Scientific
Research
Publishing

Chi-Square Distribution: New Derivations and Environmental Application

Thomas M. Semkow[1,2], Nicole Freeman[3], Umme-Farzana Syed[1], Douglas K. Haines[1], Abdul Bari[1], Abdul J. Khan[1], Kimi Nishikawa[1], Adil Khan[1], Adam G. Burn[1,2], Xin Li[1,2], Liang T. Chu[1,2]

[1]Wadsworth Center, New York State Department of Health, Albany, NY, USA
[2]Department of Environmental Health Sciences, University at Albany, State University of New York, Rensselaer, NY, USA
[3]Averill Park Central School District, Averill Park, NY, USA
Email: thomas.semkow@health.ny.gov

Abstract

We describe two new derivations of the chi-square distribution. The first derivation uses the induction method, which requires only a single integral to calculate. The second derivation uses the Laplace transform and requires minimum assumptions. The new derivations are compared with the established derivations, such as by convolution, moment generating function, and Bayesian inference. The chi-square testing has seen many applications to physics and other fields. We describe a unique version of the chi-square test where both the variance and location are tested, which is then applied to environmental data. The chi-square test is used to make a judgment whether a laboratory method is capable of detection of gross alpha and beta radioactivity in drinking water for regulatory monitoring to protect health of population. A case of a failure of the chi-square test and its amelioration are described. The chi-square test is compared to and supplemented by the *t*-test.

Keywords

Mathematical Induction, Laplace Transform, Gamma Distribution, Chi-Square Test, Gross Alpha-Beta, Drinking Water

1. Introduction

The chi-square distribution (CSD) has been one of the most frequently used distributions in science. It is a special case of the gamma distribution (see Section 2). The latter has been an important distribution in fundamental physics, for example as kinetic energy distribution of particles in an ideal gas (Maxwell-Boltzmann) [1] or the kinetic energy distribution of particles emitted from excited nuclei in nuclear reactions [2]. A historical context for the development of the CSD is de-

scribed in References [3] and [4]. Its first derivation is attributed to Bienaymé [5], who used multiple integrals over normal variables and substitutions. Abbe [6] used a method of integration in the complex plane to solve multiple integrals. The most general derivation is attributed to Helmert, who proposed a classic transformation to derive CSD, including calculation of the Jacobian determinant of transformation [7]. This transformation can be worked out into polar variables, which is described in statistical textbooks [4] [8].

The established fundamental derivations of the CSD described above lend themselves to complicated handling of multiple integrals. On the contrary, the simplified derivations use the fact that CSD is a special case of the gamma distribution. Owing to the integrable and recursive properties of the gamma distribution, as well as its moment generating function (Mgf), simplified derivations of CSD are described in the textbooks [9] [10]. Another simplified derivation uses Bayesian inference [11]. In Section 2, we refer to these methods for comparisons.

In this work, we present two new methods of derivation of the CSD. They are both within the simplified category. One of them is mathematical induction. The original derivation was done by Helmert [12] using a 2-step forward mathematical induction. We have elaborated on that and observed that the CSD has certain recursive property, which enables its derivation using a single-step induction plus the well-known theorem for beta and gamma functions. Another derivation method we describe is by the Laplace transform. This method has some similarity to the Mgf and characteristic function methods, owing to the presence of exponentiation. It uses a complex-variable integration and it is free from many assumptions of the other methods. The two new derivations of the CSD by mathematical induction and Laplace transform are described in Section 2.

Chi-square testing (CST) is closely related to and based upon the CSD. It has its origins in the discovery of the goodness-of-fit test by Pearson [13]. In the goodness-of-fit, one calculates the test statistics as

$$\chi_\nu^2 = \sum_{i=1}^m \frac{\left(O_i - E_i\right)^2}{E_i}, \tag{1}$$

where O_i is frequency of observation, E_i is expected frequency based on an assumed model distribution, for category of type *i*, and *m* is the number of categories. Both O_i and E_i are unitless. $\nu = m - 1 - p$ is the number of degrees of freedom, where *p* is number of parameters of the model distribution calculated from the data. For any model distribution, Equation (1) leads asymptotically to the CSD when the number of observations is large, which has been proved for the multinomial distribution by Pearson [13]. The goodness-of-fit CST has been extensively used in statistics and widely applied to many fields [3] [14]. It is worth noting that the interpretation of the degrees of freedom was provided by Fisher [15]. As example in physics, CST goodness-of-fit has been used to verify Poisson fluctuations of radioactivity counter [14] [16].

Another form of the chi-square variable from Equation (1) is written in the

general form as

$$\chi_v^2 = \sum_{i=1}^{n} \left(\frac{x_i - \mu_i}{\sigma_i} \right)^2, \tag{2}$$

where n is the number of observations, x_i is the observed variable, μ_i is the expected value, σ_i is the standard deviation, and $v \le n$. The variables in Equation (2) can be expressed in physical units. In the limit of large number of observations, the variable and parameters of Equation (2) are approximated by those of the normal variates, and the χ_v^2 distributes as CSD. In this work, we generalize this CST test to a combined test for variance and location as well as verify it with the *t*-test [17]. The test statistics studied are described in Section 3.

Within the context of this work, we present a unique application of the CST to the detection of radioactive contaminants in drinking water required by the Safe Drinking Water Act (SDWA) in the US. The bulk of natural alpha and beta/gamma (photon) radioactivity in drinking water originates from the possible presence of ^{238}U and ^{232}Th natural radioactive-series progeny, ^{226,228}Ra and their progeny, as well as ^{40}K radionuclides [18]. The SDWA regulations [19] establish a Maximum Contaminant Level (MCL) of 15 pCi/L (555 mBq/L) for gross alpha (GA) radioactivity, excluding U and Rn. For gross beta (GB) radioactivity, the MCL is limited by the total body or any organ radiation dose of 4 mrem/y (40 μSv/y). For both GA and GB, the Maximum Contaminant Level Goal (MCLG) is zero. Furthermore, SDWA requires Detection Limits (DL) of 3 pCi/L (111 mBq/L) and 4 pCi/L (148 mBq/L) for GA and GB radioactivity, respectively. These DLs must be met by all public health laboratories accredited for monitoring of GA and GB radioactivity in drinking water in the US. In Section 4, we detail a CST procedure to verify if the required above-mentioned DLs are met [20]. We investigate the reasons and consequences of failed CST and ameliorate such cases.

2. Chi-Square Distribution

The probability density function (Pdf) of the CSD is given by

$$\text{Pdf}\left(\chi_v^2 \mid v \right) = \frac{\left(\chi_v^2 \right)^{v/2-1} e^{-\chi_v^2/2}}{2^{v/2} \Gamma(v/2)}, \tag{3}$$

where Γ is the gamma function. The expectation value of CSD is $E\left[\chi^2 \right] = v$, and the variance $\text{Var}\left[\chi^2 \right] = 2v$ [21]. The CSD is a special case of the gamma distribution abbreviated as $\text{gamma}\left(\chi_v^2 \mid a, b \right)$ with the parameters $a = v/2$ and $b = 2$ [21].

To derive Equation (3), we start with the general definition of χ_v^2 statistics given by Equation (2) assuming normal variates. For a single normal variable x_1 with $\text{Pdf}(x_1)$, the probability of $x_1 \in [x_1, x_1 + dx_1]$ is given by

$$\text{Pdf}\left(x_1 \right) dx_1 = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 / 2} dx_1. \tag{4}$$

By substituting $\chi_1^2 = \left((x_1 - \mu_1)/\sigma_1 \right)^2$, we obtain from Equation (4)

$$\text{Pdf}\left(\chi_1^2 \mid 1\right)\mathrm{d}\chi_1^2 = \frac{2}{\sqrt{2\pi}\sigma_1}e^{-\chi_1^2/2}\left|\frac{\mathrm{d}x_1}{\mathrm{d}\chi_1^2}\right|\mathrm{d}\chi_1^2 = \frac{\left(\chi_1^2\right)^{1/2-1}e^{-\chi_1^2/2}}{2^{1/2}\,\Gamma\left(1/2\right)}\mathrm{d}\chi_1^2 \qquad (5)$$
$$= \text{gamma}\left(\chi_1^2 \mid 1/2, 2\right)\mathrm{d}\chi_1^2,$$

which has the Pdf given by Equation (3) for $\nu = 1$. In deriving Equation (5), we also used $\Gamma\left(1/2\right) = \sqrt{\pi}$, whereas factor of 2 originated from the fact that the x_1 variable ranging from minus infinity to plus infinity has been substituted with the χ_1^2 variable ranging from zero to plus infinity.

Let us assume that the $n+1$ term with the normal x_{n+1} variable was added to Equation (2), and that this addition raised the number of degrees of freedom to $\nu+1$. Then,

$$\chi_{\nu+1}^2 = \chi_\nu^2 + \left(\frac{x_{n+1}-\mu_{n+1}}{\sigma_{n+1}}\right)^2. \qquad (6)$$

Using the calculus for probability density functions [21],

$$\text{Pdf}\left(\chi_{\nu+1}^2 \mid \nu+1\right)\mathrm{d}\chi_{\nu+1}^2 = \int_{-\infty}^{+\infty}\text{Pdf}\left(\chi_\nu^2 \mid \nu\right)\mathrm{d}\chi_\nu^2\,\text{Pdf}\left(x_{n+1}\right)\mathrm{d}x_{n+1}. \qquad (7)$$

Let us define a new variable z, such as

$$\left(\frac{x_{n+1}-\mu_{n+1}}{\sigma_{n+1}}\right)^2 = \chi_{\nu+1}^2\left(1-z\right). \qquad (8)$$

By realizing that $\mathrm{d}\chi_{\nu+1}^2 = \mathrm{d}\chi_\nu^2$, and performing all substitutions, the right side of Equation (7) can be rewritten as

$$2\int_0^1\text{Pdf}\left(\chi_\nu^2 \mid \nu\right)\text{Pdf}\left(x_{n+1}\right)\left|\frac{\mathrm{d}x_{n+1}}{\mathrm{d}z}\right|\mathrm{d}z\,\mathrm{d}\chi_{\nu+1}^2$$
$$= \frac{\left(\chi_{\nu+1}^2\right)^{(\nu+1)/2-1}e^{-\chi_{\nu+1}^2/2}}{2^{(\nu+1)/2}\,\Gamma\left(\nu/2\right)\Gamma\left(1/2\right)}\mathrm{d}\chi_{\nu+1}^2\int_0^1 z^{\nu/2-1}\left(1-z\right)^{1/2-1}\mathrm{d}z. \qquad (9)$$

However, the integral on the right side of Equation (9) is the beta function, $B\left(\nu/2, 1/2\right)$, which is related to the gamma functions by [22],

$$B\left(\nu/2, 1/2\right) = \frac{\Gamma\left(\nu/2\right)\Gamma\left(1/2\right)}{\Gamma\left((\nu+1)/2\right)}. \qquad (10)$$

By inserting Equation (10) into Equation (9), simplifying, and comparing with the left side of Equation (7), one obtains

$$\text{Pdf}\left(\chi_{\nu+1}^2 \mid \nu+1\right) = \frac{\left(\chi_{\nu+1}^2\right)^{(\nu+1)/2-1}e^{-\chi_{\nu+1}^2/2}}{2^{(\nu+1)/2}\,\Gamma\left((\nu+1)/2\right)}, \qquad (11)$$

which is the Pdf given by Equation (3) for $\nu+1$ degrees of freedom and it proves Equation (3) by induction.

By substituting $\varphi_i^2 = \left((x_i - \mu_i)/\sigma_i \right)^2$, Equation (2) becomes

$$\chi_\nu^2 = \sum_{i=1}^n \varphi_i^2 \qquad (12)$$

The sum of independent random variables φ_i^2 is called a convolution and the joint distribution function for χ_ν^2 can be obtained by calculating an *n*-dimensional convolution integral. Exploring the properties of this convolution leads to simplifications, which have been used in the literature. By convoluting two gamma distributions $\chi_1^2 \equiv \varphi_i^2$ from Equation (5) and using the theorem that the convolution of two gammas is also a gamma, one obtains $\text{gamma}\left(\chi_2^2 \mid 2/2, 2\right)$ [9]. By continuing this process of convoluting with χ_1^2, it is easy to infer that the full convolution is equal to $\text{gamma}\left(\chi_\nu^2 \mid \nu/2, 2\right)$, where $\nu = n$, which the CSD given by Equation (3). This provides a simplified derivation of CSD using convolution.

Another simplified derivation of CSD uses the theorem that the Mgf of convolution is a product of individual Mgfs [10]. Thus, by calculating Mfg of χ_1^2 from Equation (5) and taking it to the *n*th power, one obtains the Mgf for χ_ν^2, where $\nu = n$. One can also calculate the Mgf of the gamma distribution and infer from a comparison that the CSD in Equation (3) is a special case of the gamma distribution [10].

In this work we provide yet another simplified derivation of the CSD using Laplace transform [23]. The Laplace transform of Equation (5) is equal to

$$\int_0^\infty \frac{\left(\chi_1^2\right)^{1/2-1} \mathrm{e}^{-\chi_1^2/2}}{2^{1/2}\Gamma\left(1/2\right)} \mathrm{e}^{-s\chi_1^2} \mathrm{d}\chi_1^2 = \left(\frac{1/2}{s+1/2}\right)^{1/2}. \tag{13}$$

Subsequently, we use a theorem that the Laplace transform of a *n*th convolution is a product of the individual transforms, *i.e.* $\left(\dfrac{1/2}{s+1/2}\right)^{n/2}$. By abbreviating $u = \chi_n^2$, the inverse Laplace transform results in the Pdf of *u*,

$$\text{Pdf}\left(u \mid n\right) = \frac{1}{2\pi i} \oint \left(\frac{1/2}{s+1/2}\right)^{n/2} \mathrm{e}^{su} \mathrm{d}s = \frac{1}{2^{n/2}} \frac{1}{2\pi i} \oint \frac{\mathrm{e}^{su}}{\left(s+1/2\right)^{n/2}} \mathrm{d}s. \tag{14}$$

To calculate the contour integral in Equation (14), we start with the Cauchy integration formula for an analytic function $f(s)$ of a complex variable *s* having a simple pole at s_0 [24]:

$$f\left(s_0\right) = \frac{1}{2\pi i} \oint \frac{f(s)}{s-s_0} \mathrm{d}s. \tag{15}$$

The $k-1$ times differentiation of Equation (15), where the differentiation can be of an integer or a fractional order [25], results in:

$$f^{(k-1)}\left(s_0\right) = \frac{\Gamma(k)}{2\pi i} \oint \frac{f(s)}{\left(s-s_0\right)^k} \mathrm{d}s. \tag{16}$$

By comparing Equation (14) to Equation (16), we infer that $f(s) = \mathrm{e}^{su}$, $s_0 = -1/2$, and $k = n/2$. By inserting these variables to Equation (16) and plugging it into Equation (14), we obtain:

$$\text{Pdf}\left(u \mid n\right) = \frac{1}{2^{n/2}\Gamma\left(n/2\right)} \left(\frac{\mathrm{d}^{n/2-1}}{\mathrm{d}s^{n/2-1}} \mathrm{e}^{su}\right)_{s=-1/2} = \frac{u^{n/2-1}\mathrm{e}^{-u/2}}{2^{n/2}\Gamma\left(n/2\right)}, \tag{17}$$

which is the CSD given by Equation (3) for $\nu = n$ and $\chi_n^2 = u$.

Another simplified derivation of the CSD uses the Bayesian inference and it is not related to the convolutions described above [11]. It uses a normal likelihood function for multiple samples. It also uses the transformational prior distributions: $\propto 1/\sigma$ for scale parameter σ and a constant for translation parameter μ [26]. Marginalizing the joint distribution (μ, σ) over μ results in the CSD, whereas marginalizing over σ results in the *t*-distribution [27].

In Section 5, we summarize the advantages and disadvantages of the simplified derivation methods of CSD described in this section.

3. Test Statistics

Several models for the CST statistics can be derived from the general Equation (2). For the expected value, we can use either the sample mean $\bar{x}$ or the population mean μ, whereas for the standard deviation we can use either individual standard deviations σ_i or the sample standard deviation σ_x. We do not know the population standard deviation for the data described in Section 4. Model test statistics $\sum\left(\left(x_i - \bar{x}\right)/\sigma_x\right)^2$ is always equal to $n-1$ and thus not useful. However, the model test statistics $\sum\left(\left(x_i - \bar{x}\right)/\sigma_i\right)^2$ can be used to test the variance. Other possibilities are to test for both the variance and location by employing model test statistics $\sum\left(\left(x_i - \mu\right)/\sigma_i\right)^2$ or $\sum\left(\left(x_i - \mu\right)/\sigma_x\right)^2$, if the population mean is known which is the case for the data in Section 4.

For the *t*-test we perform a standard one-sample test, where we calculate *t* variable as $\left(\bar{x} - \mu\right)/\left(\sigma_x/\sqrt{n}\right)$. The *t*-test is the location test. The results of all these test models using radioactivity data are presented in Section 4.

4. Chi-Square- and *t*-Test for Radioactivity Detection in Drinking Water

The most convenient method of measuring GA and GB radioactivity in drinking water is by gas proportional counting [28]. In this method, a given quantity of water is evaporated with nitric acid onto a stainless-steel planchet and dried, leaving a residue containing any radioactivity. The planchet is then counted on a gas proportional detector. Alpha and beta particles are counted simultaneously, and they are differentiated by much larger ionization caused by the former.

As stated in Section 1, this method must be able to determine GA and GB at the DL, to be verified by the CST [20] using a minimum of seven samples. EPA recommends a right-tail (RT) CST at 99% Confidence Level (CL), or 0.01 significance. To accomplish this, $n = 9$ samples of community drinking water were spiked with ^{230}Th and ^{90}Sr/^{90}Y radionuclides providing alpha and beta radioactivity, respectively. The spiking activities (*i.e.* the expected μ) were: 2.9888 ± 0.0402 pCi/L for alpha and 4.1860 ± 0.0549 pCi/L for beta, close to the required DL values. The values of spiking activities and their uncertainties were obtained from the standards traceable to the National Institute of Standards and Technology (NIST). Then the experimental procedure was followed, and the meas-

ured GA and GB activities x_i are depicted as points in **Figure 1** and **Figure 2**, respectively.

Also shown in **Figure 1** and **Figure 2** are the individual standard deviations σ_i, depicted as vertical lines. These standard uncertainties are propagated, including the Poisson statistics of radioactivity counting and background subtraction, uncertainties of the detector efficiency, cross-talk between alpha and beta particles, as well as solution-pipetting uncertainties. Therefore, they are slightly different for different samples.

The GA results are described first. The sample average for GA is given by $\bar{x} = 3.0951$ pCi/L (red horizontal thick line) which is close to the expected μ (green horizontal thick line) as seen in **Figure 1**. The sample standard deviation is given by $\sigma_x = 0.7000$ pCi/L. The results of the variance test, as defined in Section 3, are given in column 3 of **Table 1**. The number of the degrees of freedom is $\nu = 8$ because one constraint is from calculating the mean. The observed χ^2 statistics is equal to 14.0 for gross alpha. The right-tail (RT) and left-tail (LT) χ^2 are calculated from the CSD at 0.01 significance each. Since $1.6 < 14.0 < 20.1$, each tail test passes at 0.01 significance and two-tail (2T) test passes at 0.02 significance. Then, the two combined variance/location tests, as defined in Section 3 are given in columns 4 and 5 using σ_i and σ_x, respectively. $\nu = n = 9$ in these cases, because there are no constraints. They both pass for GA.

The *t*-test statistics is calculated as described in Section 3 resulting in 0.45 for GA, as given in column 6 in **Table 1**. The RT probability of 0.33 and 2T probability of 0.66 are larger than 0.01 and 0.02, respectively, ensuring the passage of the location *t*-test.

The gross beta activities plotted in **Figure 2**, with the mean $\bar{x} = 5.1274$ pCi/L (red horizontal thick line) and $\sigma_x = 0.3050$ pCi/L differ significantly from the expected μ (green horizontal thick line) beyond the observed uncertainties. That fact did not affect the variance test which passed for GB (column 3 in **Table 1**). However, the observed χ^2 of 43.1 and 93.7 exceed the calculated RT χ^2 of 21.7 (columns 4 and 5 in **Table 1**), therefore the combined variance/location tests failed. This failure is supported by the *t*-test, where the high $t = 9.26$ (column 6) resulted in very low values of the RT and 2T probabilities (columns 7 and 8) and failures of the test for GB.

To elucidate the reasons for failure of the GB CST and *t*-test, fifteen non-spiked Method Blank (MB) community water samples were prepared and measured. The average GA activity was below detection; however, the average GB was 0.8121 ± 0.2801 pCi/L. This MB was then subtracted from the spiked GB results and the corrected GB activities are plotted in **Figure 3**. The mean of the corrected GB is $\bar{x} = 4.3153$ pCi/L ($\sigma_x = 0.3050$ pCi/L), very close to the value for spiked radioactivity. The corrected observed χ^2 are now 2.7, 3.2 and 9.6 (columns 3, 4, and 5 in **Table 1**) ensuring the passage of the three CSTs. This is supported by the passage of the *t*-test also (columns 6, 7, and 8).

Table 1. The results of χ^2- and *t*-tests. Abbreviations: RT right-tail, LT left-tail, 2T two-tail. Significance is 0.01 for each tail.

| 1	2	3	4	5	6	7	8
		χ^2-test				*t*-test	
Experiment, reference	Parameter	Variance, σ_i	Variance and location, σ_i	Variance and location, σ_x		Location	
	Deg free	8	9	9	8		
	Calc RT	20.1	21.7	21.7			
	Calc LT	1.6	2.1	2.1	*t*	RT prob	2T prob
Gross Alpha, **Figure 1**	Observed	14.0	13.4	8.2	0.45	0.33	0.66
	Test result	Passed	Passed	Passed		Passed	Passed
Gross Beta, **Figure 2**	Observed	3.8	43.1	93.7	9.26	7.5E−06	1.5E−05
	Test result	Passed	Failed	Failed		Failed	Failed
Gross Beta-MB subtracted, **Figure 3**	Observed	2.7	3.2	9.6	1.27	0.12	0.24
	Test result	Passed	Passed	Passed		Passed	Passed
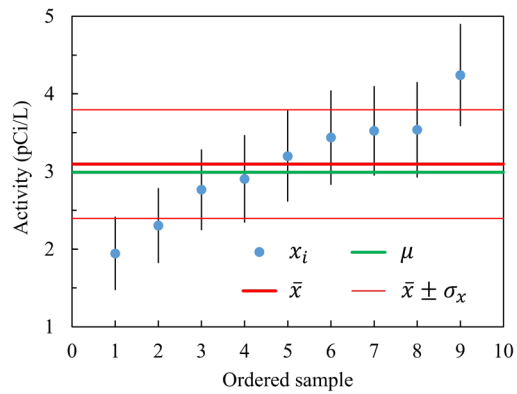
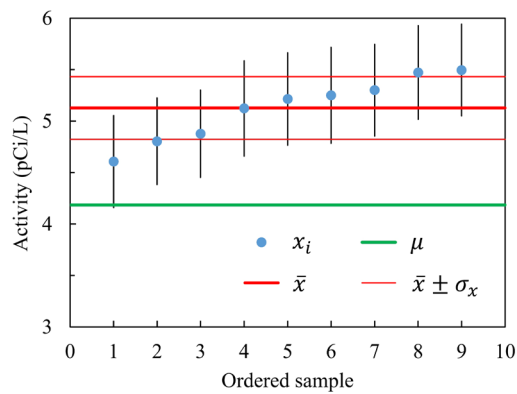**Figure 1.** Gross alpha (points) ordered according to the increased activity.

Figure 2. Gross beta (points) ordered according to the increased activity.
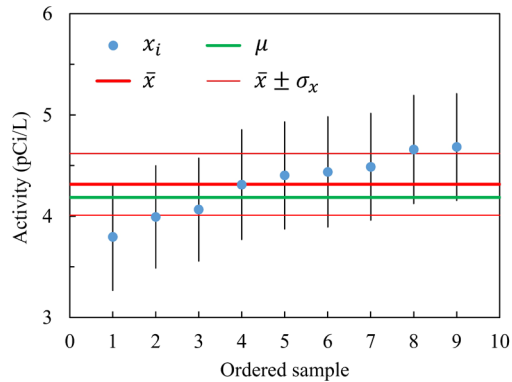
Figure 3. Gross beta (points) corrected for method blank and ordered according to the increased activity.

The reasons for the elevated GB in MB of community drinking water were investigated. Ten L of water were evaporated to 50 mL and measured using precise gamma-ray spectrometry [29]. It was determined that the concentration of the beta/gamma emitter, ^{40}K was 0.6926 ± 0.0790 pCi/L. It was also possible to identify several beta/gamma progenies of the ^{238}U series: ^{234}Th, ^{214}Pb, ^{214}Bi, and ^{210}Pb, as well as those from the ^{232}Th series: ^{228}Ac, ^{212}Pb, and ^{208}Tl. The combined activity of the beta/gamma progeny was 0.1513 ± 0.0672 pCi/L. Therefore, the sum of ^{40}K and beta/gamma progeny was 0.8440 ± 0.1037 pCi/L. The latter is consistent with the GB activity of 0.8121 ± 0.2801 pCi/L from the MB measurement to within the measured uncertainties. Also associated with the decay of ^{238}U and ^{232}Th is their alpha activity plus alpha progeny of similar activity to that of the beta/gamma progeny. This alpha activity could not have been detected by gamma spectrometry and was below the detection by GA in the MB measurement. However, the fact that GA of 3.0951 pCi/L is slightly higher than the expected 2.9888 pCi/L is an indication of that. Unlike in the case of beta activity, the small alpha progeny activity did not affect the CST or *t*-test. It should be noted that this level of naturally present radioactivity in the community water is much below the MCL, and thus poses small risk to the population.

5. Summary and Conclusions

We have described five simplified methods of deriving the chi-square distribution. Three of them: by convolution, moment generating function, and Bayesian inference are described in the literature and have been outlined here for comparison. The simplest of them seems to be the convolution method. It only uses the substitution from the normal distribution to a chi-square variable and requires a calculation of a single convolution integral on the above. It infers the form of multiple convolution on gamma distribution leading to the chi-square distribution. The moment generating function method of derivation is more advanced as it requires the knowledge of the moment generating function and the

gamma distribution. The Bayesian inference method requires the knowledge about likelihood function and prior probabilities but does not require the knowledge about the gamma distribution.

In this work, we have proposed two new methods for derivation of the chi-square distribution: by induction and by Laplace transform. The method of induction uses operational calculus with only a single integral leading to beta function. The proposed derivation applies modern formalism and seems to be simpler than the original derivation by Helmert as early as in 1876. A disadvantage of the induction method is that it requires a prior knowledge of the chi-square distribution to perform induction on it. There is a significant advantage, however. All other methods require either no constraints in the data; *i.e.* the number of degrees of freedom must be equal to the number of observations, or one constraint in case of Bayesian inference. The induction method leaves any constraints intact by adding one induction step to the existing number of degrees of freedom. The proposed derivation method by Laplace transform is more advanced because it uses integration in the complex plane. The significant advantage of the Laplace transform, and the Bayes inference methods is that they do not require prior knowledge about the gamma distribution.

We have also described a unique application of the chi-square test to environmental science. In chi-square testing, it is important to delineate systematic effects from the random uncertainties. In this work, a systematic natural contamination of laboratory method blank caused the chi-square test for combined variance/location to fail; however, it did not affect the chi-square test for variance alone. After subtracting the systematic method blank, the chi-square variance/location test was shown to have passed. This was confirmed by the location *t*-test. It is also imperative to perform analysis of uncertainty. In this work, using either individual or sample standard deviations did not affect the variance/location chi-square test. While the chi-square test provides verification if a laboratory test method is adequate to monitor gross alpha and gross beta radioactivity in drinking water, the test statistics combining variance and location is more useful than the one based on the variance alone because it can identify systematic bias.

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Hill, T.L. (1986) An Introduction to Statistical Thermodynamics. Dover Publications, New York, 122.

[2] Satchler, G.R. (1990) Introduction to Nuclear Reactions. Oxford U. P., New York, 248. https://doi.org/10.1007/978-1-349-20531-8

[3] Lancaster, H.O. (1969) The Chi-Squared Distribution. J. Wiley & Sons, New York, Chap. 1.

[4] Gorroochurn, P. (2016) Classic Topics on the History of Modern Mathematical Statistics: From Laplace to More Recent Times. J. Wiley & Sons, Hoboken, Chap. 3. https://doi.org/10.1002/9781119127963

[5] Bienaymé, I.-J. (1852) Sur la Probabilité des Erreurs d'Aprés la Méthode des Moindres Carrés. *Liouvillės Journal de Mathématiques Poures et Appliquées*, *Séries* 1, **17**, 33-78.

[6] Abbe, D.E. (1863) Ueber die Gesetzmässigkeit in der Vertheilung der Fehler bei Beobachtungsreihen. Dissertation, Jena.

[7] Helmert, F.R. (1876) Die Genauigkeit der Formel von Peters zur Berechnung des wahrscheinlichen Beobachtungsfehlers directer Beobachtungen gleicher Genauigkeit. *Astronomische Nachrichten*, **88**, 113-132. https://doi.org/10.1002/asna.18760880802

[8] Stuart, A. and Ord, K. (1994) Kendall's Advanced Theory of Statistics, Vol. 1, Distribution Theory. Arnold Hodder Headline Group, London, Chap. 11.

[9] Ross, S. (2006) A First Course in Probability. Pearson Prentice Hall, Upper Saddle River, Sec. 6.3.

[10] Berry, D.A. and Lindgren, B.W. (1996) Statistics: Theory and Methods. Wadsworth Publishing, Belmont, Sec. 5.12, 6.4.

[11] Gull, S.F. (1988) Bayesian Inductive Inference and Maximum Entropy. In: Erickson, G.J. and Smith, C.R., Eds., *Maximum-Entropy and Bayesian Methods in Science and Engineering*, Kluwer Academic, Dordrecht, 53-74. https://doi.org/10.1007/978-94-009-3049-0_4

[12] Helmert, F.R. (1876) Ueber die Wahrscheinlichkeit der Potenzsummen der Beobachtungsfehler und über einige damit im Zusammenhange stehende Fragen. *Zeitschrift für Mathematik und Physik*, **21**, 192-218.

[13] Pearson, K. (1900) On the Criterion that a Given System of Deviations from the Probable in the Case of Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling. *Philosophical Magazine Series* 5, **50**, 157-175. https://doi.org/10.1080/14786440009463897

[14] Greenwood, P.E. and Nikulin, M.S. (1996) A Guide to Chi-Squared Testing. J. Wiley & Sons, New York, Sec. 3.18.

[15] Fisher, R.A. (1922) On the Interpretation of χ^2 from Contingency Tables and the Calculation of P. *Journal of the Royal Statistical Society A*, **85**, 87-94. https://doi.org/10.2307/2340521

[16] Evans, R.D. (1985) The Atomic Nucleus. Krieger Publishing, Malabar, Chap. 27.

[17] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995) Continuous Univariate Distributions, Vol. 2. J. Wiley & Sons, New York, Chap. 28.

[18] Eisenbud, M. and Gesell, T. (1997) Environmental Radioactivity from Natural, Industrial, and Military Sources. Academic Press, San Diego, Chap. 6. https://doi.org/10.1016/B978-012235154-9/50010-4

[19] Environmental Protection Agency (2000) 40 CFR Parts 9, 141, and 142 National Primary Drinking Water Regulations; Radionuclides; Final Rule. *Federal Register*, **65**, 76708-76752.

[20] EPA (2017) Procedure for Safe Drinking Water Act Program Detection Limits for Radionuclides. Report EPA 815-B-17-003, Cincinnati.

[21] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994) Continuous Univariate Distributions, Vol. 1. J. Wiley & Sons, New York, Chap. 12, 17, 18.

[22] Johnson, N.L., Kemp, A.W. and Kotz, S. (2005) Univariate Discrete Distributions. J. Wiley & Sons, Hoboken, Chap. 1. https://doi.org/10.1002/0471715816

[23] Margenau, H. and Murphy, G.M. (1976) The Mathematics of Physics and Chemistry. Krieger Publishing, Huntington, Sec. 8.5.

[24] Dettman, J.W. (1965) Applied Complex Variables. Dover Publications, New York, Sec. 3.6.

[25] Oldham, K.B. and Spanier, J. (1974) The Fractional Calculus: Theory and Applications of Differentiation and Integration to Arbitrary Order. Dover Publications, Mineola, Sec. 3.4.

[26] Jaynes, E.T. (2004) Probability Theory: The Logic of Science. Cambridge U. P., Cambridge, Chap. 12. https://doi.org/10.1017/CBO9780511790423

[27] Student (1908) The Probable Error of a Mean. *Biometrika*, **6**, 1-25.
https://doi.org/10.1093/biomet/6.1.1

[28] Semkow, T.M. and Parekh, P.P. (2001) Principles of Gross Alpha and Beta Radioactivity Detection in Water. *Health Physics*, **81**, 567-574.
https://doi.org/10.1097/00004032-200111000-00011

[29] Khan, A.J., Semkow, T.M., Beach, S.E., Haines, D.K., Bradt, C.J., Bari, A., Syed, U.-F., Torres, M., Marrantino, J., Kitto, M.E., Menia, T. and Fielman, E. (2014) Application of Low-Background Gamma-Ray Spectrometry to Monitor Radioactivity in the Environment and Food. *Applied Radiation and Isotopes*, **90**, 251-257.
https://doi.org/10.1016/j.apradiso.2014.04.011

Appendix

A.1. Glossary

CL: Confidence Level

CSD: Chi-Square Distribution

CST: Chi-Square Test

DL: Detection Limit for radionuclides

EPA: U.S. Environmental Protection Agency

GA: Gross Alpha Radioactivity

GB: Gross Beta Radioactivity

L: Liter

LT: Left Tail

MB: Method Blank

mBq: milli-Becquerel

MCL: Maximum Contaminant Level

MCLG: Maximum Contaminant Level Goal

Mgf: Moment generating function

mL: milli-Liter

mrem: milli-rem

NIST: National Institute of Standards and Technology

pCi: pico-Curie

Pdf: Probability density function

RT: Right Tail

SDWA: Safe Drinking Water Act

STEM: Science, Technology, Engineering and Mathematics

y: year

μSv: micro-Sievert

2T: Two Tail

A.2. Variables

a, b: parameters of the gamma distribution

B: beta function

E: expectation value

E_i: expected frequency

$f(s)$: analytic function

gamma: gamma distribution

i, k: indices

m: number of categories

n: number of observations

O_i: observed frequency

p: number of parameters for model distribution

s: complex variable

s_0: pole

t: t-test variable

Var: variance

x_i : normal random variable

$\bar{x}$: sample mean

u, z: substituted variables

Γ : gamma function

μ, μ_i : expected variable: population, individual

ν : number of degrees of freedom

$\sigma, \sigma_i, \sigma_x$: standard deviation, individual, sample

φ_i^2 : individual chi-square

$\chi^2, \chi_i^2, \chi_n^2, \chi_\nu^2$: chi-square, for *i, n* observations, ν degrees of freedom