will be presented in this article.

THE STANDARD ERROR OF THE MEAN

In statistics, samples are usually drawn from much larger populations; and data are collected about the sample to find out something about the population. Furthermore, the probability theory enables the usage of the samples to estimate quantities in populations, and to determine the precision of these estimates. Using a suitable random sampling method, the sampling experiment draws repeated samples from the population. The reiterating procedures give these sampling data as well as their means.

Usually, these sample means are not all the same and would form a distribution. The distribution of all possible sample means is called the sampling distribution of the mean. In general, the sampling distribution of any statistic is the distribution of the value of the statistic arising from all possible samples. The sample mean is an estimate of the population mean. The standard deviation of its sampling distribution is called the standard error of the estimate, which provides a measure of how far from the true value the estimate is likely to be. In almost all practical situations, we do not know the true value of the population variance but only its estimate. Therefore, use the formula $s/\sqrt{n}$ to estimate the standard error (BLAND MARTIN, 2000), where s and n are the standard deviation and size of the sample. The estimate is referred to as the standard error of the mean.

The mean and standard error are often written as mean ± standard error. However, as pointed out by BLAND (BLAND MARTIN, 2000), the common expression would be rather misleading in that the true value may be up to two standard errors from the mean with reasonable probability. The standard error is often confused with the standard deviation (DAWSON *at all* 2000). The standard deviation is concerned with the variability of samples, but the standard error is used to measure the precision of estimates. Furthermore, the population mean is estimated to lie somewhere in the interval between these limits, which is called confidence interval in statistics. In the language of statistics mean ± 1.96 standard error, there is 95% confident that the mean lies between the ± 1.96 standard error limits.

CONCLUSIONS

From the above discussion, the expression of mean ± standard error is far from the definition of the measurement ± uncertainty (URONE 2001). The standard error of mean should not be treated as the uncertainty, even though they are frequently expressed in the same format. According to the definition of standard error of mean to rewrite Professor Schwartz's words, it should be read as "…later calculated the standard error of the mean as 0.012g. There is 95% confidence that the mean lies between the 2.47 ± 0.0235g limits".

The standard error of mean was treated as the gold standard to propagate significant figures of the mean value (SCHWART LOWELL 1985, PARRATT 1961). However, this article reviews the definition of the standard error of mean in statistics and finds that maybe the common expression causes the misunderstanding (BLAND MARTIN, 2000). Moreover, the newly published textbooks in physics and chemistry do not adopt the method of standard error of mean dealing with the significant figure of the mean value.

Any mathematical manipulation, such as calculating the mean value for a group of measurements, certainly could not increase accuracy of experimental values, but it can increase the confidence that the true answer is within a particular range. Therefore, propagating significant figures of the mean value by the standard error of mean is not recommended since it is difficult to apply appropriately.

ACKNOWLEDGMENT

The author would like to appreciate Prof. KWANG-TING LIU from Department of Chemical Science of National Taiwan University for his kindly providing information and comments.

BIBLIOGRAPHY

BLAND, MARTIN, *An introduction to medical statics*, 3rd, Oxford University Press, 2000.

BRADY, J.E., RUSSELL, J.W. and HOLUM, J.R., *Chemistry: The Study of Matter and Its Changes*, 3rd, John Wiley & Sons, Inc., 2000.

CUTNELL, J.D. and JOHNSON K.W., *Physics*, John Wiley & Sons, Inc., 1998.

DAWSON, BETH and TRAPP, ROBERT, *Basic and Clinical Biostatistics*, 3rd, McGraw-Hill, 2000.

EUGENE, H., *Physics*, Brooks/Cole, 2000.

JONES, L., *Chemistry molecules, matter, and change*, W.H. Freeman, New York, 2000.

MARTIN, S.S., *Chemistry: the molecular nature of matter and change*, McGraw-Hill, Boston, 2000.

PARRATT, L.G., *Probability and experimental error in science*, John Wiley & Sons, Inc. New York, 1961.

SCHWART LOWELL, M., Propagation of Significant Figures, *Journal of Chemical Education*, vol. 62, 8, pp. 693-697, 1985.

SERWAY, R.A., *Principles of Physics*, 2nd edition, Saunders College Publishing, 1998.

URONE, P.P., *College Physics*, 2nd Thomson Learning Inc., 2001.

Received 31.10.2003 / **Approved** 21.09.2004

Hungarian secondary school students' strategies in solving stoichiometric problems

Estrategias de estudiantes húngaros de escuela secundaria para resolver problemas estequiométricos

ZOLTÁN TÓTH AND EDINA KISS

Team of Chemical Methodology, Faculty of Science, University of Debrecen, Debrecen, Hungary
o319tz@tigris.klte.hu

Abstract

A random sample of 750 out of 2954 Hungarian secondary school students (grade 7 to 11, aged 13-17) from 17 schools were participated in a paper-and-pencil test with free-response problem on the composition of binary compounds. In this study the following research questions were investigated: (1) whether Hungarian students – similar to the German high school students – also created their own strategy in solving simple stoichiometric problems, or they used the algorithmic methods learned at school, and (2) how the students' strategies changed during the education. We found that contrary to German high school students, Hungarian secondary school students applied the strategies learned at school (the mole method and the proportionality method) in stoichiometric calculations. The success and the ratio of the mole method to the proportionality method increased with the age of the students. Three possible interpretations of the contradiction results are discussed.

***Key words**: stoichiometry, problem-solving strategies, composition of binary compounds*

Resumen

Una muestra aleatoria de 750 entre 2.954 estudiantes húngaros de la escuela (grado de 7 a 11, y edad 13-17) participaron en una prueba de papel y lápiz con un problema de respuesta libre sobre la composición de sustancias binarias. En este estudio se investigaron las preguntas siguientes: (1) si los estudiantes húngaros son similares a los estudiantes alemanes de escuela secundaria y pueden crear su propia estrategia para resolver problemas simples de estequiometría, o usan los métodos algorítmicos aprendidos en la escuela, y (2) cómo los estudiantes cambiaron las estrategias durante la educación. Se encontró que al contrario de los estudiantes alemanes, los estudiantes húngaros aplicaron las estrategias aprendidas en la escuela (los métodos de mol y de reglas de tres) en los cálculos estequiométricos. Los logros y la proporción del método de mol al método de reglas de tres aumentan con la edad de los estudiantes. Se discuten tres posibles interpretaciones de los resultados contradictorios.

***Palabras clave**: estequiométria, estrategias para resolver problemas, composición de sustancias binarias*

INTRODUCTION

Research shows that the problem-solving strategy a student applies depends on different factors. SCHMIDT (1994, 1997) reported that the high school students in Germany successfully used their own strategy in solving simple stoichiometric problems, but tended to use algorithmic methods in case of difficult problems. In balancing chemical equations we found (TÓTH, 2004) that Hungarian high school students created their own balancing strategy (mainly the trial-and-error) before learning the oxidation number method at school, and they stuck to their own strategies of low efficiency even in case of complicated redox equations.

In this study we investigated the questions:

1. whether Hungarian students - similar to the German high school students - also created their own strategy in solving simple stoichiometric problems, or they used the algorithmic methods learned at school, and
2. how the students' strategies changed during the education.

In this survey we used paper-and-pencil test with free-response problem on the composition of binary compounds similar to those developed by SCHMIDT (1992, 1994, 1997):

'How many grams of carbon are there in 96 g MgC_2? Write down your solution. $A_r(Mg) = 24$; $A_r(C) = 12$'