Universal Data Processing Science with Multiple-Sources Intelligent Iteration

Ph. D. & Dr. Sc. Lev G. Gelimson (AICFS)

Mathematical Journal of the "Collegium" All World Academy of Sciences, Munich (Germany), 13 (2013), 6.

Modern data processing is based on classical probability theory and statistics [1] with their defects. Regression analysis gives the least square method results different by mutually replacing the variables roles and only qualitatively proves the fact of dependence via rejecting the null hypothesis of full independence. There are no adequately estimating the errors of the dependence formula and no improving it. Single-source iteration has an inflexible algorithm often with very slow convergence.

Universal data processing science [2-4] includes a number of corresponding theories. Direct equiprecise measurement theory proves that for groups i (i = 1, 2, … , n) of data x_{ij} (j = 1, 2, … , k(i) = k_i), the variance of a separate data by known exact value a

$$\sigma_x^2 = \Sigma_{i=1}^n \Sigma_{j=1}^{k(i)} (x_{ij} - a)^2 / \Sigma_{i=1}^n k(i) = \{\Sigma_{i=1}^n k(i) \Sigma_{j=1}^{k(i)} [(x_{ij} - a)^2 / k(i)]\} / \Sigma_{i=1}^n k(i)$$

is groupwise invariant. The same holds for the general mean value

$$m = \Sigma_{i=1}^n \Sigma_{j=1}^{k(i)} x_{ij} / \Sigma_{i=1}^n k(i) = [\Sigma_{i=1}^n k(i) \Sigma_{j=1}^{k(i)} x_{ij} / k(i)] / \Sigma_{i=1}^n k(i)$$

(replacing a) if and only if either the dependence of m on x_{ij} with the Bessel formula is ignored and $\Sigma_{i=1}^n k(i)$ rather than $\Sigma_{i=1}^n k(i) - 1$ is used in the last denominator or this subtracting 1 also applies to the group weights k(i) - 1 instead of k(i):

$$\sigma_x^2 = \Sigma_{i=1}^n \Sigma_{j=1}^{k(i)} (x_{ij} - m)^2 / [\Sigma_{i=1}^n k(i) - 1]$$
$$= \{\Sigma_{i=1}^n [k(i) - 1] \Sigma_{j=1}^{k(i)} (x_{ij} - m)^2 / [k(i) - 1]\} / [\Sigma_{i=1}^n k(i) - 1].$$

For groupwise noninvariant variances $\sigma_{x(i)}^2$ about groupwise mean values m(i),

$$\sigma_x^2 = \{\Sigma_{i=1}^n [k(i) - 1] \sigma_{x(i)}^2 + \Sigma_{i=1}^n k(i) m(i)^2 - \Sigma_{i=1}^n k(i) m^2\} / [\Sigma_{i=1}^n k(i) - 1].$$

Direct nonequiprecise measurement theory gives weights $p_i k_i/\sigma_{x(i)}^2$ to data means x_i with variances $\sigma_{x(i)}^2$ and finally for weighted arithmetic mean m and its variance σ_m^2

$$m = \Sigma_{i=1}^n w_i k_i/\sigma_{x(i)}^2\, x_i / \Sigma_{i=1}^n w_i k_i/\sigma_{x(i)}^2,\ \sigma_m^2 = 1 / \Sigma_{i=1}^n w_i k_i/\sigma_{x(i)}^2.$$

Unimode theory introduces unimode u (unlike the discrete or descriptive mode) as the most probable value which generalizes the continuous distribution mode. The weighted power unimode method gives for t > 0 equation $\Sigma_{i=1}^n w_i (x_i - u)^{"t} = 0$ with using base sign conserving exponentiation $a^{"b} = |a|^b$ sign a . By $x_1 \le x_2 \le \dots \le x_n$, apply asymmetric distributions with different left and right variances σ_L^2 and σ_R^2.

Extreme data maximum and mean correction theories without ignoring outliers least-correct the equal amounts of both the least and the greatest data so that their distances from the next data do not exceed the maximum or the mean distance for the remaining data, respectively, e.g. by the classical Millikan data [5] (Figures 1, 2).

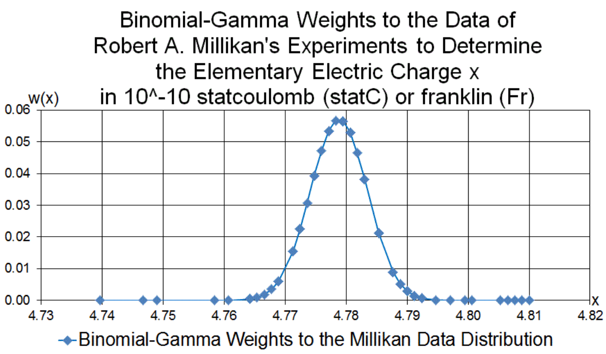

Figure 1. Weights w(x) to Millikan's Data

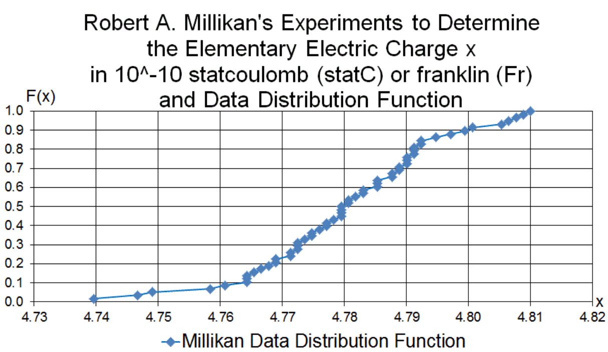

Figure 2. Millikan's Data Distribution F(x)

Multiple-sources iteration theory rationally introduces additional initial sources.
Intelligent iteration theory rationally replaces each already calculated iteration approximation with an absolutely freely chosen pseudosolution to a general problem.
The modern value for the elementary charge in 10^{-10} statcoulomb is e = 4.8032. Using Millikan's 58 data, classical statistics gives e = m = 4.7806 with standard deviation σ_e = 0.0147 but using median m_{ed} = 4.7801 and following Charlier [1] u = $3m_{ed}$ - 2m = 4.7791. Universal data processing science gives e = u = 4.784 with deviations ${}^u\sigma_L$ = 0.0053 and ${}^u\sigma_R$ = 0.0050. Therefore, applying this science to the classical Millikan data gives new results for the elementary charge as one of the fundamental physical constants. This also shows that it is very important to publish the complete experimental data and not only the final results of their statistical processing because new theories and methods can essentially improve the results.
Universal data processing science and multiple-sources intelligent iteration give correct results by data scatter, asymmetry, and outliers, e.g. in aeronautical fatigue.

Keywords: Ph. D. & Dr. Sc. Lev Gelimson, "Collegium" All World Academy of Sciences, Academic Institute for Creating Fundamental Sciences, Mathematical Journal, Universal Data Processing Science with Multiple-Sources Intelligent Iteration, probability theory, statistics, regression analysis, Legendre, Gauss, contradictory overdetermined problem, null hypothesis of full independence, single-source iteration, inflexible algorithm, direct equiprecise measurement theory, variance, universal mathematics, unimathematics, Unimathematik, overdetermined quantiset, quantity, direct nonequiprecise measurement theory, weighted arithmetic mean, unimode theory, weighted power unimode method, base sign conserving exponentiation, extreme data maximum and mean correction theories, outlier, asymmetric distribution, multiple-sources iteration theory, intelligent iteration theory, pseudosolution, general problem, elementary charge, standard deviation, median, Charlier, fundamental physical constant, statistical processing, data scatter, aeronautical fatigue, general problem theory, elastic mathematics, general strength theory, corrections and generalizations of the least square method, Cramér, Mathematical Methods of Statistics, Robert Millikan. On the Elementary Electric Charge and the Avogadro Constant.

[1] H. Cramér. Mathematical Methods of Statistics. Princeton University Press, 1945
[2] Lev G. Gelimson. General Problem Theory. Abhandlungen der WIGB (Wissenschaftlichen Gesellschaft zu Berlin), 3 (2003), Berlin, 26-32
[3] Lev Gelimson. Elastic Mathematics. General Strength Theory. The "Collegium" All World Academy of Sciences Publishers, Munich, 2004
[4] Lev G. Gelimson. Providing Helicopter Fatigue Strength: Flight Conditions. In: Structural Integrity of Advanced Aircraft and Life Extension for Current Fleets – Lessons Learned in 50 Years After the Comet Accidents, Proceedings of the 23rd ICAF Symposium, Claudio Dalle Donne (Ed.), 2005, Hamburg, Vol. II, 405-416
[5] Robert A. Millikan. On the Elementary Electric Charge and the Avogadro Constant. Phys. Rev., 2 (2), 1913, 109-143