Quest: Journal of Geometry, Mathematical and Quantum Physics
e-ISSN: 3064-7282 Vol.3 Issue 3|2026

Guide to the Chi- Square in Fluid Mechanics

Batool A. Musawi
Department of Statistical Techniques, Institute of Management/ Baghdad, Middle Technical University, Baghdad, Iraq

Abstract: The chi-square distribution, a fundamental element of statistical theory, naturally emerges from the analysis of sums of squared Gaussian variables and finds broad application across scientific disciplines. In fluid mechanics, its significance is particularly evident in turbulence, boundary-layer dynamics, and energy fluctuation studies, where velocity components and kinetic energy distributions often conform to chi-square or related gamma laws. This statistical framework provides a rigorous basis for quantifying variability, testing hypotheses, and validating experimental observations in complex flow systems. Furthermore, the chi-square test serves is a practical tool to assessing goodness-of-fit between measured fluid dynamic data and theoretical models, thereby enhancing the reliability of empirical investigations. By integrating statistical inference with physical modeling, chi-square methods deepen our understanding of randomness in fluid motion and support the development of more robust predictive models in fluid mechanics.

Keywords: Mathematical Induction, Gamma Distribution, Chi-Square Test, Gross Alpha-Beta, Pearson chi-square test, difference between two proportions, goodness of fit, contingency tables.

INTRODUCTION

The chi-square test is one of the most important non-parametric statistical tests [1]. It is used to study the relationship between categorical variables or to check whether the data distribution conforms to the expected distribution see (Table 1). This test is characterized by its ease of application and widespread use in social, marketing, medical, and scientific research The formal origins of the chi-square test can be traced back to the renowned British statistician Carl Pearson [2, 3], who first published it in 1900 in his influential paper entitled "On the Criterion that a certain system of deviations from probability in the case of a system of correlated variables may reasonably be assumed to have arisen from a random sample." Prior to Pearson's work, there had been early attempts to use similar measures, but it was Pearson who developed the precise mathematical basis for the probability distribution of the chi-square, allowing for a systematic and reproducible determination of the degree of correspondence between theoretical and experimental data, Before Karl Pearson [4], most statistical analyses relied heavily on quantitative data and the normal distribution (Gaussian distribution). The introduction of the chi-square test revolutionized applied statistics, providing for the first time a reliable method for dealing with categorical data that do not follow a normal distribution or that do not satisfy the assumptions of parametric tests. This advance enabled researchers in diverse fields, such as genetics (where pioneering scientists like Mendel used it to analyze their results) [5] and the social sciences, to draw accurate statistical inferences about distributions and correlations that were previously considered merely descriptive or subjective. Thus, Pearson established the chi-square test as a cornerstone of modern inferential statistics, profoundly influencing research methodologies in the 20th century and beyond, in

general it is a statistical test used to determine whether there is a statistically significant correlation or difference between two categorical variables, or whether the data conforms to a specific distribution. It is an effective statistical tool for studying the relationship between categorical variables or for testing the conformity of data to the expected distribution. By applying it correctly and following its conditions, researchers can draw accurate conclusions and support data-driven decision-making [6].

Test Type of chi- square	Purpose	Example in Research
Goodness-of-Fit Test	Checks if the distribution of one categorical variable matches expected proportions.	Testing if a dice is fair (equal probability for each side).
Test of Independence	Examines whether two categorical variables are related.	Studying if gender is associated with voting preference
Homogeneity Test	Compares distributions across different populations.	Comparing customer satisfaction levels across multiple stores.

Table 1: A table showing the Types of chi-square tests.

So the conditions for using the Chi-Square Test are the variables must be qualitative or categorical and the sample size must be sufficient to ensure the accuracy of the results, finally the expected values for each cell should be greater than 5 in most cases [7], Therefore, after understanding the advantages of the test, it is necessary to point out some common mistakes that the researcher may make, when using the chi-square test include: using it with quantitative variables instead of qualitative ones; ignoring sample size requirements or expected values; interpreting results without considering the level of statistical significance; relying solely on the test without considering the graph or the scientific context. Over the following decades, the chi-square test was extended to applications in contingency tables (testing independence between categorical variables), pioneered by statisticians such as Ronald Fisher in the 1920s [8]. Its adaptability made it central to fields ranging from biology and genetics (e.g., Mendelian inheritance studies) to social sciences, medicine, and engineering.

The chi-square test has become one of the most widely utilized statistical methods for the analysis of categorical data across numerous research domains. Its applications extend to the social sciences, biomedical research, economics, education, and marketing, where it is employed to examine associations between variables and to assess the consistency of observed outcomes with theoretical expectations. This broad adoption underscores its central role in empirical investigation and statistical reasoning, highlighting its importance as a versatile tool for validating hypotheses and strengthening the reliability of research findings.

As noted in [9], the chi-square test is second only to the Student's t-test in frequency of use. Nevertheless, despite its popularity, several scholars [10, 11, 12] have emphasized methodological shortcomings, including unproven bounds on sample size requirements and the assumption of independence among observations. In this work, we illustrate the test's pronounced sensitivity to both sample size and data representation. In particular, we demonstrate that scaling the concordance matrix by a constant c results in the test statistic being scaled by the same factor, which makes it possible to manipulate the outcome—either rejecting or failing to reject the null hypothesis H_0. Comparable issues have recently been identified in analysis of variance (ANOVA) and in Tukey-Kramer's procedure. Specifically, Tukey-Kramer's formula [13] for determining the critical range between observation sets i and j is influenced not only by these two sets but also by all other sets, even when they are unrelated. This dependency introduces logical inconsistencies in ANOVA when more than two groups are involved, for detailed definitions and discussion. The chi-squared (CSD) distribution is one of the most widely used distributions in science. Its importance in fundamental physics is undeniable, for example in the kinetic energy distribution of

particles in an ideal gas, as described by the Maxwell–Boltzmann law [14], or in the energy spectra of particles emitted from excited nuclei during nuclear reactions [15]. Chi-square tests play a pivotal role in fluid mechanics, particularly in the statistical validation of experimental data, model fitting, and uncertainty analysis [16, 17]. Researchers use this test to verify whether observed fluid behavior, such as turbulence measurements, velocity distributions, and particle tracking data, matches theoretical predictions or simulation outputs. Error & Uncertainty Analysis: Chi-square statistics quantify how well measured data fit expected models, especially in noisy environments such as turbulence or multiphase flows. Random Variable Analysis: Fluid mechanics often deals with stochastic processes (e.g., turbulence intensity, droplet breakup). Chi-square distribution is used to test hypotheses about these random variables. So take an example, CFD Validation that compares simulation outputs with lab data using chi- square goodness of fit to verifying Navier-Stokes solver predictions.

How It Works: (Mathematical formulas for chi- square)

Section 1:

Pearson's tests [18] (goodness of fit test, independence test, and ratio homogeneity test) can be organized according to the same basic formula for the test statistic in a single form (Table 2). For example, consider a scenario where a T result from polynomial experiments is classified according to two different criteria, within one of the same m categories, in one of the n groups. The results can be presented in a two-way consistency table consisting of m rows and n columns.

	Group 1	Group 2	Group 3	……	Group j	……	Group n	Row total
Cate. 1	O_{11}	O_{12}	O_{13}	…….	O_{1j}		O_{1n}	$O_{1.}$
Cate. 2	O_{21}	O_{22}	O_{23}	…….	O_{2j}		O_{2n}	$O_{2.}$
. . .				……..	……..			
Cate. i	O_{i1}	O_{i2}	O_{i3}	……..	O_{ij}		O_{in}	$O_{i.}$
. . .				……	O_{i3}			
Cate. m	O_{m1}	O_{m2}	O_{m3}	……	O_{mj}		O_{mn}	$O_{m.}$
Column Total	$O_{.1}$	$O_{.2}$	$O_{.3}$	……..	$O_{.j}$		$O_{.n}$	T

Table 2: A table showing the general structure of an $m \times n$ order compatibility matrix, s.t. O_{ij} refers to the observant-ion in the cell located in row i and column j.

For example, Consider a region with a population of one million, comprised of four neighborhoods: A, B, C, and D. A random sample of 700 women from this region was selected to study the impact of science on social life, and their occupations were recorded as either "employees," "homemakers," or "students." The null hypothesis states that each person's residential neighborhood is independent of their occupational classification. The data are tabulated as follows, show table 3 [19]:

	A	B	C	D	Total
employees	140	80	154	125	499
homemakers	80	70	101	50	301
students	50	90	95	65	300
	270	240	350	240	1100

Table 3: A table showing the Distribution of population density divisions.

Section 2: (the hypotheses)

1. **Formulate Hypotheses**

- Null hypothesis (H_0): No relationship or difference exists.
- Alternative hypothesis (H_a): A relationship or difference exists.

2. **Calculate Expected Frequencies**

- Based on the assumption of independence or equal distribution.

3. **Compute Chi-Square Statistic** $X^2 = \sum(O - E)2$

Let, $R_i = \sum_{i=1}^{n} O_{ij}$, $i \in \{1,2,\dots\dots,m\}$ it is row sum for population i, $C_j = \sum_{j=1}^{m} O_{ij}$, $j \in \{1,2,\dots\dots,n\}$ it is column sum for category and $T = \sum_{i=1}^{m}\sum_{j=1}^{n} O_{ij}$ (total sample size). The null hypothesis is:

$H_0: \pi_{1j} = \pi_{2j} = \dots\dots\dots = \pi_{mj} \qquad \forall\, j \in \{1, \dots\dots, n\}.$

Where, $\pi_{ij} \in [0,1]$ is the proportion of indivduals in group I falling into category and $\sum_{i=1}^{n} \pi_{ij} = 1$ for each *i* under H_0 the expected frequency E_{ij} in cell (i,j) is given by:

$$E_{ij} = \frac{R_i \,.\, C_j}{N}, \quad \forall\, i = 1 \dots\dots m \; and \;\; j = 1 \dots\dots n$$

So the test statistic define as $X^2{}_{stat} = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$ (1)

Under the test assumptions the test statistic $X^2{}_{stat}$ approxomateiy follows a chi – square distribution

with (m-1)(n-1) degrees of freedom under H_0 and let $\alpha \in (0,1)$ be the significance level. Then the decision rule is:

$Reject \;\; H_0 \;\; if \;\; X^2{}_{stat} > X^2{}_{1-\alpha\,,(m-1)(n-1)}$

Where $X^2{}_{1-\alpha\,,(m-1)(n-1)}$ is the chi – square distribution with (m-1)(n-1) degrees of freedom.

Theorem:

It is clear that for any real constant $c \geq 0$, the relation $X^2{}_{stat}\ (cA) = c\ X^2{}_{stat}\ (A)$ holds. The nullhypothesis H_0 stated is rejected at agiven significance level α, if and only if $(\Leftrightarrow)$ $X^2{}_{stat}\ (A) > X^2{}_{crit}\ (A)$, where $X^2{}_{crit}\ (A) = X^2{}_{1-\alpha\,,(m-1)(n-1)}$ represents the (1−α)-quantile of the chi-square distribution with (m−1)(n−1) degrees of freedom, This critical value depends solely on m, n, and α, but is independent of the specific entries of matrix A.

Remark:

The following five conditions are mutually equivalent [20]:

(a) the rows of A are perfectly proportional;

(b) the columns of A are perfectly proportional;

(c) $O_{ij} = E_{ij} \; for \; all \;\; i = 1 \dots\dots m \; and \; j = 1 \dots\dots n;$

(d) The chi-square statistic satisfies $X^2{}_{stat}\ (A) = 0;$

(e) For any positive real constant c, we have $X^2{}_{stat}\ (cA) = 0.$

Moreover, if none of the above conditions$(a - e)$ hold, then $X^2{}_{stat}\,(A) > 0$. In this case, for a fixed significance level α, the null hypothesis H_0 will be rejected for the scaled matrix cA with

confidence 1−α when c is sufficiently large, while it will not be rejected when c is sufficiently small.

Proof:

Returning to the equation (1), the numerator and denominator of this equation represent quadratic and linear functions of the matrix components of A and $X^2{}_{stat}\ (cA) = cX^2{}_{stat}(A)$. So the proportionality test of the rows cA should not linearly depend on c.

Let's walk through a numerical chi-square example in fluid mechanics using turbulence velocity distribution data. Suppose we are studying turbulent velocity fluctuations in a wind tunnel. Theory suggests that the velocity fluctuations follow a normal distribution. We collect 100 velocity samples and categorize them into bins. We group the velocity fluctuations (normalized by mean velocity) into categories: see table 4

Velocity rang ($\acute{u}/_{U}$)	Observed Frequency (O)
-4 to -2	12
-2 to 0	28
0 to 2	40
2 to 4	20
	Total =100

Table 4: A table showing the Distribution of the velocities.

If the distribution is normal, the expected probabilities for each bin can be calculated from the standard normal distribution.

- Probability $(-4 < z < -2) \cong 0.135$
- Probability $(-2 < z < 0) \cong 0.341$
- Probability $(0 < z < 2) \cong 0.341$
- Probability $(2 < z < 4) \cong 0.135$

Multiply each probability by 100 (sample size), see table 5,

Velocity rang ($\acute{u}/_{U}$)	Observed Frequency (E)
-4 to -2	13.5
-2 to 0	34.1
0 to 2	34.1
2 to 4	13.5

Table 5: A table showing the Distribution of velocities.

By the formula:

$$x^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \quad (2)$$

Compute for each bin:

- Bin 1: $\frac{(12-13.5)^2}{13.5} = \frac{2.25}{13.5} \approx o.167$
- Bin 2: $\frac{(28-34.1)^2}{34.1} = \frac{37.21}{34.1} \approx 1.09$
- Bin 3: $\frac{(40-34.1)^2}{34.1} = \frac{34.51}{34.1} \approx 1.02$
- Bin 4: $\frac{(20-13.5)^2}{13.5} = \frac{47.25}{13.5} \approx 3.13$

$$x^2 = o.167 + 1.09 + 1.02 + 3.13 = 5.41$$

$df = number\ of\ categories - 1 - parameters\ estimated$, our categories = 4 and the parameters estimated (mean, variance) = 2, the degrees of freedom will be:

$df = 4 - 1 - 2 = 1$

Critical value at $(df = 1, \alpha = 0.05) \rightarrow 3.84, and\ our\ X^2 = 5.41$

So, we will reject null H_0 therefore, the turbulence velocity distribution does not perfectly follow a normal distribution in this experiment. This example shows how chi-square can be applied in fluid mechanics to validate turbulence models. Even through theory suggests Gaussian fluctuations, experimental data may deviate. Now, let's take another example of CFD Validation of Turbulent Jet Velocity, if we have:

- Experimental data: Measured velocity distribution at a cross-section of a turbulent jet.
- CFD simulation data: Predicted velocity distribution from a turbulence model (say, k−ϵ). We want to test if the CFD predictions statistically agree with the experimental data using a chi-square goodness-of-fit test.

Now, If we suppose we divide the jet velocity profile into 5 bins (normal velocity ranges) as follows:

Velocity Range (u/U)	Experimental Frequency (Observed, O)	CFD Prediction (Expected, E)
0.0 – 0.2	**18**	**20**
0.2 – 0.4	**22**	**24**
0.4 – 0.6	**30**	**28**
0.6 – 0.8	**20**	**18**
0.8 – 1.0	**10**	**9**
		T=100

Table 6: A table showing the Distribution of normal velocities ranges.

With eq. 2 of the formula $x^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$, we can compute for each bin:

- Bin 1: $\frac{(18-20)^2}{20} = \frac{4}{20} \approx 0.20$
- Bin 2: $\frac{(22-25)^2}{25} = \frac{9}{25} \approx 0.36$
- Bin 3: $\frac{(30-28)^2}{28} = \frac{4}{28} \approx 0.14$
- Bin 4: $\frac{(20-18)^2}{18} = \frac{4}{18} \approx 0.22$
- Bin 5: $\frac{(10-9)^2}{9} = \frac{1}{9} \approx 0.11$

$$x^2 = 0.20 + 0.36 + 0.14 + 0.22 + 0.11 = 1.03$$

$df = number\ of\ categories - 1$, Categories = 5 and $df = 5 - 1 = 4$.So the critical value at $(df = 4, \alpha = 0.05) \rightarrow 9.49, and\ our\ X^2 = 1.03 < 9.49$

So, it is fail to reject null H_0 therefore, the CFD simulation matches the velocity distribution well. In this way we shows how chi- square in CFD validation:

- Low chi- square values give us CFD predication are statistically consistent with experiments.

- High chi- square values give us CFD model may need refinement e.g. turbulence model choice, mesh resolution).
- It is quantitative way to validate CFD models beyond just visual comparison of plots.

In this study, we introduce a simplified technique for deriving the composite power density function (CSD). The proposed approach is based on mathematical induction. While Helmert [21] initially obtained the CSD through a two-step forward induction process, our refinement reveals an inherent iterative characteristic of the function. This property enables the derivation to be carried out through a single-step induction, supported by the fundamental theorem involving the beta and gamma functions.

Section 3: (Chi-Square Distribution)

The probability density function (PDF) of the composite power density (CSD) is expressed as follows:

$$pdf\ (X_v^2|\ v) = \frac{(X_v^2)^{v/2-1} e^{-X_v^2/2}}{2^{v/2}\Gamma(v/2)}\ ,\ \ \Gamma \text{ (gamma function) (3)}$$

Here, CSD $E[X^2] = v$ (the expectation value), Var $[X^2] = 2v$ (the variance) [21]. The CSD is a special case of the gamma distribution abbreviated as gamma $(X_v^2|a,b)$ with the parameters $a = v/2$, and b = 2 [22].

Let $X_v^2 = \sum_{i=1}^{n} \frac{(X_i-\mu_i)^2}{\sigma_i}$ be the statistics with normal variable x_1; $x_1 \in [x_1, x_1 + dx_1]$ and its pdf is given by $(bdf(x_1)dx_1 = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\left[\frac{x_1-\mu_1}{\sigma_1}\right]^2/2} dx_1)$. If we put $X_1^2 = ((x_1-\mu_1)/\sigma_1)^2$, eq. (3) will be obtain as:

$$pdf(X_1^2|1)dX_1^2 = \frac{2}{\sqrt{2\pi}\sigma_1} e^{-X_1^2/2} \left|\frac{dx_1}{dX_1^2}\right| dX_1^2 = \frac{(X_1^2)^{1/2-1} e^{-\frac{X_1^2}{2}}}{2^{\frac{1}{2}}\Gamma(1/2)} dX_1^2 = gamma(X_1^2 \mid 1/2, 2)\, dX_1^2. \text{ (4)}$$

Now for $v = 1$ and used $\Gamma = (1/2) = \sqrt{\pi}$, The factor of two arises from the substitution of the variable x_1, which originally ranged from $-\infty\ to\ +\infty$, with the squared variable X_1^2, whose domain extends from $0\ to\ +\infty$. Now, let us assume that an additional term, corresponding to the normal variable x_{n+1}, is appended to $X_v^2 = \sum_{i=1}^{n} \frac{(X_i-\mu_i)^2}{\sigma_i}$. This inclusion increases the number of degrees of freedom by one $v + 1$, thereby modifying the structure of the distribution accordingly. So $X_{v+1}^2 = X_v^2 + (\frac{X_{n+1}-\mu_{n+1}}{\sigma_{n+1}})^2$, and by using the calculus for probability density functions [21],

$$pdf(X_{v+1}^2|v+1)dX_{v+1}^2 = \int_{-\infty}^{\infty} pdf\ (X_v^2|v)dX_v^2\ pdf\ (X_{n+1})dX_{n+1} \text{ (5)}$$

Let 's define a new variable z, such as, $(\frac{X_{n+1}-\mu_{n+1}}{\sigma_{n+1}})^2 = X_{n+1}^2(1-z)$ and $dX_{v+1}^2 = dX_v^2$, therefor the right side of Eq. 5 can be rewritten as

$$2\int_0^1 pdf\ (X_v^2|v)pdf\ (X_{n+1}) \left|\frac{dX_{n+1}}{dz}\right| dz\ dX_{n+1}^2 = \frac{(X_{n+1}^2)^{v-1/2-1} e^{-X_{v+1}^2/2}}{2^{(v+1/2}\Gamma(v/2)\Gamma(1/2)} dX_{v+1}^2 \int_0^1 z^{v/2-1}(1-z)^{1/2-1} dz$$

(6)

Now, the right side of Eq. 6 is the beta function, $\beta(v/2, 1/2)$, can be given as:[22]

$$\beta(v/2, 1/2) = \frac{\Gamma(v/2)\Gamma(1/2)}{\Gamma((v+1)/2)} \text{ (7)}$$

By substituting Eq. (7) into Eq. (6), simplifying the resulting expression, and then comparing it with the left-hand side of Eq. 5, one obtains the following relationship:

$$pdf(\,X_{v+1}^2\,|v+1) = \frac{(X_{v+1}^2)^{(v+1)/2-1}\; e^{-X_{v+1}^2/2}}{2^{(v+1)/2}\,\Gamma((v+1)/2)} \quad (8)$$

This result corresponds to the pdf. given in Eq.3, for $v+1$ degrees of freedom, thereby establishing Eq. 3, through induction. By substituting $\varphi_i^2 = ((x_i-\mu_i)/\sigma_i)^2$, transforms into $X_v^2 = \sum_{i=1}^{n} \varphi_i^2$, which represents the generalized chi-square form with parameters μ_i and σ_i.

The sum of independent random variables φ_i^2 is expressed as a convolution, and the joint distribution function for X_v^2 can be obtained by evaluating an n-dimensional convolution integral. Investigating the properties of this convolution leads to useful simplifications that have been exploited in prior studies. In particular, by convolving two gamma distributions $X_1^2 = \varphi_i^2$ from Eq. 5, and applying the theorem that the convolution of two gamma distributions yields another gamma distribution, one obtains: $(X_2^2\,|2/2,2)$gamma.

By continuing this process of convolution with X_1^2, it follows directly that the complete convolution is given by$(X_v^2\,|v/2,2)$, $v = n$, which is exactly the composite power density (CSD) expressed in Equation (3). This provides a simplified derivation of the CSD using convolution.

An alternative simplification relies on the theorem that the moment generating function (MGF) of a convolution is equal to the product of the individual MGFs [23]. Thus, by calculating the MGF of X_1^2 from Eq. 4, and raising it to the n-th power, one obtains the MGF of X_v^2, where $v = n$. Comparing this with the known MGF of the gamma distribution shows that the CSD in Eq. 3, is a special case of the gamma distribution [24]. In this work, we present yet another simplified derivation of the CSD using the Laplace transform [25], which serves as the foundation for extending the result to the general case of

$$\int_0^{\infty} \frac{(X_1^2)^{1/2-1} e^{-X_1^2/2}}{2^{1/2}\,\Gamma(1/2)}\; e^{-sX_1^2}\, dX_1^2 = \left(\frac{1/2}{s+1/2}\right)^{1/2} \quad (9)$$

Now, by used the Laplace transform of a nth i.e. $\left(\frac{1/2}{s+1/2}\right)^{n/2}$, with $u = X_n^2$, the inverse Laplace transform results in the Pdf of u, this yields

$$pdf(u|n) = \frac{1}{2\pi i}\oint \left(\frac{1/2}{s+1/2}\right)^{n/2} e^{su}\, ds = \frac{1}{2^{n/2}}\frac{1}{2\pi i}\oint \frac{e^{su}}{(s+1/2)^{\frac{n}{2}}}\, ds \quad (10)$$

To calculate the contour integral in Eq. (10), If we start with the Cauchy integration formula for an analytic $f(s)$ of a complex variable s having a simple pole at s_0

$$f(s_0) = \frac{1}{2\pi i}\oint \frac{f(s)}{s-s_0}\, ds \quad (11)$$

The k −1 times differentiation of Eq. (11), where the differentiation can be of an integer or a fractional order [25], results in:

$$f^{(k-1)}\,(s_0) = \frac{\Gamma(k)}{2\pi i}\oint \frac{f(s)}{(s-s_0)^k}\, ds \quad (12)$$

By comparing Eq. (10) to Eq. (12), we infer $f(s) = e^{su}$,$s_0 = -1/2$, and $k = n/2$. By inserting these variables to Eq. (12) and plugging it into Eq. (10), we obtain:

$$pdf(u|n) = \frac{1}{2^{n/2}\Gamma(n/2)}\left[\frac{d^{\,n/2-1}}{ds^{\,n/2-1}}\, e^{su}\right]_{su-1/2} = \frac{u^{n/2-1}e^{-u/2}}{2^{n/2}\,\Gamma\left(^n/_2\right)} \quad (13)$$ This is the CSD that given by eq. (3) $v = n$ and $X_n^2 = u$.

Section 4: (examples)

If we consider the following 2×2 information see table 7:

	Group 1 Group 2	Row Total
Category A	**1 1**	**2**
Category B	**1 11**	**12**
Column Total	**2 12**	**14**

Table 7: A table showing 2×2 contingency matrix.

If we start the null hypothesis that the proportions of outcomes are the same across the two groups. The expected value for each cell number the null hypothesis are:

$$E_{11} = \frac{2 \times 2}{14} = \frac{4}{14} = \frac{2}{7}$$

$$E_{12} = \frac{2 \times 12}{14} = \frac{24}{14} = \frac{12}{7}$$

$$E_{21} = \frac{12 \times 2}{14} = \frac{24}{14} = \frac{12}{7}$$

$$E_{22} = \frac{12 \times 12}{14} = \frac{144}{14} = \frac{72}{7}$$

And the chi- square statistic is found to be:

$$X^2_{stat} = \frac{(1-\frac{2}{7})^2}{\frac{2}{7}} + \frac{(1-\frac{12}{7})^2}{\frac{12}{7}} + \frac{(1-\frac{12}{7})^2}{\frac{12}{7}} + \frac{(11-\frac{72}{7})^2}{\frac{72}{12}} = \frac{175}{72}$$

With the degree of freedom =1, compare this result with the critical value from the chi- square distribution table at significance level $\alpha = 0.05$, which is approximately 3.841,we have that:

$$X^2_{stat} \approx 2.43 < 3.841$$

Thus we fail to reject the null hypothesis. There isn't enough evidence to suggest a significant difference in proportions between the two groups. But if we change the last information to modify it see table 8:

	Group 1 Group 2	Row Total
Category A	**2 2**	**4**
Category B	**2 22**	**24**
Column Total	**4 24**	**28**

Table 8: A table showing 2×2 modify contingency matrix.

If we start the null hypothesis that the proportions of outcomes are the same across the two groups. The expected value for each cell number the null hypothesis are:

$$E_{11} = \frac{4 \times 4}{28} = \frac{16}{28} = \frac{4}{7}$$

$$E_{12} = \frac{4 \times 24}{28} = \frac{96}{28} = \frac{24}{7}$$

$$E_{21} = \frac{24 \times 4}{28} = \frac{96}{28} = \frac{24}{7}$$

$$E_{22} = \frac{24 \times 24}{28} = \frac{576}{28} = \frac{144}{7}$$

And the chi- square statistic is found to be:

$$X^2_{stat} = \frac{(2-\frac{4}{7})^2}{\frac{4}{7}} + \frac{(2-\frac{24}{7})^2}{\frac{24}{7}} + \frac{(2-\frac{24}{7})^2}{\frac{24}{7}} + \frac{(22-\frac{144}{7})^2}{\frac{144}{12}} = \frac{175}{36}$$

With the degree of freedom =1, compare this result with the critical value from the chi- square distribution table at significance level $\alpha = 0.05$, which is approximately 3.841,we have that

$$X^2_{0.05,1} \approx 4.86 < 3.841$$

Thus we have to reject the null hypothesis. There is the sufficient evidence to suggest the proportions differ between the two groups.

Conclusion:

1. Two methods for calculating chi-squared have been discussed, including the simplified method for deducing the chi-squared distribution and the moment generation function, and have been summarized here for comparison.
2. The method of deriving the moment generation function is more advanced than some other methods, because it requires knowledge of the moment generation function and the gamma distribution, unlike the Bayesian inference method, which requires knowledge of the probability function and the prior probabilities, without knowledge of the gamma distribution.
3. The chi-square test provides a precise framework for testing independence and fit hypotheses by measuring the differences between observed and predicted frequencies, making it one of the most important basic statistical methods for analyzing categorical data.
4. The chi-squared distribution is traditionally used in statistics to analyze categorical data, and also in fluid mechanics to study stochastic processes and energy distributions. In turbulent flows, the chi-squared distribution is particularly useful because particle velocities and kinetic energies often follow probability laws that can be derived from the sum of squares of Gaussian variables—the basis of the chi-squared distribution.
5. Broadening this framework, chi-squared methods enable researchers to validate experimental measurements, assess the consistency of observed flow data with theoretical probability models, and evaluate the agreement between computational fluid dynamics (CFD) simulations and experimental results. In this way, chi-squared analysis helps bridge the gap between statistical inference and physical modeling, enhancing the reliability of fluid mechanics research and contributing to the development of robust predictive models for complex flow systems.

Reference:

1. Hill, T.L. (1986) An Introduction to Statistical Thermodynamics. Dover Publications, New York, 122.
2. K. Pearson: *Mathematical contributions to the theory of evolution — On a form of spurious correlation which may arise when indices are used in the measurement of organs*, Proc. R. Soc. Lond., 60 (359-367) (1897), 489–498.
3. K. Pearson: *Contributions to the mathematical theory of evolution,* Philos. Trans. R. Soc. A, 185 (1894), 71–110.
4. K.Pearson: X.*Contributions to the mathematical theory of evolution.—II. Skew variation in homogeneous material*, Philos. Trans. R. Soc. A, 186 (1895), 343–414.

5. 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.
6. Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995) Continuous Univariate Distributions, Vol. 2. J. Wiley & Sons, New York, Chap. 28.
7. R. L. Plackett: Karl Pearson and the chi-squared test, Int. Stat. Rev., 51 (1) (1983), 59–72.
8. Fisher, R.A. (1922) On the Interpretation of $\chi 2$ from Contingency Tables and the Calculation of P. *Journal of the Royal Statistical Society A*, 85, 87-94. https://doi.org/10.2307/2340521
9. Lancaster, H.O. (1969) The Chi-Squared Distribution. J. Wiley & Sons, New York, Chap. 1.
10. Bienaymé, I.-J. (1852) Sur la Probabilité des Erreurs d'Aprés la Méthode des Moin dres Carrés. *Liouville's Journal de Mathématiques Poures et Appliquées, Séries* 1, 17, 33-78.
11. Abbe, D.E. (1863) Ueber die Gesetzmässigkeit in der Vertheilung der Fehler bei Beobachtungsreihen. Dissertation, Jena.
12. Evans, R.D. (1985) The Atomic Nucleus. Krieger Publishing, Malabar, Chap. 27.
13. V. Gurvich, M. Naumova: *Logical contradictions in the one-way ANOVA and Tukey-Kramer multiple comparisons tests with more than two groups of observations*, *Symmetry*, 13 (2021), Article ID: 1387.
14. Satchler, G.R. (1990) Introduction to Nuclear Reactions. Oxford U. P., New York, 248. https://doi.org/10.1007/978-1-349-20531-8.
15. Ross, S. (2006) A First Course in Probability. Pearson Prentice Hall, Upper Saddle River, Sec. 6.3.
16. Semkow, T.M., Freeman, N., Syed, U.-F., Haines, D.K., Bari, A., Khan, A.J., Nishikawa, K., Khan, A., Burn, A.G., Li, X. and Chu, L.T. (2019) Chi-Square Distribution: New Derivations and Environmental Application. Journal of Applied Mathematics and Physics, 7, 1786-1799. https://doi.org/10.4236/jamp.2019.78122.
17. Margenau, H. and Murphy, G.M. (1976) The Mathematics of Physics and Chemistry. Krieger Publishing, Huntington, Sec. 8.5.
18. 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 Rea sonably Supposed to Have Arisen from Random Sampling. *Philosophical Magazine Series* 5, 50, 157-175. https://doi.org/10.1080/14786440009463897.
19. Jaynes, E.T. (2004) Probability Theory: The Logic of Science. Cambridge U. P., Cambridge, Chap. 12. https://doi.org/10.1017/CBO9780511790423.
20. Student (1908) The Probable Error of a Mean. *Biometrika*, 6, 1-25. https://doi.org/10.1093/biomet/6.1.1.
21. 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.
22. 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.
23. Berry, D.A. and Lindgren, B.W. (1996) Statistics: Theory and Methods. Wadsworth Publishing, Belmont, Sec. 5.12, 6.4.
24. 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.

25. 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.