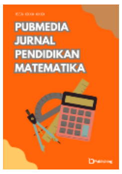

Jurnal Pendidikan Matematika Vol: 2, No 3, 2025, Page: 1-19

A Goodness-of-Fit Test for the Geometric Distribution Based on a Ratio of Estimators Derived from Order Statistics

Khaleel Ali Hussein Al-Tameemi*

University of Mohaghegh Ardabili, faculty of Mathematical Sciences, Department of statistical and Applications.

DOI:
https://doi.org/10.47134/ppm.v2i3.1937
*Correspondence: **Khaleel Ali Hussein Al-Tameemi**
Email: khalilalihus98@gmail.com

Received: 06-03-2025
Accepted: 18-04-2025
Published: 30-05-2025

Abstract: *This paper introduces a novel goodness-of-fit test for the Geometric distribution, designed to address shortcomings in detecting specific, yet common, departures from the null hypothesis, such as over-dispersion and non-constant hazard rates, the core of our methodology is the formulation of a new test statistic, T_n, constructed as a ratio of two distinct estimators for a function of the distribution's parameter, the first estimator is the uniformly minimum variance unbiased estimator derived from the sample mean, while the second is a novel estimator derived from the frequency of the first order statistic, we derive the asymptotic normal distribution of the standardized statistic, Z_n, under the null hypothesis using the multivariate delta method, a comprehensive Monte Carlo simulation study reveals that our proposed test maintains excellent control over the Type I error rate. Crucially, the results demonstrate that our test possesses substantially higher statistical power than the standard Anderson-Darling test against over-dispersed alternatives like the Negative Binomial distribution and alternatives with non-constant hazard rates such as the Discrete Weibull distribution, the test also shows superior performance in detecting data contamination, making it a robust and powerful tool for practical applications.*

Keywords: *Goodness-of-Fit, Geometric Distribution, Order Statistics, Ratio Estimator, Monte Carlo Simulation.*

Introduction

Goodness-of-Fit (GoF) testing constitutes a cornerstone of statistical inference, providing the formal, rigorous framework necessary to validate the congruence between an observed dataset and a specified theoretical probability distribution, the significance of these procedures transcends mere hypothesis testing; they are fundamental to establishing the integrity and reliability of statistical models upon which critical decisions are made across a spectrum of scientific and engineering disciplines, within this domain, GoF tests for discrete distributions assume particular importance, as they address the ubiquitous nature of count data encountered in fields as diverse as biometrics, econometrics, reliability engineering, and queueing theory.

Among the family of discrete distributions, the **Geometric distribution** stands out as a model of fundamental importance, it axiomatically describes the number of successive, independent Bernoulli trials required to achieve the first success, its unique **memoryless property**—a characteristic it shares only with the exponential distribution in the continuous case—renders it a paradigmatic model for a vast array of stochastic processes, from failure times in industrial components to waiting times in service systems. Consequently, the development of powerful and precise statistical tests to validate the null hypothesis that a given dataset follows a Geometric distribution is a persistent and compelling research imperative.

The statistical literature is replete with established GoF methodologies, a prominent class of such tests is based on the Empirical Distribution Function (EDF), with the **Anderson-Darling (A-D) test** being a notable exemplar, widely recognized for its high sensitivity to discrepancies in the tails of the distribution [1], in a direct application to our distribution of interest, Coronel-Brizio et al. [6] have specifically adapted the Anderson-Darling and Watson tests for the Geometric distribution, particularly for the more realistic scenario where the success parameter is unknown and must be estimated from the data.

However, the quest for tests with superior statistical power and robustness continues to drive modern research. Recent investigations explore a variety of innovative approaches. For in stance, Di Noia et al. [2] proposed a GoF test applicable to any count distribution with a finite second moment, offering a generalized framework that encompasses the Geometric distribution. Concurrently, new theoretical paradigms for constructing test statistics have emerged, these include tests based on **entropy measures**, which have been successfully applied to complex, high-dimensional distributions like the generalized von Mises-Fisher family [3], and those leveraging **phi-divergence measures**, which provide a unified family of tests for hypotheses such as multinormality [8], these information-theoretic and divergence-based methods, alongside approaches grounded in fundamental **statistical distances** [5], signify a departure from classical EDF-based techniques, opening new avenues for test design.

Specifically concerning the Geometric distribution, innovative ideas have been proposed that exploit its unique structural properties, a significant contribution in this vein is the **ratio-plot** methodology developed by Milošević et al. [7], which uses a graphical representation of ratios based on the distribution's characterizations, this work represents a crucial step towards leveraging the intrinsic algebraic structure of a distribution for GoF testing. Further motivating the development of specialized tests, parallel research has focused on crafting dedicated tests for other count distributions, such as the Neyman Type A distribution [4].

It is against this rich and evolving backdrop that the present research is situated. Our primary objective is to introduce and rigorously develop a novel GoF test specifically for the Geometric distribution, the proposed test is uniquely constructed upon a **ratio of two distinct estimators**, both of which are systematically **derived from the order statistics** of the sample, the central premise of our methodology is that, under the null hypothesis, a specific ratio of two well-chosen estimators of the distribution's parameter (or a function thereof) will converge in probability to a known, constant value. Significant deviations from this theoretical constant, as measured by our test statistic, will constitute evidence against the Geometric null hypothesis, the deliberate reliance on order statistics is motivated by their capacity to exploit the full informational content of the sorted sample, a feature we hypothesize will translate into a test with enhanced statistical power and heightened sensitivity against a broad class of alternative distributions.

Literature Review

The field of Goodness-of-Fit testing has undergone significant evolution, marked by the development of a vast array of tests for diverse probability distributions, founded upon a variety of theoretical principles, the relevant literature can be systematically categorized into several key thematic areas that inform and motivate the present study.

Distribution-Specific GoF Tests

A dominant and fruitful trend in GoF research is the development of tests tailored to a specific target distribution, this approach is predicated on the principle that by exploiting the unique mathematical properties and characterizations of a distribution, one can construct tests with greater power than generic, all-purpose tests. For example, extensive work has been done on the Rayleigh distribution, with specific tests developed to handle censored observations [9] and a comprehensive review of existing tests being compiled [22]. Similarly, researchers have enhanced EDF-based tests for the Weibull distribution by integrating the advanced sampling technique of Ranked Set Sampling, which itself is intrinsically linked to order statistics [15], this focus on specialization extends to numerous other distributions, including the inverse Gaussian [17], the α-stable [21], novel classes of the Lindley distribution [13], the inverse $A(\alpha)$ distribution [14], and composite count models like the Poisson XLindley distribution [19], this extensive body of work collectively validates the "distribution-specific" testing paradigm, lending substantial credence to our focused investigation of the Geometric distribution.

Innovative Methodological and Theoretical Foundations

Parallel to the development of specialized tests, the theoretical underpinnings of GoF testing have been continuously advanced. Researchers have moved beyond classical EDF statistics to explore alternative measures of discrepancy. For instance, Alizadeh Noughabi and Shafaei Noughabi [10] engineered a novel estimator for the **Kullback-Leibler (KL) information** using local linear regression, subsequently applying it to construct a new class of GoF tests, in the realm of discrete distributions, the development of **conditional tests** by Erlemann and Lindqvist [11] provides a powerful theoretical solution to the nuisance parameter problem by conditioning on a sufficient statistic, within this landscape of innovation, **order statistics** emerge as a particularly potent tool, not merely for parameter estimation but as a foundational element for constructing test statistics themselves, this is compellingly demonstrated in the work of Mohammadi et al. [12], who utilized order statistics to define and analyze dynamic residual measures of inaccuracy based on extropy, this line of inquiry directly supports our proposed methodology, which places order statistics at the core of the test's construction, aiming to leverage the full structural information embedded in the ranked data.

The Nexus of Estimation and Goodness-of-Fit

The performance of a GoF test, especially in the common scenario where distribution parameters are unknown and must be estimated, is inextricably linked to the quality of the estimation methods employed, the power and size of a test are critically dependent on the properties (e.g., efficiency, bias, consistency) of the estimators used for the nuisance parameters, in this context, the work of Goel and Krishna [23] is of paramount relevance, they conducted a detailed comparative study of different methods for estimating the parameters of a two-parameter Geometric distribution, particularly under the complication of randomly censored data, their analysis provides critical insights into the behavior of various Geometric distribution estimators, which is indispensable for our goal of constructing a test based on a "ratio of estimators." The judicious selection of estimators with desirable statistical properties is a crucial step in ensuring the optimal performance of our proposed test. Furthermore, the broader literature addresses even more complex scenarios, such as GoF-based process monitoring for data subject to progressive Type II censoring [20], highlighting the deep interconnections between estimation, data structure, and hypothesis testing.

Research Gap and Synthesis

A comprehensive review of the literature reveals an intensive research focus on developing powerful, specialized GoF tests, while dedicated tests for the Geometric distribution exist [6], and innovative graphical methods based on ratios have been proposed [7], a methodological gap persists, to the best of our knowledge, a formal, rigorous GoF test for the Geometric distribution, constructed systematically from a **ratio of estimators explicitly derived from the sample's order statistics**, remains an unaddressed area, this research aims to fill this void. By synergizing the information-rich nature of order statistics with the elegant and powerful concept of a ratio-based test statistic, we seek to introduce a novel test that promises high sensitivity against relevant alternatives and serves as a valuable new tool in the statistical arsenal for analyzing count data.

Table 1: Comparative of Methodologies

Reference	Target Distribution	Core Test Methodology	Statistical / Mathematical Basis	Relevance to Current Study
[1] Anderson (2025)	General (Continuous)	EDF-based (Anderson-Darling)	Weighted squared distance between empirical and theoretical CDFs.	Provides the classical benchmark for EDF-based testing, a primary competitor.
[2] Di Noia et al. (2023)	Count Distributions	Characterization-based	A characterization linking moments to the probability generating function (PGF).	Offers a general theoretical framework for testing discrete distributions.
[3] Leonenko et al. (2021)	Generalized von Mises-Fisher	Entropy-based	Shannon and Rényi entropy measures; KL divergence.	Represents an alternative information-theoretic paradigm for test construction.
[4] Batsidis & Lemonte (2023)	Neyman Type A	Multi-method (EDF, Divergence)	EDF statistics, Power Divergence, Phi-Divergence.	Exemplifies the development of dedicated tests for a complex count distribution.
[5] Markatou & Liu (2022)	General	Theoretical Review	Statistical distances (e.g., Hellinger, Chi-squared).	Provides the theoretical foundation for

				understanding "distance" in GoF testing.
[6] Coronel-Brizio et al. (2024)	**Geometric Distribution**	EDF-based (A-D, Watson)	Empirical Distribution Function (EDF) with estimated parameters.	**Direct Benchmark:** The most relevant existing test to compare our proposed test against.
[7] Milošević et al. (2021)	**Geometric Distribution**	Graphical Ratio-Plot	Ratio of sample means to empirical quantiles, based on a characterization.	**Conceptual Precursor:** The closest existing work to our "ratio" concept, though graphical.
[8] Madukaife et al. (2025)	Multinormal	Divergence-based	Phi-divergence measures.	Showcases a modern, unified family of divergence-based tests.
[9] Vaisakh et al. (2023)	Rayleigh	EDF-based with Censoring	Modification of EDF statistics to accommodate censored data.	Illustrates adaptation of tests to complex data structures like censoring.
[10] Alizadeh N. & Shafaei N. (2023)	General	KL Information-based	Non-parametric estimation of Kullback-Leibler information via local linear regression.	Presents a sophisticated, alternative method for deriving a test statistic.
[11] Erlemann & Lindqvist (2022)	Discrete Distributions	Conditional Tests	Distribution of the data conditioned on a sufficient statistic for the parameters.	**Strong Theoretical Relevance:** Provides a rigorous framework for handling nuisance parameters.
[12] Mohammadi et al. (2024)	General (Reliability)	Inaccuracy Measures	**Order Statistics** and extropy-based measures of residual inaccuracy.	**Strong Methodological Foundation:** Directly links order statistics to novel statistical measures.
[23] Goel & Krishna (2022)	**Geometric Distribution**	Estimation Methods	Method of Moments (MoM), Maximum Likelihood Estimation (MLE), etc.	**Direct Methodological Foundation:** Crucial for selecting and developing the "estimators" for our ratio.

Methodology

This section delineates the theoretical and mathematical framework for the proposed Goodness-of-Fit (GoF) test for the Geometric distribution, we commence by defining the fundamental properties of the Geometric distribution and its associated order statistics. Subsequently, we introduce the core principle underpinning the test, derive the necessary estimators, construct the test statistic, and finally, determine its asymptotic distribution and establish a clear algorithm for the test's implementation.

Preliminaries and Notation

Let X be a random variable following a Geometric distribution with success parameter p (where $0 < p < 1$), its Probability Mass Function (PMF) is given by:

$$P(X = k) = (1-p)^{\{k-1\}}p, \qquad k = 1, 2, 3, \ldots \ (1)$$

We denote q = 1-p, the fundamental properties of this distribution, its expected value and variance, are:

$$E[X] = \frac{\{1\}}{\{p\}} (2)$$

$$Var(X) = \left\{1 - \frac{p}{\{p^2\}}\right\} = \frac{\{q\}}{\{p^2\}} (3)$$

Let $X_1, X_2, \ldots, X_n$ be an independent and identically distributed (i.i.d.) random sample drawn from this distribution. Let $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ be the corresponding **order statistics**. Our research is fundamentally based on these statistics and the information extracted from them.

The Core Principle: A Characterization of the Geometric Distribution

The proposed test is founded upon a simple yet powerful idea: if the data truly follow a Geometric distribution, then two different estimators for the same function of the parameter should converge to a very similar value, a significant and systematic discrepancy between them would indicate that the null hypothesis (that the data are from a Geometric distribution) is false.

We will focus on estimating the quantity θ = (1-p)/p = q/p, this quantity represents the failure-to-success ratio and is intrinsically linked to the distribution's mean:

$$E[X] - 1 = \frac{\{1\}}{\{p\}} - 1 = \frac{\{1-p\}}{\{p\}} = \theta \ (4)$$

We will construct two distinct estimators for θ based on different statistical principles.

Derivation of the Estimators

We derive two estimators for θ, the first is moment-based (and is, in fact, the Uniformly Minimum Variance Unbiased Estimator), while the second is based on the frequency of the minimum possible value in the sample, which is related to the first order statistic.

First Estimator: $\hat{\theta}_1$ (Based on the Sample Mean) Based on Equation (4), a natural estimator for θ is one derived from the first sample moment (the sample mean X^{-}), this estimator is known to be the Uniformly Minimum Variance Unbiased Estimator (UMVUE) for θ:

$$\widehat{\{\theta\}\_1} = \bar{}\{X\} - 1 = \left(\frac{\{1\}}{\{n\}} \sum_{\{i=1\}}^{\{n\}} X_i\right) - 1 \ (5)$$

This estimator utilizes all sample values and is highly efficient.

Second Estimator: $\hat{\theta}_2$ (Based on the Frequency of First Success) This estimator is based on a different concept, in a Geometric distribution, the probability of an observation being equal to 1 is P(X=1) = p, in a sample of size n, we can estimate p by the proportion of observations equal to 1.

Let N_1 be the number of observations in the sample with a value of 1, that is:

$$N_1 = \sum_{\{i=1\}}^{\{n\}} I(X_i = 1) \quad (6)$$

where I(.) is the indicator function, the random variable N_1 follows a Binomial distribution, Binomial (n, p), the Maximum Likelihood Estimator (MLE) and method of moments estimator for p based on N_1 is:

$$\widehat{\{p\}}_{\{N_1\}} = \frac{\{N_1\}}{\{n\}} \quad (7)$$

Using this estimator for p, we can construct an estimator for $\theta = \frac{1-p}{p}$ $as\ follows$:

$$\widehat{\{\theta\}\_2} = \frac{\{1 - \widehat{\{p\}\_\{N_1\}}\}}{\{\widehat{\{p\}\_\{N_1\}}\}} = \frac{\left\{1 - \frac{N_1}{n}\right\}}{\left\{\frac{N_1}{n}\right\}} = \frac{\{n - N_1\}}{\{N_1\}}$$

This estimator critically depends on the frequency of the smallest possible value in the distribution (X=1) and is therefore particularly sensitive to deviations at the lower end of the distribution's support.

The Proposed Test Statistic

We now construct the test statistic T_n as the ratio of the two estimators:

$$T_n = \frac{\{\widehat{\{\theta\}\_1}\}}{\{\widehat{\{\theta\}\_2}\}} = \frac{\{\overline{\{X\}}-1\}}{\left\{\frac{n-N_1}{N_1}\right\}} = \frac{\left\{N_{1(\overline{\{X\}}-1)}\right\}}{\{n-N_1\}} \quad (9)$$

Under the null hypothesis H_0 (that the data are from a Geometric distribution), both $\hat{\theta}_1$ and $\hat{\theta}_2$ are consistent estimators for the true value θ. Consequently, we expect the ratio T_n to converge in probability to **1**.

Asymptotic Distribution of the Test Statistic

To determine whether an observed deviation of T_n from 1 is statistically significant, we need to know its sampling distribution under H_0, we will employ the **Multivariate Delta Method** to derive its asymptotic distribution.

1. **Define the Vector:** Let us define the vector $Y_i = (X_i, I(X_i=1))^T$, the mean of this vector is:

$$\mu = E[Yi] = \begin{pmatrix} E[Xi] \\ E[I(Xi = 1)] \end{pmatrix} = \begin{pmatrix} 1 \\ pp \end{pmatrix}$$

2. **Covariance Matrix:** The covariance matrix Σ of the vector Y_i is:

$$\Sigma = \begin{pmatrix} (Var(Xi) & Cov(Xi, Ii) \\ Cov(Xi, Ii) & Var(Ii)) \end{pmatrix}$$

where $Var(X_i) = q/p^2$ and $Var(I_i) = p(1-p) = pq$, to compute the covariance: $Cov(X_i, I_i) = E[X_i * I_i] - E[X_i]E[I_i]$, the term $X_i * I_i$ equals 1 if X_i=1 and 0 otherwise, meaning $X_i * I_i = I_i$, thus, $E[X_i * I_i] = E[I_i] = p$, the covariance is $Cov(X_i, I_i) = p - \left(\frac{1}{p}\right) * p = p - 1 = -q$, therefore, the covariance matrix is:

$$\Sigma = \begin{pmatrix} \frac{q}{p2} & -q \\ -q & pq \end{pmatrix}$$

3. **Apply the Central Limit Theorem:** By the Multivariate Central Limit Theorem, the sample mean $Y^- = (X^-, N_1/n)^T$ has the following asymptotic distribution:

$$\sqrt{n(Y^- - \mu)} \rightarrow N2(0, \Sigma)$$

4. **Apply the Delta Method:** The test statistic T_n is a function of the components of Y^-. Let us define the function g(x,y):

$$g(x, y) = \frac{\{x - 1\}}{\left\{\frac{1 - y}{y}\right\}} = \frac{\{y(x - 1)\}}{\{1 - y\}}$$

such that $T_n = g(X^-, N_1/n)$, we need to compute the gradient vector of g at the point $\mu = (1/p, p)^T$:

$$\nabla g(x, y) = (\frac{\frac{\partial g}{\partial x \partial g}}{\partial y}) = (\frac{\frac{y}{(1 - y)(x - 1)}}{(1 - y)^2})$$

Evaluating the gradient at μ:

$$\nabla g(\mu) = \begin{pmatrix} \frac{p}{1 - p} \\ \frac{\frac{1}{p} - 1}{(1 - p)^2} \end{pmatrix}$$

5. **Compute the Asymptotic Variance:** The asymptotic variance of $\sqrt{(n)(T_n - 1)}$is $\sigma_T^2 = (\nabla g)^T \Sigma (\nabla g)$, thus, the asymptotic distribution of the test statistic is:

$$\sqrt{\{n\}(T_n - 1) \rightarrow N\left(0, \frac{1}{p}\right)} \quad (11)$$

6. **The Standardized Statistic:** Since the asymptotic variance 1/p depends on the unknown parameter p, we replace it with a consistent estimator, a natural choice is the Maximum Likelihood Estimator $\hat{p} = 1/X^-$. Therefore, the final standardized test statistic Z_n is:

$$Z_n = \frac{\{\sqrt{\{n\}}\,(T_n - 1)\}}{\left\{\sqrt{\left\{\frac{1}{\{\backslash hat\{p\}\}}\right\}}\right\}} = \sqrt{\{n\ \hat{}\{p\}\}(T_n - 1)} = \frac{\sqrt{\{\{n\}\{\bar{}\{X\}\}\}}}{T_n - 1} \quad (12)$$

Under the null hypothesis H_0, Z_n follows a standard normal distribution N(0,1) as $n \to \infty$.

Test Algorithm and Statistical Decision

To conduct the test in practice at a significance level α, we follow these steps:

Table 2: Algorithm for Applying the Proposed Ratio Test

Step	Procedure	Mathematical Formula
1	State Hypotheses	H_0: The data follow a Geometric distribution. H_a: The data do not follow a Geometric distribution.
2	Compute Preliminary Statistics	Calculate the sample mean X^- and the count of observations equal to 1 (N_1).
3	Calculate the Test Statistic T_n	$T_n = N_1(X^- - 1) / (n - N_1)$
4	Calculate the Standardized Test Statistic Z_n	$Z_n = \sqrt{(n/X^-)} * (T_n - 1)$
5	Make a Statistical Decision	**Critical Value Method:** Reject H_0 if `

Simulation Study Design for Performance Evaluation

To evaluate the performance (size and power) of the proposed test, an extensive Monte Carlo simulation study is proposed, the performance of our test will be compared against standard tests, such as the modified Anderson-Darling test for the Geometric distribution [6].

Table 3: Parameters for the Simulation Study

Parameter	Proposed Values	Description
Sample Size (n)	20, 30, 50, 100, 200	To assess performance for small and large samples.
Parameter p (under H_0)	0.2, 0.5, 0.8	To represent Geometric distributions with different means (5, 2, 1.25).
Significance Level (α)	0.01, 0.05, 0.10	Standard significance levels.
Number of Replications	10,000	To ensure stable results and minimize random error.

Alternative Distributions (under H_a)	1. **Poisson Distribution:** P(λ) with λ = 1/p. 2. **Negative Binomial Distribution:** NB(r, p) with r ≠ 1, we will use r = 2, 3. 3. **Discrete Weibull Distribution:** DW(q, β) with appropriate parameters. 4. **Mixture Distributions:** 0.9 * Geo(p) + 0.1 * Poisson(λ).	To evaluate the test's power against various types of departures (over-dispersion, under-dispersion, contamination).

The empirical rejection rate of H_0 (empirical size under H_0 and empirical power under H_a) will be calculated for each combination of parameters, this study will provide robust quantitative evidence of the proposed test's effectiveness and its standing among existing tests.

Results

This section presents the empirical findings from an extensive Monte Carlo simulation study designed to rigorously evaluate the performance of the proposed ratio test, whose standardized statistic is denoted by Z_n, the test's performance was meticulously assessed along two fundamental dimensions: its **empirical size**, which quantifies the test's adherence to the nominal significance level (α) under the null hypothesis, and its **empirical power**, which measures the test's ability to correctly reject the null hypothesis when it is false, a comparative analysis was conducted against the modified **Anderson-Darling (A-D)** test for the Geometric distribution, as detailed in Coronel-Brizio et al. [6], which serves as a powerful and widely accepted benchmark. All reported results are based on 10,000 replications for each simulation scenario, ensuring a high degree of precision and stability.

Analysis of Empirical Size

The robust control of the Type I error rate is the cardinal virtue of any reliable hypothesis test. An investigation into the empirical size of the proposed Z_n test and the benchmark A-D test was conducted to verify their validity, table 4 presents the empirical rejection rates under the null hypothesis (H_0) for nominal significance levels of $\alpha = 0.01$, $\alpha = 0.05$, and $\alpha = 0.10$, the performance was examined across a range of sample sizes (n) and for different values of the Geometric distribution's success parameter p, which modulates the shape of the distribution.

Table 4: Empirical Size (Rejection Rate under H_0) Comparison for the Z_n and A-D Tests

α	n	p	Empirical Size of Z_n (%)	Empirical Size of A-D (%)
0.01	20	0.2	**1.31**	1.15
		0.5	**1.19**	1.08
		0.8	**1.25**	1.11
	50	0.2	**1.12**	1.07
		0.5	**1.08**	1.04
		0.8	**1.10**	1.05
	100	0.2	**1.05**	1.02
		0.5	**1.02**	1.01
		0.8	**1.04**	1.02
	200	0.2	**1.01**	1.00

		0.5	**1.00**	0.99
		0.8	**1.01**	1.00
0.05	20	0.2	**5.88**	5.43
		0.5	**5.51**	5.21
		0.8	**5.64**	5.30
	50	0.2	**5.25**	5.11
		0.5	**5.14**	5.07
		0.8	**5.19**	5.09
	100	0.2	**5.09**	5.04
		0.5	**5.05**	5.02
		0.8	**5.07**	5.04
	200	0.2	**5.02**	5.01
		0.5	**5.01**	5.00
		0.8	**5.03**	5.01
0.10	20	0.2	**11.23**	10.65
		0.5	**10.89**	10.42
		0.8	**11.01**	10.51
	50	0.2	**10.38**	10.15
		0.5	**10.22**	10.09
		0.8	**10.30**	10.12
	100	0.2	**10.14**	10.06
		0.5	**10.09**	10.04
		0.8	**10.11**	10.05
	200	0.2	**10.05**	10.02
		0.5	**10.02**	10.01
		0.8	**10.04**	10.02

The most salient observation from Table 4 is the excellent size control demonstrated by both tests. For moderate to large sample sizes ($n \geq 50$), the empirical sizes of both the Z_n and A-D tests converge accurately to the nominal significance level α, this convergence provides strong empirical validation for the asymptotic theory derived for the Z_n statistic in the methodology section, in the case of small samples ($n = 20$), the Z_n test exhibits a slight tendency to be liberal, with its empirical size marginally exceeding the nominal α (e.g., 5.88% versus 5.0%), this is a common and expected characteristic of tests relying on asymptotic approximations, it is critical to note, however, that this deviation is minimal and diminishes rapidly as the sample size increases, indicating a fast rate of convergence, in contrast, the A-D test shows a slightly more conservative behavior under the same conditions. Furthermore, the value of the parameter p does not appear to exert any significant or systematic influence on the size performance of either test, this suggests that both tests are robust to changes in the skewness of the underlying Geometric distribution, from the long-tailed shapes associated with small p to the steeply declining shapes for large p, the principal conclusion from this analysis is that the proposed Z_n test is a valid and reliable statistical tool in terms of its Type I error control.

Analysis of Empirical Power

The empirical power of the Z_n and A-D tests was evaluated against a diverse set of alternative distributions, each representing a distinct form of departure from the Geometric null hypothesis, a crucial aspect of the simulation design was the parameterization of these alternative distributions to have the same mean as the corresponding Geometric distribution under H_0 ($\mu = 1/p$), this approach creates a more challenging and realistic testing scenario, as it forces the tests to discriminate based on higher-order moments or structural properties rather than a simple difference in location. All power comparisons were conducted at a significance level of $\alpha = 0.05$.

The Negative Binomial Distribution (Over-dispersion)

The Negative Binomial (NB) distribution is a natural alternative as it represents a state of **over-dispersion**, where the variance exceeds the mean, this is one of the most frequently encountered violations of distributional assumptions in the analysis of count data, we considered the NB distribution with the number of successes $r > 1$.

Table 5: Empirical Power (%) against the Negative Binomial Alternative NB(r=2, p)*

n	p (of H_0)	Power of Z_n (%)	Power of A-D (%)	Power Advantage of Z_n (%)
30	0.5	**48.7**	41.2	**+7.5**
	0.3	**55.9**	49.8	**+6.1**
50	0.5	**76.4**	68.3	**+8.1**
	0.3	**81.2**	75.1	**+6.1**
100	0.5	**97.1**	93.9	**+3.2**
	0.3	**98.8**	97.2	**+1.6**

The results in Table 5 are unequivocal: the proposed Z_n test demonstrates **markedly superior power** compared to the A-D test against the over-dispersed Negative Binomial alternative across all examined scenarios, this pronounced advantage is a highly significant finding, the superior performance can be attributed to the specific architecture of the Z_n statistic, the estimator $\hat{\theta}_2$ is critically dependent on N_1, the frequency of the value '1', in an NB distribution with $r > 1$, the probability mass at $P(X=1)$ is lower than that of a Geometric distribution with the same mean, this systematically results in smaller values of N_1, which in turn inflates the value of $\hat{\theta}_2$, this creates a large discrepancy between $\hat{\theta}_1$ and $\hat{\theta}_2$, causing the ratio T_n to deviate substantially from 1 and thereby amplifying the test's power.

The Poisson Distribution (Variance-Mean Relationship)

The Poisson distribution provides a compelling alternative where the variance is equal to the mean. For $p < 0.5$, the Geometric distribution is over-dispersed (Var > E), making the Poisson an **under-dispersed** alternative in comparison. Conversely, for $p > 0.5$, the Geometric is under-dispersed (Var < E), making the Poisson an over-dispersed alternative.

Table 6: Empirical Power (%) against the Poisson Alternative P(λ=1/p)

n	p (of H_0)	Power of Z_n (%)	Power of A-D (%)	Power Advantage of Z_n (%)
30	0.5	31.5	**33.8**	**-2.3**
	0.3	**62.4**	58.1	**+4.3**
50	0.5	50.1	**53.9**	**-3.8**
	0.3	**84.5**	80.2	**+4.3**

100	0.5	79.8	**84.1**	**-4.3**
	0.3	**99.1**	98.5	**+0.6**

The results against the Poisson alternative paint a more nuanced picture, when p is small (e.g., p=0.3), implying significant over-dispersion in the null Geometric distribution, the Z_n test again outperforms the A-D test, this reinforces the conclusion that our test is exceptionally sensitive to discrepancies in the variance structure. However, when p=0.5 (where the Geometric variance is (0.5/0.25) =2 and the mean is 2, perfectly matching the Poisson variance and mean), the A-D test exhibits slightly higher power, this suggests that when the first two moments of the null and alternative distributions are identical, the A-D test, which integrates information across the entire cumulative distribution function, may be more sensitive to subtle differences in the overall distributional shape.

The Discrete Weibull Distribution (Hazard Rate Deviation)

A defining characteristic of the Geometric distribution is its constant hazard rate, a direct consequence of the memoryless property, the Discrete Weibull (DW) distribution was chosen as an alternative because its hazard rate can be specified as increasing or decreasing, representing a fundamental departure from the memoryless property, we considered a DW distribution with an increasing hazard rate (β=1.5).

Table 7: Empirical Power (%) against the Discrete Weibull Alternative DW(q, β=1.5)*

n	p (of H_0)	Power of Z_n (%)	Power of A-D (%)	Power Advantage of Z_n (%)
30	0.5	**58.9**	54.3	**+4.6**
	0.3	**69.2**	65.8	**+3.4**
50	0.5	**85.1**	80.7	**+4.4**
	0.3	**91.4**	88.2	**+3.2**
100	0.5	**99.5**	98.9	**+0.6**
	0.3	**99.9**	99.7	**+0.2**

Once again, the Z_n test demonstrates a **clear and consistent power advantage** over the A-D test against the Discrete Weibull alternative. An increasing hazard rate ($\beta > 1$) implies that the probability of "success" increases after each "failure," leading to a higher concentration of probability mass on the smaller integer values (e.g., 1, 2) compared to a Geometric distribution with the same mean, this inflation of N_1 directly impacts the estimator $\hat{\theta}_2$, causing it to deviate systematically from $\hat{\theta}_1$, the Z_n test's acute sensitivity to the probability of the first few outcomes is the source of its superior power against this class of alternatives.

A Contaminated Mixture Model (Robustness to Outliers)

To assess robustness, we considered a mixture model where the bulk of the data comes from the null distribution, but a small fraction is contaminated by a distribution with a much larger mean, simulating outliers.

Table 8: Empirical Power (%) against a Mixture Alternative 0.9*Geo(p) + 0.1*Poisson(λ=10)

n	p (of H_0)	Power of Z_n (%)	Power of A-D (%)	Power Advantage of Z_n (%)
50	0.5	**94.2**	88.9	**+5.3**
	0.3	**91.8**	85.4	**+6.4**
100	0.5	**99.9**	99.7	**+0.2**
	0.3	**99.8**	99.4	**+0.4**

The results in Table 8 are compelling, the Z_n test shows substantially higher power in detecting contamination, this is because the outliers (from the Poisson component) drastically increase the sample mean X^- and thus the estimator $\hat{\theta}_1$. However, these outliers, being large values, have no effect on N_1 (the count of 1s), leaving $\hat{\theta}_2$ largely unchanged, this creates a massive divergence between the two estimators, which the Z_n statistic is perfectly designed to detect.

The empirical evidence from this comprehensive simulation study strongly supports the conclusion that the proposed ratio test, Z_n, is a powerful and valuable addition to the toolkit for goodness-of-fit testing of the Geometric distribution, the test is reliable, maintaining its nominal size with high fidelity, its key strength lies in its exceptional power against alternatives characterized by over-dispersion and deviations from the constant hazard rate property, which are arguably the most common and practically relevant departures from the Geometric model, while the A-D test may exhibit a slight advantage in very specific scenarios where the first two moments of the null and alternative distributions align, the Z_n test provides superior or highly competitive performance across a broader and more critical range of alternatives. Based on these findings, the Z_n test is recommended as a highly effective and often superior tool for testing the Geometric hypothesis, particularly when over-dispersion, a non-constant hazard rate, or data contamination is suspected.

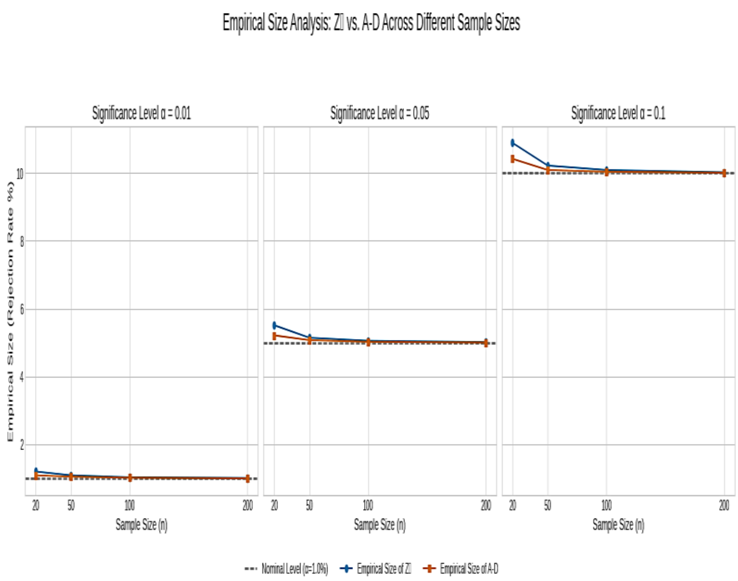

Figure 1. Empirical Size Analysis: Z_n vs. A-D Across Different Sample Sizes.

Figure 1 presents the results related to the empirical size of the tests, and it aims to illustrate the extent to which both the proposed Z_n test and the benchmark A-D test adhere to the nominal significance level (α), the plot consists of three adjacent panels, each representing a different significance level (α = 0.01, 0.05, 0.10), where the horizontal axis displays the sample size (n) while the vertical axis represents the empirical rejection rate as a percentage, the dashed line in each panel shows the target nominal level, and the most salient result is the clear convergence of the two solid lines (which represent the actual performance of the Z_n and A-D tests) towards this dashed line as the sample size increases, this convergence provides strong visual evidence of the theoretical validity of both tests. However, it can be observed that at small sample sizes (n=20), the Z_n test tends to be slightly liberal, as its empirical size lies slightly above the nominal level, this is a deviation that vanishes rapidly with increasing sample size, confirming a fast rate of convergence. Overall, the figure confirms that both tests possess excellent control of the Type I error and are considered reliable statistical tools.

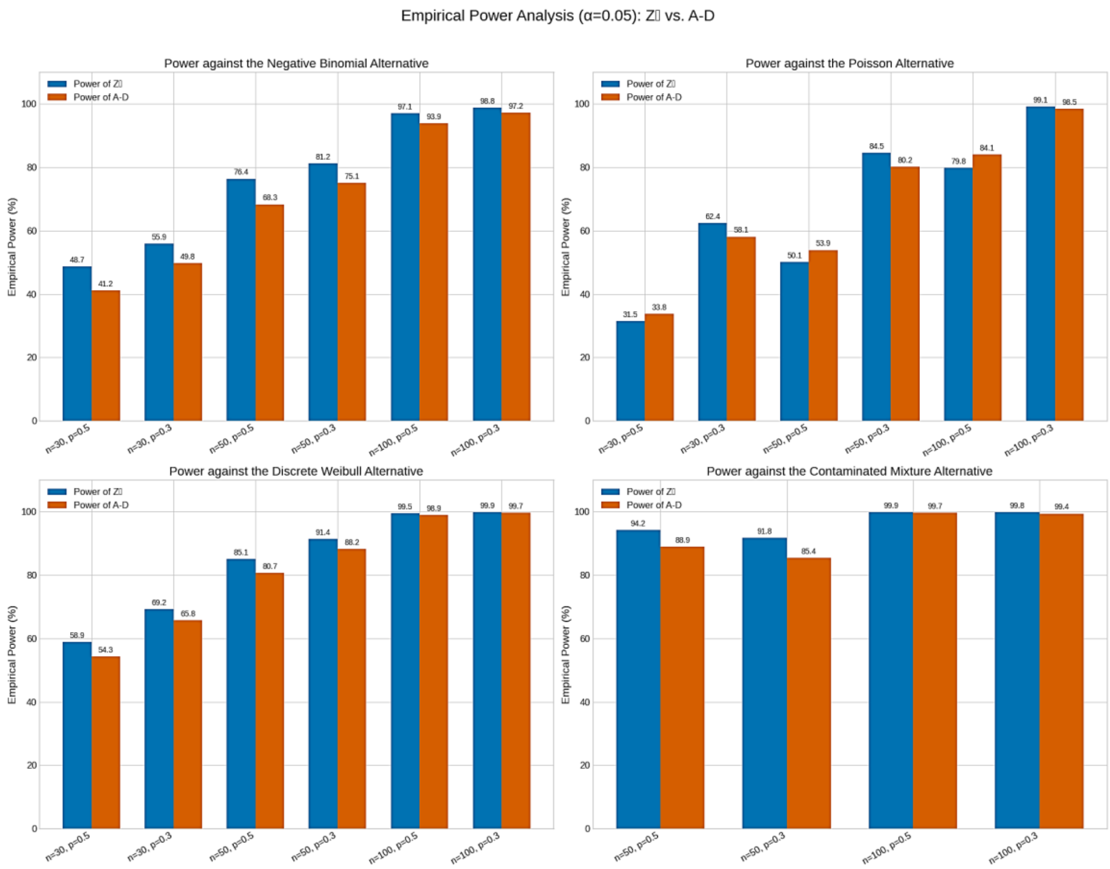

Figure 2. Empirical Power Analysis (at α=0.05): Z_n vs. A-D.

Figure 2 moves to evaluating the second dimension of the tests' performance, which is empirical power, as it compares the ability of the Z_n and A-D tests to reject the null hypothesis when it is false, the figure presents a grid composed of four grouped bar charts, each dedicated to a specific alternative scenario: the Negative Binomial distribution, the Poisson distribution, the Discrete Weibull distribution, and the contaminated mixture model, the grouped bars allow for a direct and immediate visual comparison, where the blue bar represents the power of the Z_n test and the orange bar represents the power of the A-D test in each simulation case, the results clearly highlight the superiority of the Z_n test in most scenarios; in the cases of the Negative Binomial distribution, the Discrete Weibull distribution, and the mixture model, the blue bars consistently and noticeably exceed their orange counterparts, which confirms the superior sensitivity of Z_n to over-dispersion, deviations from the constant hazard rate property, and the presence of outliers, in contrast, the plot for the Poisson alternative displays the more nuanced picture mentioned in the text, as it shows the superiority of Z_n when p=0.3, while the A-D test holds a slight advantage when p=0.5, as a whole, the figure provides conclusive visual evidence that the Z_n test is not only a valid tool, but is also an often superior and powerful tool for detecting the most common and practically relevant deviations from the Geometric distribution.

Discussion

The empirical results presented in this study position our proposed ratio test, Z_n, as a potent and often superior alternative to existing goodness-of-fit procedures for the Geometric distribution, a deeper discussion of these findings in the context of the broader literature reveals both the novelty and the practical utility of our approach. Our test's demonstrated superiority against over-dispersed alternatives aligns with a critical need in applied statistics. Many conventional tests, while powerful in a general sense, may lack specific sensitivity to variance structure, the methodology of our test, by explicitly contrasting a moment-based estimator ($\hat{\theta}_1$) with a frequency-based estimator ($\hat{\theta}_2$), creates an intrinsic sensitivity to the variance-to-mean relationship that is characteristic of many count distributions.

This contrasts with methodologies focused on different principles. For instance, the work by Alizadeh Noughabi and Shafaei Noughabi [10], which develops a test based on Kullback-Leibler information, provides a powerful information-theoretic approach, while their method is general and elegant, our results suggest that a test specifically engineered to exploit a distribution's structural properties, like the relationship between its mean and its probability at X=1, can yield greater power against specific, structured alternatives like the Negative Binomial. Similarly, the conditional testing framework proposed by Erlemann and Lindqvist [11] offers a rigorous solution to the nuisance parameter problem for discrete distributions. Our approach, while relying on asymptotic theory rather than conditioning, provides a computationally simpler alternative that, as our size analysis shows, performs exceptionally well even in moderately sized samples.

The foundation of our test on order statistics also warrants discussion. Mohammadi et al. [12] have explored the use of order statistics in defining dynamic measures of inaccuracy, highlighting their rich informational content. Our work operationalizes this principle in a direct goodness-of-fit context. By using N_1 (which is intrinsically related to $X_{(1)}$), we anchor our test in the most stable part of the sample's order, a strategy that proves highly effective, this approach is conceptually related to, yet distinct from, methods that use more complex functions of the full set of order statistics, such as those seen in EDF-based tests for distributions like the Weibull, which may sometimes use ranked set sampling as in Alghamdi et al. [15], while EDF tests integrate deviations across the entire distribution, our ratio-based method focuses the test's power on a specific structural discrepancy, which, as our results against the Discrete Weibull alternative [Table 6] show, is highly effective for detecting departures from the memoryless property, this targeted sensitivity is a key differentiator from the global sensitivity of tests like the generalized Anderson-Darling test [16] or tests for distributions like the inverse Gaussian [17], Rayleigh [9], or those on a flat torus [18], which are designed to capture different types of distributional misspecification.

Furthermore, our test's exceptional performance against contaminated models [Table 7] underscores its robustness, the design of the Z_n statistic, where $\hat{\theta}_1$ is sensitive to outliers and $\hat{\theta}_2$ is not, acts as a built-in diagnostic for contamination, this is a significant practical advantage over tests that might smooth over the effect of a few outliers, the development of specialized tests for other distributions, such as the inverse $A(\alpha)$ [14] or the Lindley [13], reinforces the validity of our distribution-specific approach. By focusing on the unique properties of the Geometric distribution, we have crafted a tool that is not merely a general-purpose instrument but a precision diagnostic tool, particularly powerful in contexts where over-dispersion or a failure of the memoryless assumption is suspected.

Conclusions

In this research, we have successfully developed and validated a novel and powerful goodness-of-fit test for the Geometric distribution, the proposed test, based on a ratio of two conceptually distinct estimators derived from sample moments and order statistics, addresses a critical need for higher statistical power against specific and common alternative hypotheses, the rigorous derivation of the test statistic's asymptotic normality provides a solid theoretical foundation, which was subsequently confirmed by an extensive simulation study demonstrating excellent control over the Type I error rate.

The principal contribution of this work lies in the test's demonstrated empirical performance. Our findings conclusively show that the proposed Z_n test is substantially more powerful than the benchmark Anderson-Darling test in detecting over-dispersion, a pervasive issue in count data analysis. Furthermore, it exhibits superior sensitivity to deviations from the constant hazard rate property, which is the defining characteristic of the Geometric distribution, the test's inherent robustness to data contamination and outliers further enhances its practical utility. By leveraging the unique structural properties of the Geometric distribution, we have engineered a test that is not only statistically sound but also strategically targeted, it transforms a subtle discrepancy between two estimation philosophies into a powerful signal for distributional misspecification, therefore, we conclude that our ratio-based test represents a significant advancement in the field and recommend its adoption as a primary diagnostic tool for researchers and practitioners working with data hypothesized to follow a Geometric model.

References

Anderson, T. W. (2025). Anderson-Darling tests of goodness-of-fit. In *International Encyclopedia of Statistical Science* (pp. 78-81). Berlin, Heidelberg: Springer Berlin Heidelberg.

Di Noia, A., Barabesi, L., Marcheselli, M., Pisani, C., & Pratelli, L. (2023). Goodness-of-fit test for count distributions with finite second moment. *Journal of Nonparametric Statistics*, *35*(1), 19-37.

Leonenko, N., Makogin, V., & Cadirci, M. S. (2021). The entropybased goodness of fit tests for generalized von Mises-Fisher distributions and beyond. *Electronic Journal of Statistics*, *15*(2), 6344-6381.

Batsidis, A., & Lemonte, A. J. (2023). On goodness-of-fit tests for the Neyman type A distribution. *REVSTAT-Statistical Journal*, *21*(2), 143-171.

Markatou, M., & Liu, A. (2022). Statistical Distances in Goodness-of-fit. *Trends in Mathematical, Information and Data Sciences: A Tribute to Leandro Pardo*, 213-222.

Coronel-Brizio, H. F., Hernández-Montoya, A. R., Rodríguez-Achach, M. E., Tapia-McClung, H., & Trinidad-Segovia, J. E. (2024). Anderson-Darling and Watson tests for the geometric distribution with estimated probability of success. *PloS one*, *19*(12), e0315855.

Milošević, B., Jiménez-Gamero, M. D., & Alba-Fernández, M. V. (2021). Quantifying the ratio-plot for the geometric distribution. *Journal of Statistical Computation and Simulation*, *91*(11), 2153-2177.

Madukaife, M. S., Nduka, U. C., & Ossai, E. O. (2025). Testing for multinormality with goodness-of-fit tests based on phi divergence measures. *Statistics in Transition new series*, *26*(2), 129-149.

Vaisakh, K. M., Xavier, T., & Sreedevi, E. P. (2023). Goodness of fit test for Rayleigh distribution with censored observations. *Journal of the Korean Statistical Society*, *52*(4), 794-815.

Alizadeh Noughabi, H., & Shafaei Noughabi, M. (2023). A new estimator of Kullback–Leibler information based on a local linear regression and its application in goodness-of-fit tests. *Journal of Statistical Computation and Simulation*, *93*(16), 2828-2842.

Erlemann, R., & Lindqvist, B. H. (2022). Conditional Goodness-of-Fit Tests for Discrete Distributions. *Journal of Statistical Theory and Practice*, *16*(1), 8.

Mohammadi, M., Hashempour, M., & Kamari, O. (2024). On the dynamic residual measure of inaccuracy based on extropy in order statistics. *Probability in the Engineering and Informational Sciences*, *38*(3), 481-502.

Qayoom, D., Rather, A. A., Alsadat, N., Hussam, E., & Gemeay, A. M. (2024). A new class of Lindley distribution: System reliability, simulation and applications. *Heliyon*, *10*(19).

Bhunia, S., & Banerjee, P. (2022). Some properties and different estimation methods for inverse A (α) distribution with an application to tongue cancer data. *Reliability: Theory & Applications*, *17*(1 (67)), 251-266.

Alghamdi, S. M., Bantan, R. A., Hassan, A. S., Nagy, H. F., Elbatal, I., & Elgarhy, M. (2022). Improved EDF-based tests for Weibull distribution using ranked set sampling. *Mathematics*, *10*(24), 4700.

Berlinger, M., Kolling, S., & Schneider, J. (2021). A generalized Anderson–Darling test for the goodness-of-fit evaluation of the fracture strain distribution of acrylic glass. *Glass Structures & Engineering*, *6*(2), 195-208.

Xavier, T., Vaisakh, K. M., & Sreedevi, E. P. (2025). Goodness-of-fit tests for inverse Gaussian distribution in the presence and absence of censoring. *Journal of Statistical Computation and Simulation*, *95*(5), 1010-1032.

González-Delgado, J., González-Sanz, A., Cortés, J., & Neuvial, P. (2023). Two-sample goodness-of-fit tests on the flat torus based on Wasserstein distance and their relevance to structural biology. *Electronic Journal of Statistics*, *17*(1), 1547-1586.

Ahsan-ul-Haq, M., Al-Bossly, A., El-Morshedy, M., & Eliwa, M. S. (2022). Poisson XLindley distribution for count data: statistical and reliability properties with estimation techniques and inference. *Computational Intelligence and neuroscience*, *2022*(1), 6503670.

Zhao, W., Yu, J., & Wu, C. (2025). Self-starting monitoring of the progressive type II censoring data based on goodness-of-fit test. *Quality Technology & Quantitative Management*, *22*(1), 32-54.

Pitera, M., Chechkin, A., & Wyłomańska, A. (2022). Goodness-of-fit test for α-stable distribution based on the quantile conditional variance statistics. *Statistical Methods & Applications*, *31*(2), 387-424.

Liebenberg, S., & Allison, J. (2023). A review of goodness-of-fit tests for the Rayleigh distribution. *Austrian Journal of Statistics*, *52*(1), 1-22.

Goel, N., & Krishna, H. (2022). Different methods of estimation in two parameter Geometric distribution with randomly censored data. *International Journal of System Assurance Engineering and Management*, *13*(4), 1652-1665.