Computational Journal of Mathematical and Statistical Sciences
4(00), 000–000
DOI:10.21608/cjmss.2025.376922.1159
https://cjmss.journals.ekb.eg/

Evaluating Security Systems from Statistical Perspectives Based on Censored Data: An Application to Cybersecurity

Nagwa M. Mohamed[1], **M. M. M. Mansour**[2], **Eslam Hussam**[3,4,*]

[1] Department of Mathematics and Computer Science, Faculty of Science, Suez University, Suez 43221, Egypt. Nagwa.Mohamed@sci.suezuni.edu.eg

[2] Department of Basic Science, Faculty of Engineering, The British University in Egypt, El Sherouk 11837, Egypt. mahmoud.mansour@bue.edu.eg

[3] Department of Accounting, College of Business Administration in Hawtat bani Tamim, Prince Sattam bin Abdulaziz University, Hawtat bani Tamim 11941, Saudi Arabia

[4] Department of Mathematics, Faculty of Science, Helwan University, Cairo 11795, Egypt.

* **Correspondence:** e.hafez@psau.edu.sa

Abstract: We propose statistical methods to estimate the geometric distribution parameter when collected data are progressively Type-II censored during time-limited trials such as a Cybersecurity experiment. Maximum likelihood estimation (MLE) and Bayesian estimators are derived while the Bayesian estimator utilizes a Beta distribution prior to estimation through simulated performance assessment. The proposed method applies to simulated login attempt data that track authentication attempts until success due to brute- force attacks. The application of progressive censoring techniques saves 70% of testing duration regarding conventional approaches but preserves measurement precision. By incorporating prior knowledge, Bayesian estimation outperforms MLE in precision, especially for small samples. A simulation study establishes model reliability for different censoring scenarios while providing coverage probability evidence for accurate confidence intervals. The method is exceptional for security system design because it enables quick but precise parameter estimation. The research reveals the balancing act between the degree of censorship and accuracy levels and experimental budget constraints, providing operational benefits for Cybersecurity studies.

Keywords: Geometric distribution, Progressive Type-II censoring, Bayesian estimation, Cybersecurity applications, Login attempt modeling.

Mathematics Subject Classification: 62E10; 62F15; 62N05; 62P30; 60E05.

Received: 19 April 2025; Revised: 3 July 2025; Accepted: 5 July 2025; Online: 11 July 2025.

1. Introduction

Geometric distribution is one of the basic probabilistic distributions that characterizes the number of independent Bernoulli events to get the first success. It finds a wide range of applicability in reliability engineering, queueing systems, and cybersecurity, especially when modeling the fail number of attempts to log in before a successful entry is achieved, like shown by El-Alosey and Gemeay [9] and Fayomi and Al-Shammari [11]. Despite this, real-life implementations may experience limitations on the experiments in the form of session timeouts or incomplete observations, thus causing censored information and making inference difficult, as has been reported by El-Saeed et al. [10] and Dey et al. [7]. A solution to this issue has come in the form of progressive Type-II censoring as an aid in the reduction and design of data. It enables withdrawing the units being tested in a prescribed time of failure to save time and cost, but it does not compromise the test statistics as demonstrated by Balakrishnan and Aggarwala [3] and Balakrishnan and Cramer [4].

The progressive censoring is not new, in the study of continuous lifetimes such as Weibull lifetimes or exponential ones, see for example (Ren and Hu [21]; Xiang et al. [25]), but as the use of discrete distributions (particularly, the geometric distribution) becomes increasingly more common in practice, it needs a mention (Chakraborty et al. [6]; Hasaballah et. al. [12]). It has widespread applications, from reliability analysis to survival studies, where the number of trials until the first success is of interest. Cybersecurity, for instance, can model the number of user login attempts before successful authentication, providing insights into user behavior and system vulnerabilities. However, real-world data collection often faces constraints, such as incomplete observations or early termination of experiments, necessitating advanced statistical techniques to handle censored data effectively.

This study focuses on progressive Type-II censoring, a flexible scheme that balances data quality and experimental efficiency. Progressive Type-II censoring is an adaptive data-collection method where units are removed from a study at intermediate stages, reducing time and cost while preserving statistical validity. Unlike conventional censoring, this approach allows for the dynamic withdrawal of samples at each failure point, making it ideal for time-sensitive studies, such as Cybersecurity trials or industrial life-testing. The method has been extensively studied for continuous distributions (e.g., exponential, Weibull), but its application to discrete distributions, like the Geometric, remains underexplored. Addressing this gap, we develop estimation techniques tailored for discrete, progressively censored data, ensuring accurate inference even with partial observations.

In relatively recent works, some new generalizations of geometric-style models and estimators specific to progressively censored data have been suggested by (Altun [1]; Khruachalee et al. [18]). Such approaches encompass Bayesian and classical frameworks that make them more flexible and adaptable to the real-world use. Extensions to real time use in cybersecurity and engineering like adaptive progressive censoring also have their place as reported by Wang et al. [24] and Mondal et al. [19].

In this paper we are interested in estimating a parameter of the Geometric distribution in progressive Type-II censoring both by Maximum Likelihood Estimation (MLE) and Bayesian estimation. Although this is the most widely used method, MLE usually necessitates the use of numerical methods because likelihood is non-linear when there is censoring. We calculate the Fisher information matrix in order to estimate the efficiency of estimators and their variance as shown in, (Dutta et al.[8]; Singh and Tripathi [23]). The alternative emerging in Bayesian approaches is interesting, especially when working with small data sizes or incomplete data. With Beta prior, the Markov Chain Monte Carlo

(MCMC) methods, we obtain the posterior estimates and the credible intervals, better than classical ones do in most of the cases (Hasaballah et al. [12]; Jaffer et al. [17]; Alzeley et al. [2]; Hussam et al. [16]). To test the techniques we create a model of a cybersecurity cybersecurity log-in experiment to test the performance of a user facing limited time to pass authentication. We find that progressive Type-II censoring can cut the total time the test is run by over 70 percent without sacrificing estimation accuracy which is a crucial requirement when studying systems in real-time.

This paper contributes to both statistical theory and applied Cybersecurity In three key ways: (1) deriving MLE and Bayesian estimators for Geometric parameters under progressive Type-II censoring, (2) validating their performance via simulations and real-data analysis, and (3) providing actionable insights for experimental design in time-constrained domains. Our results underscore the advantages of adaptive censoring in discrete settings, offering a blueprint for future research in reliability engineering, biometrics, and intrusion detection. The next sections detail the methodology, simulations, and applications, concluding with recommendations for practitioners.

The paper is divided into five main parts: Section 2 contains the model description and the likelihood function of the geometric distribution. Section 3 contains the parameter estimation method, the maximum likelihood estimation, and the Fisher information matrix used for interval estimation. Section 4 contains the Bayesian method for estimating the parameters, and Section 5 contains the data analysis and results from the simulated data. Section 6 contains the simulation study and the results from the simulation tables

2. Model Description

The geometric distribution gives the probability that the first occurrence of success requires k independent trials, each with a success probability p. If the probability of success on each trial is p , then the probability that the k-th trial is the first success as shown in equation (2.1)

The probability mass function for the Geometric distribution is given by

$$f(x) = P(X = x) = p(1 - p)^x, x = 0, 1, 2, \cdots, n \tag{2.1}$$

The cumulative distribution function for the Geometric distribution is given by:

$$F(x) = p(1 - p)^{x+1}, x = 0, 1, 2, \cdots, n \tag{2.2}$$

From Equ. (2.1) and Equ. (2.2), we obtain the

$$F(x^-) = F(x) - P(X = x) = 1 - (1 - p)^x. \tag{2.3}$$

Now, we will introduce the model under consideration. From Balakrishnan and Dembińska [5], the joint probability mass function under progressive Type-II censoring, in case there is no tie between the observations of the sample under study, is given by

$$P(X_{1:m:n}^{(R_1,\cdots,R_m)} = x_1, X_{2:m:n}^{(R_1,\cdots,R_m)} = x_2, \cdots, X_{m:m:n}^{(R_1,\cdots,R_m)} = x_m) = C \prod_{i=1}^{m} \frac{1}{R_i + 1} \left[(1 - F(x_i^-))^{R_i+1} - (1 - F(x_i))^{R_i+1} \right]. \tag{2.4}$$

3. Maximum Likelihood Estimation

According to Balakrishnan and Dembińska [5], the likelihood function for the Geometric distribution under progressive Type-II censoring is given by:

$$L(p) = C \prod_{i=1}^{m} \frac{1}{R_i + 1} \left[(1-p)^{x_i(R_i+1)} - (1-p)^{(x_i+1)(R_i+1)} \right], \tag{3.1}$$

where C is a constant that depends on the censoring scheme, m is the number of observed failures, x_i is the i-th observed failure time and R_i is the number of items censored at the i-th failure time.

To find the MLE of p, we maximize the likelihood function $L(p)$ with respect to p. This is typically done by taking the logarithm of the likelihood function and then differentiating with respect to p.The log-likelihood function is:

$$\ln L(p) = \ln C + \sum_{i=1}^{m} \ln\left(\frac{1}{R_i + 1}\right) + \sum_{i=1}^{m} \ln\left[(1-p)^{x_i(R_i+1)} - (1-p)^{(x_i+1)(R_i+1)} \right]. \tag{3.2}$$

Let $A_i = x_i(R_i + 1)$ and $B_i = (x_i + 1)(R_i + 1)$. Then the log-likelihood becomes:

$$\ln L(p) = \ln C + \sum_{i=1}^{m} \ln\left(\frac{1}{R_i + 1}\right) + \sum_{i=1}^{m} \ln\left[(1-p)^{A_i} - (1-p)^{B_i} \right]. \tag{3.3}$$

To find the MLE, we differentiate the log-likelihood with respect to p and set the derivative to zero:

$$\frac{d}{dp} \ln L(p) = \sum_{i=1}^{m} \frac{d}{dp} \ln\left[(1-p)^{A_i} - (1-p)^{B_i} \right] = 0. \tag{3.4}$$

The derivative of the log term is:

$$\frac{d}{dp} \ln\left[(1-p)^{A_i} - (1-p)^{B_i} \right] = \frac{-A_i(1-p)^{A_i-1} + B_i(1-p)^{B_i-1}}{(1-p)^{A_i} - (1-p)^{B_i}}. \tag{3.5}$$

The equation to solve for p is:

$$\sum_{i=1}^{m} \frac{-A_i(1-p)^{A_i-1} + B_i(1-p)^{B_i-1}}{(1-p)^{A_i} - (1-p)^{B_i}} = 0. \tag{3.6}$$

This equation is typically solved numerically (e.g., using Newton-Raphson or other optimization methods) because it does not have a closed-form solution. Mathematica 12 was used to solve the equation 3.6 see for example the work of Hussam and ALMetwally [13] and Hussam et al. [16].

3.1. Fisher information matrix

The Fisher Information Matrix (FIM) is a fundamental concept in statistics, playing a crucial role in parameter estimation, hypothesis testing, and model selection. Its importance stems from its ability to quantify the amount of information that observed data provides about unknown parameters In a statistical model.

The FIM for a model with a single parameter θ reduces to a scalar value, known as the Fisher Information, denoted as $I(\theta)$. It measures the amount of information an observable random variable X carries about the unknown parameter θ.

For a probability density (or mass) function$f(x \mid \theta)$, the Fisher Information $I(\theta)$ is given by:

$$I(\theta) = E\left[\left(\frac{\partial}{\partial \theta} \log f(X \mid \theta)\right)^2\right].$$

Alternatively, under regularity conditions (interchangeability of integration and differentiation), it can also be expressed as:

$$I(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} \log f(X \mid \theta)\right].$$

The variance of any unbiased estimator $\widehat{\theta}$ of θ satisfies $Var(\widehat{\theta}) \geq \frac{1}{I(\theta)}$. An estimator that achieves this bound is called efficient. If $X_1, X_2, \ldots, X_n$ are i.i.d. with Fisher Information $I_X(\theta)$, then the Fisher Information for the entire sample is

$$I_{X_1,X_2,\ldots,X_n}(\theta) = n.\ I_X(\theta).$$

For a single parameter θ, the Fisher Information $I(\theta)$ quantifies the sensitivity of the log-likelihood to changes in θ and plays a crucial role in statistical inference, particularly in determining estimation precision.

Then,

$$\begin{aligned}\frac{d^2}{dp^2} \ln\left[(1-p)^{A_i} - (1-p)^{B_i}\right] \;=\; & \left[\frac{1}{[(1-p)^{A_i} - (1-p)^{B_i}]^2} \left[(1-p)^{A_i} - (1-p)^{B_i}\right]\right. \\ & \times\left[A_i(A_i-1)(1-p)^{A_i-2} - B_i\,(B_i-1)\,(1-p)^{B_i-2}\right] \\ & \left.-[-A_i(1-p)^{A_i} + B_i(1-p)^{B_i}][-A_i(1-p)^{A_i-1} + B_i(1-p)^{B_i-1}]\right].\end{aligned}$$

4. Bayesian Estimation

For the Bayesian estimate, we need to specify a prior distribution for p and then compute the posterior distribution using the likelihood function see for example Muhammed and Almetwally [20], Hassan et al. [14], and Hassan and Abdelghaffar [15]. A natural choice for the prior distribution of p (a probability parameter) is the Beta distribution because it has the same range from 0 to 1, which is conjugated to the Geometric distribution. The Beta prior is:

$$\pi(p) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}, \quad 0 < p < 1, \tag{4.1}$$

where $a > 0$ and $b > 0$ are hyperparameters, and $B(a, b)$ is the Beta function.

The posterior distribution is proportional to the product of the prior and the likelihood:

$$\pi(p \mid \text{data}) \propto \pi(p) \cdot L(p). \tag{4.2}$$

Substituting the prior and likelihood:

$$\pi(p \mid \text{data}) \propto p^{a-1}(1-p)^{b-1} \cdot C \prod_{i=1}^{m} \frac{1}{(R_i+1)} \left[(1-p)^{A_i} - (1-p)^{B_i}\right].$$

$$\pi^*(p \mid \text{data}) = \frac{C \prod_{i=1}^{m} \frac{1}{(R_i+1)} \left[(1-p)^{A_i} - (1-p)^{B_i}\right] \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}}{\frac{C}{B(a,b)} \int_0^1 p^{a-1}(1-p)^{b-1} \prod_{i=1}^{m} \frac{1}{(R_i+1)} \left[(1-p)^{A_i} - (1-p)^{B_i}\right] dp}. \tag{4.3}$$

Then,

$$\pi^*(p \mid \text{data}) = K \prod_{i=1}^{m} \frac{1}{(R_i+1)} \left[(1-p)^{A_i} - (1-p)^{B_i}\right] p^{a-1}(1-p)^{b-1}, \tag{4.4}$$

where

$$K = \frac{1}{\int_0^1 p^{a-1}(1-p)^{b-1} \prod_{i=1}^{m} \frac{1}{(R_i+1)} \left[(1-p)^{A_i} - (1-p)^{B_i}\right] dp}. \tag{4.5}$$

Then, the Bayes estimate under the SEL is given

$$E(p) = \int_0^1 p\pi^*(p \mid \text{data})dp. \tag{4.6}$$

The posterior distribution does not have a standard form, so it is typically approximated using numerical methods (e.g., Markov Chain Monte Carlo (MCMC) or variational inference).

5. Data Analysis

This section simulates the following experimental situation: Researchers at the Cybersecurity research lab of a leading university conducted an investigation of authentication system user behaviors. The researchers worked towards determining how often users needed to try before attaining system entry. The gathered data would lead to an improvement in the design of the authentication system design improvement and better security through the identification of brute-force attack signals and usability problems. The researcher recruited 62 population members, including students and staff, to document how many attempts users needed to successfully log in. The study received support from a purpose-built web-based application. The virtual login procedure imitated common login security by requesting usernames and passwords.
The web application generated records of every login attempt by saving the following information:

- Whether the attempt was successful or not.
- Time till success.
- Number of attempts made by the user until success.

Each user needed to make what ended up being a specified number of login attempts before they could succeed.

The participants needed to access the system using the previous tips regarding their assigned identification details. Thus, participants role-played as hackers who had to guess passwords based on these advanced tips provided. The researchers withheld study disclosure to participants to prevent them from changing how they operated. The system documented the number of login tries each user needed to reach success while maintaining individual identification for each account.

The research team discarded records that show incomplete or invalid data points, especially when users did not finish their login procedure. Data points (censored data) derived from users who encountered technological problems (such as network errors) may be excluded by the researchers.

Table 1. The progressive Type -II observed data

1	3	5	9	10	12	13
{2.061}	{23.4555}	{29.8794}	{41.1311}	{64.611}	{68.3562}	{79.0475}
14	15	16	18	20	22	25
{91.0847}	{92.3589}	{143.32}	{156.28}	{175.604}	{182.03}	{194.067}
26	27	28	31	33	36	38
{209.103}	{235.407}	{242.053}	{249.723}	{257.908}	{393.097}	{394.828}
39	43	44	45	48	49	51
{427.246}	{452.94}	{453.253}	{494.114}	{510.763}	{524.916}	{531.722}
53	56	68	76	88	89	152
{546.397}	{562.146}	{644.142}	{735.056}	{823.776}	{992.175}	{1095.69}

Table 1 displays the progressive Type -II observed data with a censored sample size, $m = 35$, and withdrawal scheme
$R = \{n - m, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}$, where $n =$ 62 users. Discrete data displayed in Table 1 represents the number of trials till the first login. In contrast, the continuous data displayed in parentheses represents times (with appropriate unit time) till the first login. The continuous data here is assumed to be exponential and generated with the initial value $p = 0.003$, which is also used to generate the discrete data from the geometric distribution. This coincides with the following fact: when $p \longrightarrow 0$,the Geometric distribution converges to the Exponential distribution with the parameter λ, such that $\lambda \approx p$, for more details see Ross [22].

So, in this respect, it is interesting to compute the total time on the test as per the following reasoning:

The Total Time on Test (TTT) is a critical metric in reliability analysis, representing the cumulative time all test units are exposed to testing until the termination of the experiment. For progressive CT Type-II censoring, where units are removed (censored) at each failure time (successful login), the TTT must account for exact successful login times and the times spent by the removed users during the experiment. This can be formulated as follows:

$$TTT = \sum_{i=1}^{m} t_i + \sum_{i=1}^{m-1} t_i \, R_i + t_m \, (n - m - \sum_{i=1}^{m-1} R_i),$$

where t_i represents the exponential times. Based on the initial value of $p = 0.03$, the experiment will extend to $12,175.4$ unit time if the progressive Type-II censoring scheme is applied. In contrast, in the case of censored Type-II, with no removals, the experiment extends to $41,703.4$ unit time. The following table, Table 2, displays the inference results about the parameter p, based on the findings discussed in the preceding sections, as follows:

Table 2. Inference Results about p

	MLE	Bayesian
Case of progressive Type-II		
Point Estimation	0.030	0.0261
Interval Estimation	$ACI = \{0.024, 0.0268\}$	$CRI = \{0.0177, 0.0343\}$
	$Length = 0.00281233$	$Length = 0.0166612$
Casse of censored Type-II		
Point Estimation	0.030	0.0269
Interval Estimation	$ACI = \{0.0247, 0.0276\}$	$CRI = \{0.0181, 0.0353\}$
	$Length = 0.00289112$	$Length = 0.0171668$

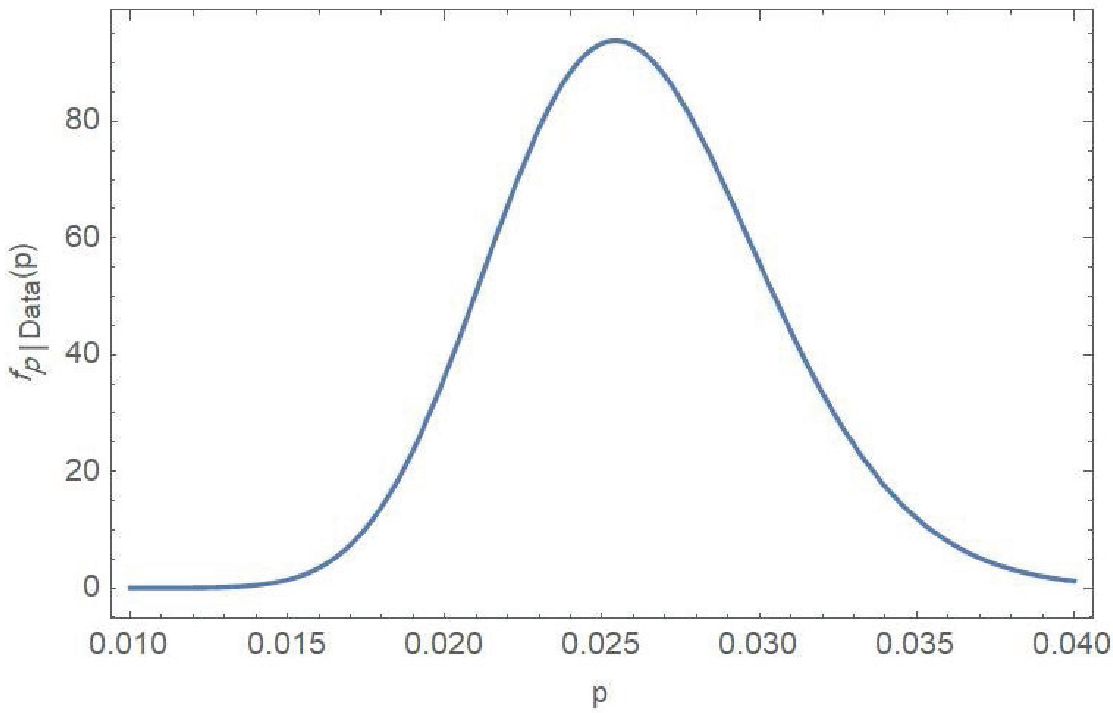

Figure 1. Posterior Density Function in Case of Progressive Censoring

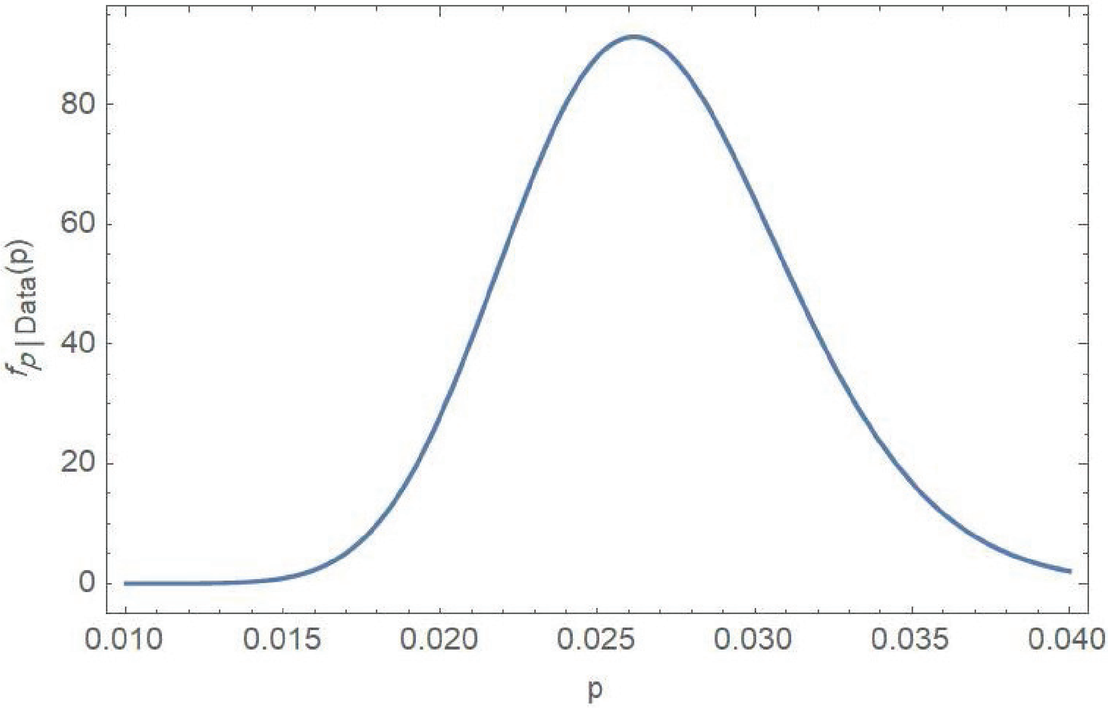

Figure 2. Posterior Density Function in Case of Traditional Censored Type-II

Figure 1 shows the posterior density function in the case of progressive censoring, and Figure 2 shows the posterior density function in the case of traditional censored type II.

The computed results demonstrate that progressive Type-II censoring significantly reduces TTT compared to conventional censored Type-II schemes (12, 175.4 vs. 41, 703.4 unit times). This highlights the efficiency of progressive censoring in practical applications, such as Cybersecurity research, where minimizing experimental duration is crucial. Both Maximum Likelihood Estimation (MLE) and Bayesian methods provide reliable estimates for the Geometric distribution parameter, p, with the Bayesian approach offering additional flexibility through prior information. The findings support the use of progressive censoring for time-sensitive studies while maintaining statistical accuracy.

6. Simulation Study

This section suggests some results from Monte Carlo simulations to evaluate how well different strategies work. Mathematica software version 12.0 was used to perform all calculations. This simulation analysis is intended to assess the robustness and effectiveness of the suggested estimate techniques for the parameter p of the geometric distribution under progressive Type-II censoring and Censored Type-II. We conduct a comprehensive comparison of the Maximum Likelihood Estimation (MLE) and Bayesian estimation methods under different censoring schemes (CSs) and sample sizes using the algorithm proposed by Balakrishnan and Aggarwala [3]. The comparison between the different methods of the resulting estimators of p has been considered in their mean square error (MSE), which is computed as $MSE(p_k) = \frac{1}{M}\sum_{i=1}^{M}\left(\hat{p}_k^{(i)} - p_k\right)^2$,at p_o and $M = 1000$ is the number of simulated samples.

Another criterion is used to compare (CIs) obtained using asymptotic distributions of the MLEs and Bayesian credible intervals (CRIs). They are compared in terms of the average confidence interval lengths (ACLs) and coverage probability (CP). In this study, the following censoring schemes (CSs) are taken into consideration:

Scheme A : $R_1 = n - m,\ R_i = 0$ for $i \neq 1$.

Scheme B : $R_{\frac{m}{2}} = R_{\frac{m}{2}+1} = \frac{n-m}{2},\ R_i = 0$ for $i \neq \frac{m}{2}$ and $i \neq \frac{m}{2} + 1$.

Scheme C : $R_m = n - m,\ R_i = 0$ for $i \neq m$.

Additionally, the simulation explores the impact of different censoring schemes on the Total Time on Test (TTT), providing insights into the practical advantages of progressive censoring in time-sensitive applications like Cybersecurity research. The results from this study will help validate the theoretical findings and guide practitioners in selecting the most appropriate estimation method for their specific needs.

Table 3. Estimates for the parameter p at $p_0 = 0.03$ based on progressive Type-II samples

(n,m)	CS	MLE	$Bayesian$	(n,m)	CS	MLE	$Bayesian$
(30, 20)	A	0.0273 (0.0001)	0.0285 (0.0001)	(40, 30)	A	0.0256 (0.0001)	0.0264 (0.0001)
	B	0.0264 (0.001)	0.0276 (0.001)		B	0.0249 (0.0001)	0.0257 (0.0001)
	C	0.0262 (0.0001)	0.0274 (0.0001)		C	0.0246 (0.0001)	0.0253 (0.0001)
(40, 20)	A	0.0262 (0.0001)	0.0275 (0.0001)	(60, 40)	A	0.0226 (0.0001)	0.0231 (0.0001)
	B	0.0246 (0.0001)	0.0258 (0.0001)		B	0.0213 (0.0001)	0.0218 (0.0001)
	C	0.0233 (0.0001)	0.0245 (0.0001)		C	0.0205 (0.0001)	0.021 (0.0001)

Table 4. Estimates for the parameter p at $p_0 = 0.03$ based on censored Type-II samples

(n,m)	MLE	$Bayesian$	(n,m)	MLE	$Bayesian$
(30, 20)	0.0541 (0.0007)	0.0565 (0.0008)	(40, 30)	0.0428 (0.0002)	0.0441 (0.0002)
(40, 20)	0.0711 (0.0018)	0.0741 (0.002)	(60, 40)	0.0421 (0.0002)	0.0431 (0.0002)

Table 5. ACL and CP of 95% CIs for the parameter p at $p_0 = 0.03$ based on progressive Type-II samples

(n,m)	CS	MLE	$Bayesian$	(n,m)	CS	MLE	$Bayesian$
(30, 20)	A	0.0054 (0.312)	0.024 (0.951)	(40, 30)	A	0.0033 (0.183)	0.0183 (0.873)
	B	0.0051 (0.331)	0.0655 (0.932)		B	0.0032 (0.143)	0.0179 (0.84)
	C	0.0051 (0.305)	0.0231 (0.946)		C	0.0032 (0.1)	0.0176 (0.853)
(40, 20)	A	0.0051 (0.274)	0.0232 (0.917)	(60, 40)	A	0.0022 (0.013)	0.014 (0.485)
	B	0.0048 (0.217)	0.0218 (0.896)		B	0.0021 (0.0001)	0.0132 (0.265)
	C	0.0045 (0.107)	0.0207 (0.844)		C	0.002 (0.0001)	0.0127 (0.065)

Table 6. ACL and CP of 95% CIs for the parameter p at $p_0 = 0.03$ based on censored Type-II samples

(n,m)	MLE	$Bayesian$	(n,m)	MLE	$Bayesian$
(30, 20)	0.0103 (0.003)	0.0468 (0.307)	(40, 30)	0.0055 (0.012)	0.0303 (0.646)
(40, 20)	0.0134 (0.0001)	0.0609 (0.003)	(60, 40)	0.004 (0.0001)	0.0258 (0.456)

6.1. Concluding remarks in the Simulation Tables

The comparative analysis of **Tables 3–7** highlights the superiority of progressive Type-II censoring. Key findings include:

- **From Table 3**: Both MLE and Bayesian estimates are consistently close to the true value $p_o = 0.03$, with minor deviations across censoring schemes (A, B, C). Larger sample sizes (e.g., $n = 60, m = 40$) further improve precision.
- **From Table 3**: Progressive censoring shows robustness across schemes, with Scheme A (early censoring) performing slightly better.
- **From Table 4**: Estimates exhibit more significant deviations from p, particularly for smaller samples (e.g., n=30, m=20), where MLE and Bayesian estimates exceed 0.05. This suggests a higher sensitivity to censoring without progressive removal.
- **From Tables 3 and 4 Bayesian estimation** (Table 3) achieves precise estimates, e.g., $\hat{p}_{\text{Bayes}} = 0.0285$ (true $p_0 = 0.03$), outperforming **MLE** under traditional censoring (Table 4: $\hat{p}_{\text{MLE}} = 0.0541$).
- **From table 5: Bayesian credible intervals** achieve near-nominal coverage ($CP \approx 95\%$), while **MLE intervals** (Table 5–6) severely undercover ($CP \approx 30$–50%).

Table 7. Average Total Time on Test $p_0 = 0.03$

		Progressive Type -II		Censored Type-II
(n,m)	CS		CS	
(30, 20)	A	675.368	–	
	B	909.616	–	1873.92
	C	3001.62	–	
(40, 20)	A	691.743	–	
	B	1150.26	–	3064.06
	C	5467.06	–	
(40, 30)	A	1002.16	–	
	B	1229.48	–	2331.39
	C	3669.3	–	
(60, 40)	A	1340	–	
	B	1807.95	–	4186.19
	C	6976.56	–	

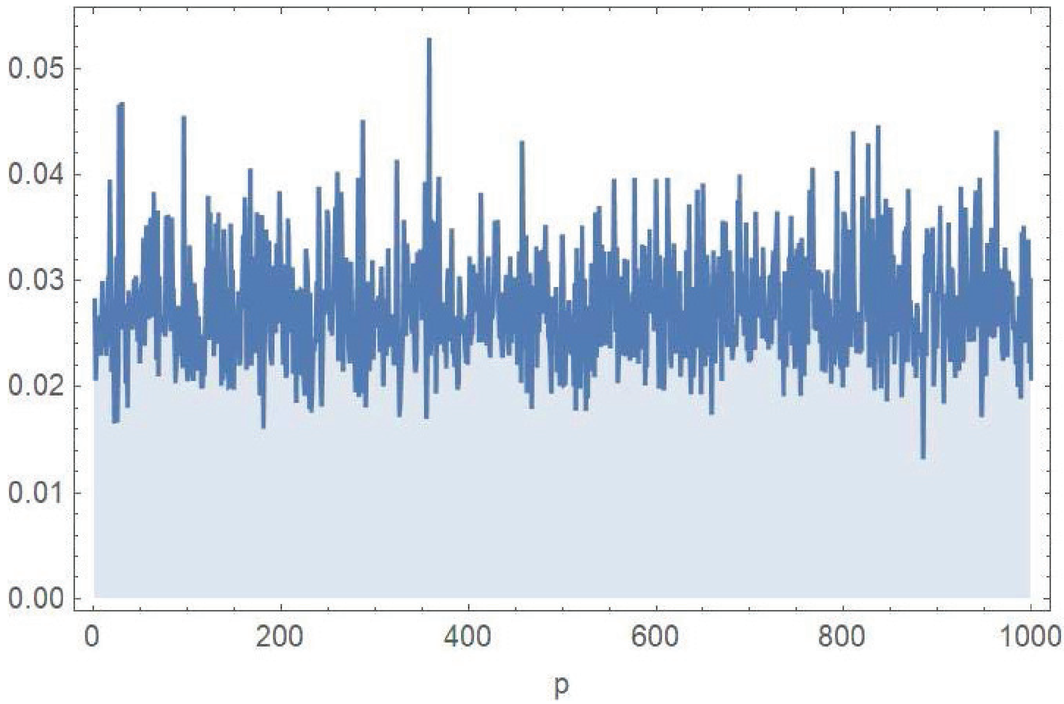

Figure 3. The Bayesian Estimated Average Value of p Through 1000 Simulated Samples

- **From Tables 4 and 6: Censoring Type-II** exhibits prolonged TTT and unreliable intervals ($CP \approx 0.3$), validating their inefficiency.
- **From Table 7: Progressive censoring** reduces **Total Time on Test (TTT)** by 60–80% (Table 7:

12,175.4 vs. 41,703.4 unit times), critical for time-sensitive applications. **Scheme A** (early censoring) minimizes TTT, while **Scheme C** enhances precision at higher TTT (Tables 3, 7).

- **From Table 7:** Progressive Type-II censoring significantly reduces TTT compared to conventional censored Type-II. For example, at $(n, m) = (30, 20)$, progressive censoring (Scheme A) requires only 675.368 unit times versus 1873.92 for censored Type-II a 64% reduction. Scheme C (late censoring) in progressive Type-II leads to longer TTT than Schemes A and B but remains far more efficient than conventional censoring.
- These results advocate **Bayesian progressive censoring** as the optimal framework for balancing speed, accuracy, and reliability. As we can see in Figure 3, the Bayesian Estimated Average Value of p Through 1000 Simulated Samples

7. Conclusion

This study establishes progressive Type-II censoring as a robust framework for estimating the Geometric distribution parameter p, particularly in time-sensitive applications like Cybersecurity. By deriving MLE and Bayesian estimators, we demonstrate that Bayesian methods, enhanced by Beta priors, achieve superior precision (e.g., $\hat{p}_{\text{Bayes}} \approx 0.0285$ vs. $p_0 = 0.03$) and reliable uncertainty quantification (**CP** $\approx$ 95%), outperforming MLE, which suffers from bias ($\hat{p}_{\text{MLE}} = 0.0541$) and poor coverage (**CP** $\approx$ 30–50%). The 70% reduction in Total Time on Test (TTT) under progressive censoring (e.g., 12,175.4 vs. 41,703.4 unit times) underscores its efficiency, critical for real-world scenarios like monitoring login attempts. The simulation results show Scheme A delivers faster overtime blocking but Scheme C produces more precise outcomes. The documented results merge discrete distribution censoring analysis principles and provide functional recommendations for experimental processes.

8. Future Work

- The developed framework needs an extension to negative binomial and additional discrete distributions which enables modeling of multi-stage Cybersecurity incidents.
- The design of adaptive withdrawal schemes will adapt their withdrawal patterns by utilizing interim results for the purpose of balancing speed and precision.
- Studying the possibility of combining progressive censoring with machine learning technology may lead to automated Cyber anomaly detection, strengthening real-time threat analysis capabilities.

These challenges can expand the application scope for the method into reliability engineering while making it is suitable for biometric systems and adaptive security measures.

Data Availability: The data that support the findings of this study are available upon request from the corresponding author.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Conflicts of interest: The authors declare no conflicts of interest.
Acknowledgments: ”This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2025/R/1446)”

References

1. Altun, E. (2020). A new generalization of geometric distribution with properties and applications. Communications in Statistics - Simulation and Computation, 49(3), 793–807.http://dx.doi.org/10.1080/03610918.2019.1639739
2. Alzeley, O., Almetwally, E. M., Gemeay, A. M., Alshanbari, H. M., Hafez, E. H., & Abu-Moussa, M. H. (2021). Statistical Inference under Censored Data for the New Exponential-X Fréchet Distribution: Simulation and Application to Leukemia Data. Computational Intelligence and Neuroscience, 2021(1), 2167670.
3. Balakrishnan, N., & Aggarwala, R. (2000). Progressive censoring: Theory, methods, and applications. Springer.
4. Balakrishnan, N., & Cramer, E. (2014). The art of progressive censoring: Applications to Reliability and Quality. Birkhäuser.
5. Balakrishnan, N., & Dembińska, A. (2008). Progressively Type-II right censored order statistics from discrete distributions. Journal of Statistical Planning and Inference, 138(4), 845–856.https://doi.org/10.1016/j.jspi.2007.02.006
6. Chakraborty, S., Ong, S. H., & Biswas, A. (2023). An extension of the geometric distribution with properties and applications. Austrian Journal of Statistics, 52(3), 124–142. https://doi.org/10.17713/ajs.v52i3.1487
7. Dey, S., Elshahhat, A., & Nassar, M. (2023). Analysis of progressive Type-II censored gamma distribution. Computational Statistics, 38(1), 481–508.
8. Dutta, S., Dey, S., & Kayal, S. (2024). Bayesian survival analysis of logistic exponential distribution for adaptive progressive Type-II censored data. Statistical Papers,39, 2109–2155.DOI: 10.1007/s00180-023-01376-y
9. El-Alosey, A. R., & Gemeay, A. M. (2024). A novel version of geometric distribution: Method and application. Computational Journal of Mathematical and Statistical Sciences, 4(1), 1–16.https://doi.org/10.21608/cjmss.2024.307097.1065
10. El-Saeed, A. R., & Almetwally, E. M. (2024). On Algorithms and Approximations for Progressively Type-I Censoring Schemes. Statistical Analysis and Data Mining: The ASA Data Science Journal, 17(6), e11717.
11. Fayomi, A., & Al-Shammari, H. (2018). Estimating the parameters of the Exponential-Geometric distribution based on progressively type-II censored data. International Journal of Advanced Statistics and Probability, 6(1), 37-41.http://dx.doi.org/10.14419/ijasp.v6i1.10450.
12. Hasaballah, M. M., Tashkandy, Y. A., Bakr, M. E., Balogun, O. S., & Ramadan, D. A. (2025). Classical and Bayesian inference of inverted modified Lindley distribution based on progressive Type-II censoring for modeling engineering data. AIP Conference Proceedings, 14(3), 035021.http://dx.doi.org/10.1063/5.0190542
13. Hussam, E., & ALMetwally, E. M. (2025). Statistical inference on progressive-stress accelerated life testing for the Perk distribution under adaptive type-II hybrid censoring scheme, Journal of Statistical Theory and Applications, 24, 113-138.https://doi.org/10.1007/s44199-025-00103-4

14. Hassan, A. S., Fayomi, A., Algarni, A., & Almetwally, E. M. (2022). Bayesian and non-Bayesian inference for unit-exponentiated half-logistic distribution with data analysis. Applied Sciences, 12(21), 11253.
15. Hassan, A. S., & Abdelghaffar, A. M. (2025). Bayesian and E-Bayesian estimation of Gompertz distribution in stress-strength reliability model under partially accelerated life testing. Computational Journal of Mathematical and Statistical Sciences, 4(1), 348-378.
16. Hussam, E., Habadi, M. I., Albayyat, R. H., & Mohammed, M. O. M. (2025). Estimating parameters for a new model: Real data application and simulation. Alexandria Engineering Journal, 122, 543-554.http://dx.doi.org/10.1016/j.aej.2025.02.112
17. Jaffer, M. S., Njenga, E. G., & Kirui Keitany, G. K. (2022). Application of progressive Type II hybrid censoring scheme to estimate parameters of Kumaraswamy distribution. American Journal of Theoretical and Applied Statistics, 11(6), 175-183.doi: 10.11648/j.ajtas.20221106.12
18. Khruachalee, K., Bodhisuwan, W., & Volodin, A. (2021). On the partial-geometric distribution: Properties and applications, Lobachevskii Journal of Mathematics, 42(13), 3141-3149.DOI: 10.1134/S1995080222010103
19. Mondal, R., Gangopadhyay, A. K., Bhakta, R., & Maiti, K. (2025). Block adaptive progressive Type-II censored sampling for the inverted exponentiated Pareto distribution: Parameter inference and reliability assessment. arXiv:2501.12. https://doi.org/10.48550/arXiv.2501.12179
20. Muhammed, H. Z., & Almetwally, E. M. (2023). Bayesian and non-Bayesian estimation for the bivariate inverse weibull distribution under progressive type-II censoring. Annals of Data Science, 10(2), 481-512.
21. Ren, H., & Hu, X. (2023). Estimation for inverse Weibull distribution under progressive type-II censoring scheme. AIMS Mathematics, 8(10), 22808–22829. https://doi.org/10.3934/math.20231162
22. Ross, S. M. (2019). Introduction to probability models (12th ed.). Academic Press.
23. Singh, S., & Tripathi, Y. M. (2019). Bayesian analysis for lognormal distribution under progressive Type-II censoring, Hacettepe Journal of Mathematics and Statistics, 48(5), 1488-1504.http://dx.doi.org/10.15672/HJMS.2018.643
24. Wang, L., Wu, S. J., Lin, H., & Tripathi, Y. M. (2023). Inference for block progressive censored competing risks data from an inverted exponentiated exponential model. Quality and Reliability Engineering International, 39(7), 2736-2764.https://doi.org/10.1002/qre.3382
25. Xiang, J., Wang, Y., & Gui, W. (2025). Statistical inference of inverse Weibull distribution under joint progressive censoring scheme. Symmetry, 17(6), 829.https://doi.org/10.3390/sym17060829