photonics

Article

Statistical Analysis of the Photon Loss in Fiber-Optic Communication

Artur Czerwinski [1,*] and Katarzyna Czerwinska [2]

[1] Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University in Torun, ul. Grudziadzka 5, 87-100 Torun, Poland
[2] Foundation for Responsible Education and Upbringing "Ad Veritatem", ul. Piskorskiej 4/4, 87-100 Torun, Poland
* Correspondence: aczerwin@umk.pl

Abstract: In optical communication systems, photons are lost due to the attenuation of the transmission medium. To efficiently implement quantum information protocols, we need to be able to precisely describe such processes. In this paper, we propose statistical methods to estimate the attenuation coefficient of the fiber link. By following the Beer–Lambert law, we utilize the properties of the exponential distribution to estimate the rate parameter based on observable data. In particular, we determine the explicit forms of unbiased estimators that are suitable for censored (truncated) sets of data. Moreover, we focus on minimum-variance methods that ensure a reliable estimation of the attenuation coefficient.

Keywords: attenuation; Beer–Lambert law; exponential distribution; unbiased estimators; photon loss; optical transmission; parametric point estimation

check for updates

Citation: Czerwinski, A.; Czerwinska, K. Statistical Analysis of the Photon Loss in Fiber-Optic Communication. *Photonics* **2022**, *9*, 568. https://doi.org/10.3390/photonics9080568

Received: 18 July 2022
Accepted: 11 August 2022
Published: 12 August 2022

1. Introduction

Optical fiber, which has emerged as the most commonly used transmission medium for long-distance communications, is a thin string of glass that guides a beam of light along its length [1]. The subfield of fibers has been developing rapidly both in terms of manufacturing and applications. Due to its flexibility and the ability to come in a bundle, optical fibers are used as a medium for telecommunication and computer networking. Through the use of wavelength-division multiplexing, each fiber can accommodate many independent communication channels, each working with a different wavelength of light [2]. However, fibers can also be used as: sensors [3,4], carriers of optical power (power-over-fiber technology) [5], light tubes [6], or components of a inspection instrument in imaging optics [7].

In quantum optics, photons are used to encode information by exploiting different degrees of freedom, including polarization, temporal, spatial, or spectral [8]. Most celebrated applications of quantum properties of light relate to quantum key distribution (QKD), which allows two parties to exchange a secret key [9]. QKD started in 1984 from a protocol that was based on polarization states of single photons [10]. Soon after, other approaches were proposed, including a protocol that employs quantum entanglement [11]. Nowadays, practical implementations of QKD protocols make the secure distribution of the secret key over optical fibers that are several hundred kilometers long [12] feasible.

Regardless of whether we operate in the quantum or classical regime, we need to take into account the photon loss that is connected with any transmission through an optical fiber. The reduction in the intensity of the light signal as it travels through the medium is called attenuation. Empirical research has demonstrated that attenuation in optical fibers is caused primarily by absorption and scattering. This phenomenon is an important factor that limits the transmission of a photonic signal across large distances. Every realization of a photon-based communication scheme requires devising an optical power budget that should comprise the acceptable attenuation [13].

In general, in quantum information theory, we discuss the quantum erasure channels (QEC) that relate to the lossy transmission of any physical particles used to encode information [14]. While photons are well-suited for the transmission of quantum information, protecting them from an external environment is very difficult due to the extreme fragility of photonic states. In particular, the photon loss is the most problematic type of error that arises in any quantum communication protocol. However, quantum error correction codes have proved to be efficient at conserving quantum information; see, for example, Refs. [15–17]. As for photonic states, general criteria for an error correction code that encodes qubits in bosonic states were introduced [18]. Another proposal to deal with the photon loss in quantum communication involves an error-correction scheme that encodes a two-photon state by using four photons, up to one of which can be lost in the transmission [19]. Alternatively, one can implement a three-photon code that protects one logical qubit against a photon loss [20].

In QKD, the photon loss caused by attenuation is a critical factor not only because of the reduction in the number of particles but also due to direct relations with other phenomena affecting the transmission. For instance, several scattering processes have an impact on the secure key rate in the presence of classical data signals, including Raman scattering, [21,22], guided acoustic wave Brillouin scattering [23], and double Rayleigh backscattering noise [24]. In each case, the secure key rate can be calculated only by taking into account the attenuation characteristics of the transmission medium.

In this paper, we propose approaching the problem of photon loss from the perspective of statistics. In quantum physics, estimation methods are usually implemented in the context of quantum state tomography [25–27] or quantum metrology [28,29]. The present paper achieves two goals—the first is that the attenuation coefficient is represented by means of a statistic that works as an estimator; the second is that optimal estimators among many possible are selected based on quantitative criteria. The key advantages of implementing the statistical treatment for the attenuation coefficient include the reliability and accuracy of estimation achieved through the minimum-variance requirement and unbiasedness, respectively.

2. Exponential Distribution and Fiber Attenuation

The methodology of the research was formulated on the grounds of the mathematical properties of the exponential distribution.

Definition 1. *We say that a random variable X has the exponential distribution with the parameter $\lambda > 0$, which is denoted by $X \sim \mathrm{Exp}(\lambda)$ if the probability density function (PDF) is given by [30]:*

$$f_\lambda(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \leqslant 0 \end{cases}$$

and cumulative distribution function (CDF):

$$F_\lambda(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0. \\ 0, & x \leqslant 0. \end{cases}$$

The parameter of the distribution, i.e., λ, is often called the rate parameter, whereas the inverse figure, λ^{-1}, is named the scale parameter. Next, one can also compute the expected value (mean) and the variance, respectively:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_\lambda(x)\mathrm{d}x = \frac{1}{\lambda}, \tag{1}$$

$$\mathrm{Var}[X] = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 f_\lambda(x)\mathrm{d}x = \int_{-\infty}^{\infty} x^2 f_\lambda(x)\mathrm{d}x - \mathbb{E}^2[X] = \frac{1}{\lambda^2}, \tag{2}$$

Another key characteristics of a probability distribution of a random variable are raw moments, μ_r for $r \in \mathbb{N}$, which are defined as $\mu_r := \mathbb{E}[X^r]$. These quantities can be computed from the moment generating function (MGF) [31]. For the exponential distribution, the MGF, denoted by $m_X(t)$, is

$$\begin{aligned} m_X(t) = \mathbb{E}[e^{tX}] &= \int_0^\infty e^{tx}\lambda e^{-\lambda x}\mathrm{d}x = \lambda \int_0^\infty e^{(t-\lambda)x}\mathrm{d}x \\ &= \lambda \frac{e^{(t-\lambda)x}}{t-\lambda}\Bigg|_0^\infty = \begin{cases} \infty, & t-\lambda \geqslant 0, \\ \frac{\lambda}{\lambda - t} & t - \lambda < 0 \end{cases}. \end{aligned}$$

If an MGF exists, then $m_X(t)$ is continuously differentiable in some neighborhood of the origin. If we differentiate the MGF r times with respect to t, and let $t \to 0$, we obtain

$$\frac{\mathrm{d}^r}{\mathrm{d}t^r} m_X(0) = \mathbb{E}[X^r] \equiv \mu_r,$$

which, for the exponential distribution, leads to

$$\mu_r = \frac{r!}{\lambda^r}.$$

Another important property of the exponential distribution is memorylessness, which can be formulated by means of the conditional probability.

Theorem 1. *An exponentially distributed random variable X obeys the relation*

$$P(X > x_1 + x_2 | X > x_1) = P(X > x_2), \qquad \forall x_1, x_2 \geq 0.$$

Proof.

$$\begin{aligned} P(X > x_1 + x_2 | X > x_1) &= \frac{P(X > x_1 + x_2 \cap X > x_1)}{P(X > x_1)} = \\ &= \frac{P(X > x_1 + x_2)}{P(X > x_1)} = \\ &= \frac{e^{-\lambda(x_1+x_2)}}{e^{-\lambda x_1}} = \\ &= e^{-\lambda x_2} \equiv P(X > x_2), \end{aligned}$$

□

In the proof, we used the fact that $P(X > x) = 1 - F_\lambda(x) = e^{-\lambda x} \equiv \overline{F}_\lambda(x)$. The function $\overline{F}_\lambda(x)$ is called the complementary cumulative distribution function (CCDF) and plays a central role in the context of physical applications of the exponential distribution.

When we consider the transmission of light through an optical fiber, we encounter losses, which are caused by absorption and scattering. This phenomenon is called the attenuation of an optical fiber. To quantify the photon loss, we follow the Beer–Lambert law, which measures the amount of light lost between the input and output:

$$I(L) = I_0 \, 10^{-\alpha L}, \tag{3}$$

where $I(L)$ denotes the output power after the transmission through a fiber of the length L, I_0 represents the input power, and α accounts for the attenuation. The attenuation coefficient α depends on the properties of the fiber. Recently, the Beer–Lambert law has already been implemented to study how a beam of photons is being attenuated by investigating the evolution of the Fock state representing the number of photons [32].

In physics, the exponential distribution can be applied to every process that involves an exponential decay; for example, a sample of a radionuclide that undergoes radioactive decay to a different state. Here, the Beer–Lambert law (3) can be equivalently written as $I(L) = I_0 \exp(-\lambda L)$, where $\lambda \equiv \ln 10 \, \alpha$. Furthermore, in the quantum regime, it implies that a single photon has been lost on the path L with the probability $1 - \exp(-\lambda L)$, whereas $\exp(-\lambda L)$ denotes the probability that it has survived. In other words, the Beer–Lambert law provides the CCDF for this process. Therefore, in the present manuscript, the distance that a photon covers before vanishing can be treated as a variable with the exponential distribution $L \sim \text{Exp}(\lambda)$.

The process of attenuation on the fundamental quantum level is in agreement with the memorylessness of the exponential distribution. If a photon travels through a fiber, one may assign time-dependent probabilities to the basis vectors in the Fock space [32]

$$\varrho(L) = \left(1 - e^{-\lambda L}\right) |0\rangle\langle 0| + e^{-\lambda L} \, |1\rangle\langle 1| \, , \tag{4}$$

where $|0\rangle$ represents the state when the photon has already vanished, and $|1\rangle$ is the situation when the photon keeps traveling. However, if, at some point, we witness the presence of the photon, the state collapses to $|1\rangle$ and the decay starts over again. Thus, the statistical property of memorylessness is equivalent to the quantum collapse that occurs upon measurement.

3. General Notes on Statistical Estimators

Estimators are used for calculating an estimate of a given quantity based on observed data. Let $\theta = (\theta_1, \ldots, \theta_k)$ denote the parameters that need to be determined (the estimand). Moreover, let Θ stand for the parameter space. The unknown parameters characterize the distribution of the random variable X. The sample space of all possible values of X is denoted by Ω. An estimator, denoted by $\hat{\theta}$, is a function of the data that maps the observed outcomes onto the parameter space, i.e., $\hat{\theta} : \Omega \to \Theta$ [33].

Suppose that a sample of n observations is drawn, resulting in the values: $X_1, X_2, \ldots, X_n$. To clarify the notation, we denote the mean of the sample by $\overline{X}$. In addition, the observed values can be arranged in the non-decreasing sequence, which is denoted by: $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$. In particular, we have $X_{1:n} = \min\{X_1, X_2, \ldots, X_n\}$ and $X_{n:n} = \max\{X_1, X_2, \ldots, X_n\}$.

Henceforth, $X_{k:n}$ shall be called the $k-$th order statistic [34]. In general, any measurable function $\mathcal{S} : \Omega \to \mathbb{R}^k$ of the sample shall be called a statistic and denoted by $\mathcal{S}(X_1, X_2, \ldots, X_n)$. From this definition, we see that an estimator is a kind of statistic.

We revise several types of estimators that can be applied in the exponential model.

3.1. Estimator of Moments

Assume that, for every $\theta \in \Theta$, there exists a finite moment $\mathbb{E}_\theta[X^r]$ for $r = 1, 2, \ldots, n$. In general, every moment will be a known function of the k parameters.

Definition 2. *Then, the estimators* $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)$ *of the parameters* $\theta = (\theta_1, \ldots, \theta_k)$ *determined by the method of moments are the solutions of the system:*

$$\begin{cases} \mathbb{E}_\theta[X] = \overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i, \\ \mathbb{E}_\theta[(X - \mathbb{E}_\theta[X])^j] = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^j, \quad j = 2, \ldots, k \end{cases}$$

with respect to θ.

The fact that $\hat{\theta}$ is an estimator of moments shall be denoted by $\hat{\theta} \sim EM[\theta]$.

3.2. Quantile Estimator

First, let us define the empirical distribution function (EDF), denoted by $\hat{F}_n(y)$, which is associated with the empirical measure of a sample

$$\hat{F}_n(y) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{(-\infty;y]}(X_i),$$

where the indicator function $\mathbf{1}_A : X \to \{0, 1\}$ is defined as

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A\,, \\ 0 & \text{if } x \notin A\,. \end{cases}$$

For the EDF, one can find that the p-th quantile is given by $X_{[np]+1:n}$, where $[y]$ denotes the integer part of y. Then, we adopt the following definition of the quantile estimator (denoted as $\hat{\theta} \sim QE[\theta]$).

Definition 3. *The estimator* $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)$ *of the parameters* $\theta = (\theta_1, \ldots, \theta_k)$ *determined by the quantile method are the solutions of the system*

$$F_\theta^{-1}(p_i) = X_{[np_i]+1:n} \quad \textit{where} \quad i = 1, 2, \ldots, k, \tag{5}$$

for some selected $0 < p_1 < \ldots < p_k < 1$. *The symbol* F_θ^{-1} *represents the quantile function associated with the CDF of the random variable, i.e.,* $F_\theta(x)$.

3.3. Minimum-Variance Unbiased Estimator

Let us start by defining the property of unbiasedness, assuming that $g(\theta)$ represents some function of the parameters θ.

Definition 4. $g(\hat{\theta})$ *is an unbiased estimator of* $g(\theta)$ *(denoted by* $UE[g(\theta)]$*) if and only if*

$$\mathbb{E}_\theta[g(\hat{\theta})] = g(\theta) \qquad \forall\, \theta \in \Theta.$$

In other words, an estimator is said to be unbiased when the estimator's expected value and the true value of the parameter being estimated are the same. In addition to unbiasedness, an estimator can satisfy the minimum-variance property.

Definition 5. $g(\hat{\theta})$ *is a minimum-variance unbiased estimator of* $g(\theta)$ *(denoted* $g(\hat{\theta}) \sim MVUE[\theta]$*) if and only if*

$$\forall\, \hat{U}(X) \sim UE[g(\theta)] \quad \forall\, \theta \in \Theta \quad \mathbb{E}[(g(\hat{\theta}) - g(\theta))^2] \leqslant \mathbb{E}_\theta[(\hat{U}(X) - g(\theta))^2].$$

3.4. Maximum Likelihood Estimator

The likelihood function of n random variables $X_1, X_2, \ldots, X_n$ is defined to be the joint density of the n random variables, say $f_\theta(X_1, X_2, \ldots, X_n)$, which is considered to be a function of θ. In particular, if $X_1, X_2, \ldots, X_n$ is a random sample from the density $f_\theta(x)$, then the likelihood function is $f_\theta(X_1) f_\theta(X_2) \ldots f_\theta(X_n)$. To remind ourselves that we think of the likelihood function as a function of θ (while the observed sample is fixed), we use the notation $\mathcal{L}(\theta) : \Theta \to \mathbb{R}$ for the likelihood function [31].

Definition 6. *$\hat{\theta}$ is a maximum likelihood estimator ($\hat{\theta} \sim MLE[\theta]$) of θ if and only if*

$$\forall\, x \in \Omega \qquad \mathcal{L}(\hat{\theta}(x)) = \sup_{\theta \in \Theta} \mathcal{L}(\theta).$$

Finally, let us note that, when $\hat{\theta} \sim MLE[\theta]$, then $g(\hat{\theta}) \sim MLE[g(\theta)]$.

4. Estimators for the Exponential Distribution

Having summarized general definitions of estimators, we can move on to the exponential distribution, which is the center of attention for this paper. As explained in Section 2, the exponential distribution is characterized by one parameter, λ, which is considered the estimand in our framework.

4.1. Estimator of Moments

By following the definition of the estimator of moments, Definition 2, and implementing the expected value for the exponential distribution (1), we arrive at the following equation:

$$\mathbb{E}_\lambda[X] = \frac{1}{\lambda} = \overline{X},$$

which can be solved straightforwardly as

$$\hat{\lambda}_{EM} = \frac{1}{\overline{X}}.$$

4.2. Quantile Estimator

Let us begin with finding the quantile function, which is the inverse to the CDF of the exponential distribution. We obtain

$$F_\lambda^{-1}(x) = -\frac{1}{\lambda} \ln{(1-x)}.$$

From Definition 3, we have

$$F_\lambda^{-1}(p) = X_{[np]+1:n} \equiv \hat{\xi}_p$$

for any $0 < p < 1$. One obtains

$$\hat{\lambda}_{QE}^{p} = -\frac{\ln{(1-p)}}{\hat{\xi}_p},$$

which, in a particular case of $p = 1/2$, takes the form

$$\hat{\lambda}_{QE}^{0.5} = -\frac{\ln\frac{1}{2}}{\hat{\xi}_{\frac{1}{2}}} = \frac{\ln 2}{\hat{\xi}_{\frac{1}{2}}}.$$

4.3. Minimum-Variance Unbiased Estimator

To define $MVUE[\theta]$ for the exponential model, we need to revise several definitions and theorems.

Definition 7 (Sufficient statistic.)**.** *A statistic $\mathcal{S}(X_1, X_2, \ldots, X_n)$ is a sufficient statistic for a family of distributions $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$, if, for any set of events $F \in \mathcal{F}^n$ and for any value of the statistic, the conditional distribution*

$$P_\theta((X_1, \ldots, X_n) \in F | \mathcal{S} = s)$$

does not depend on θ.

Definition 8 ([31], Complete family of densities)**.** *Let $X_1, X_2, \ldots, X_n$ denote a random sample from the density $f_\theta(x)$ with the parameter space Θ, and let $S = \mathcal{S}(X_1, X_2, \ldots, X_n)$ be a statistic. The family of densities P_θ of $\mathcal{S}$ is defined to be complete if and only if $\mathbb{E}_\theta[g(S)] = 0$ for all $\theta \in \Theta$ implies that $P_\theta[g(S) = 0] = 1$ for all $\theta \in \Theta$, where $g(S)$ is a statistic. In addition, the statistic S is said to be complete if and only if its family of densities is complete.*

Theorem 2 (Lehmann–Scheffe theorem)**.** *Let $X_1, X_2, \ldots, X_n$ be a random sample from the density $f_\theta(x)$. Suppose $S = \mathcal{S}(X_1, X_2, \ldots, X_n)$ is a complete sufficient statistic for θ. If ϕ is a statistic such that $\mathbb{E}[\phi(S)] = \theta$, then $\phi(S) \sim MVUE[\theta]$.*

Theorem 3 (Factorization theorem)**.** *A statistic $S = \mathcal{S}(X_1, X_2, \ldots, X_n)$ is sufficient if and only if the joint density factors as*

$$f_\theta(x_1, \ldots, x_n) = g_\theta(\mathcal{S}(x_1, \ldots, x_n))h(x_1, \ldots, x_n), \qquad \theta \in \Theta,$$

where the function $h(x_1, \ldots, x_n)$ is non-negative and does not involve the parameter θ, and the function $g_\theta(\mathcal{S}(x_1, \ldots, x_n))$ is non-negative and depends on $X_1, X_2, \ldots, X_n$ only through the function $\mathcal{S}(X_1, X_2, \ldots, X_n)$.

Furthermore, we utilize the concept of the exponential family, which, in general, can be extended to a vector parameter.

Definition 9. *A class of distributions is said to belong to the exponential family if the PDF can be written as*

$$f_\theta(x) = \exp\left(\sum_{i=1}^{k} c_i(\theta)T_i(x) - B(\theta)\right)h(x),$$

where $h(x), B(\theta), c_i(\theta)$, and $T_i(x)$ are known functions.

The statistics $\{c_i(\theta)\}$ can help to introduce a natural family of parameters. For a parameter vector $\theta = (\theta_1, \ldots, \theta_k) \in \Theta$, we define a set $\tilde{\Theta}$ to be the natural family of parameters:

$$\tilde{\Theta} = \{\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_k) = (c_1(\theta), \ldots, c_k(\theta)) : \theta \in \Theta\}.$$

The main application of the exponential family does not relate to finding sufficient statistics, but to showing that a sufficient statistic is complete since this concept is useful in obtaining the best estimators.

Theorem 4. *If $\mathcal{P}$ is an exponential family of densities and a natural family of parameters $\tilde{\Theta} = C(\Theta)$ contain a non-empty $k-$dimensional open interval, then the statistic $(T_1(x), \ldots, T_k(x))$ is complete.*

By following the above definitions and theorems, one can compute a sufficient and complete statistic for the exponential distribution. The joint density can be expressed as

$$f_\lambda(x_1, \ldots, x_n) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right)\mathbf{1}_{\mathbb{R}_+}(x_{1:n}) = \exp\left(-\lambda \sum_{i=1}^{n} x_i + n \ln \lambda\right)\mathbf{1}_{\mathbb{R}_+}(x_{1:n}).$$

Based on Theorem 3, we know that the function

$$T = \mathcal{T}(X_1, \ldots, X_n) = \sum_{i=1}^{n} X_i$$

is a sufficient statistic. Furthermore, since $-\lambda \in (-\infty, 0)$, then, from Theorem 4, we have that S is a complete statistic. In addition, we know that the sum of n random variables such that $X \sim \text{Exp}(\lambda)$ has the distribution $\Gamma(n, \lambda)$, where n is the shape parameters and λ represents the rate parameter. The PDF for the distribution $\Gamma(n, \lambda)$ is given by

$$f(x) = \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x} \mathbf{1}_{[0,\infty)}(x),$$

where $\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} \mathrm{d}x$. Now, we proceed to computing the expected value of $\frac{n-1}{T}$:

$$\begin{aligned} \mathbb{E}_\lambda\left[\frac{n-1}{T}\right] &= (n-1) \int_0^\infty \frac{1}{x} \frac{\lambda^n}{\Gamma(n)} x^{n-1} e^{-\lambda x} \mathrm{d}x \\ &= (n-1) \frac{\lambda}{n-1} \int_0^\infty \frac{\lambda^{n-1}}{\Gamma(n-1)} x^{n-2} e^{-\lambda x} \mathrm{d}x = \lambda, \end{aligned}$$

where we use the property of the Γ distribution

$$\Gamma(n) = (n-1)\Gamma(n-1)$$

as well as the value of the integral

$$\int_0^\infty \frac{\lambda^{n-1}}{\Gamma(n-1)} x^{n-2} e^{-\lambda x} \mathrm{d}x = 1,$$

which comes straightforwardly from integrating the PDF of $\Gamma(n-1, \lambda)$. As a result, we see that $\hat{\lambda}_{MVUE} = \frac{n-1}{T}$ is an unbiased estimator of the parameter λ. In addition, by following the Lehmann–Scheffe theorem, we notice that this estimator

$$\hat{\lambda}_{MVUE} = \frac{n-1}{\sum_{i=1}^{n} X_i} = \frac{n-1}{n\overline{X}} \tag{6}$$

is a minimum-variance estimator since it is a function of a sufficient and complete statistic, i.e., $\hat{\lambda}_{MVUE} \sim MVUE[\lambda]$.

For applications of the estimation theory, the class of MVUE is especially important due to the fact that it minimizes the statistical dispersion. The spread of the results is a key factor that limits the accuracy of the analysis. However, the estimator (6) guarantees the lowest achievable scatter. To quantify this aspect, let us compute

$$\begin{aligned} \text{Var}[\hat{\lambda}_{MVUE}] = \text{Var}\left[\frac{n-1}{\sum_{i=1}^{n} X_i}\right] &= \mathbb{E}\left[\left(\frac{n-1}{\sum_{i=1}^{n} X_i}\right)^2\right] - \left(\mathbb{E}\left[\frac{n-1}{\sum_{i=1}^{n} X_i}\right]\right)^2 \\ &= \mathbb{E}\left[\left(\frac{n-1}{\sum_{i=1}^{n} X_i}\right)^2\right] - \lambda^2, \end{aligned}$$

where the first element can be calculated based on the properties of the Γ distribution

$$\begin{aligned}\mathbb{E}\left[\left(\frac{n-1}{\sum_{i=1}^{n} X_i}\right)^2\right] &= (n-1)^2\,\mathbb{E}\left[\left(\frac{1}{\sum_{i=1}^{n} X_i}\right)^2\right] = (n-1)^2\int_0^\infty \frac{1}{x^2}\frac{\lambda^n}{\Gamma(n)}x^{n-1}e^{-\lambda x}\mathrm{d}x\\ &= (n-1)^2\frac{\lambda^2}{(n-1)(n-2)}\int_0^\infty \frac{\lambda^{n-2}}{\Gamma(n-2)}x^{n-3}e^{-\lambda x}\mathrm{d}x\\ &= \frac{n-1}{n-2}\lambda^2,\end{aligned}$$

which finally leads to

$$\mathrm{Var}[\hat{\lambda}_{MVUE}] = \frac{n-1}{n-2}\lambda^2 - \lambda^2 = \frac{\lambda^2}{n-2}. \tag{7}$$

4.4. Maximum Likelihood Estimator

The likelihood function for the exponential distribution takes the form

$$\mathcal{L}(\lambda) = f(\lambda; X_1, \ldots, X_n) = \lambda e^{-\lambda X_1} \ldots \lambda e^{-\lambda X_n} = \lambda^n e^{-\lambda \sum_{i=1}^{n} X_i}.$$

Since the natural logarithm is an increasing function, we can maximize either $\mathcal{L}(\theta)$ or $\ln\mathcal{L}(\theta)$. For convenience, we compute

$$\ln\mathcal{L}(\lambda) = n\ln\lambda - \lambda\sum_{i=1}^{n} X_i.$$

Then, we obtain

$$\frac{\partial \ln\mathcal{L}(\lambda)}{\partial\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} X_i.$$

One can easily solve $\frac{\partial \ln\mathcal{L}(\lambda)}{\partial\lambda} = 0$ by finding

$$\hat{\lambda}_{MLE} = \frac{n}{\sum_{i=1}^{n} X_i} = \frac{1}{\overline{X}}.$$

Since $\frac{\partial^2 \ln\mathcal{L}(\lambda)}{\partial\lambda^2} = -\frac{n}{\lambda^2} < 0$, we know that $\hat{\lambda}_{MLE}$ corresponds to the maximum value of the likelihood function. One can note that the estimator $\hat{\lambda}_{MLE}$ is equivalent to $\hat{\lambda}_{EM}$.

5. Estimators for the Type II Censored Data

In this part, we propose estimators of the exponential distribution for the type II censored data. Let $X_1, \ldots, X_n$ be a random sample from the same distribution. We say that we perform type II censoring of the data if we observe exactly r first results (naturally, $r < n$), assuming that the observations are arranged in non-decreasing order. In other words, type II censored data consist of the r consecutive order statistics: $X_{1:n}, \ldots, X_{r:n}$ [35–38].

The focus on censored data makes the estimation theory more practicable. In real experiments, we have to deal with uncertainties and deficiencies that limit our capability to measure a sufficient set of data. Therefore, it appears relevant to search for estimators suited for censored data, which can be applied in realistic circumstances.

Before determining the exact forms of estimators, we need to revise the properties of the order statistics.

Theorem 5. *Suppose $X_1, \ldots, X_n$ is a sample taken from a distribution characterized by the PDF denoted by f_λ and CDF F_λ. Then, the k-th order statistic, $X_{k:n}$, has its own PDF and CDF in the forms [39]:*

$$f_{k:n}(x) = n\binom{n-1}{k-1} f_\lambda(x) F_\lambda^{k-1}(x)(1 - F_\lambda(x))^{n-k},$$
$$F_{k:n}(x) = P(X_{k:n} \leqslant x) = \sum_{i=k}^{n} \binom{n}{i} F_\lambda^i(x)(1 - F_\lambda(x))^{n-i}.$$

Furthermore, we can also find a modified formula for the joint density that fits the censoring.

Theorem 6. *The joint density for the sequence of $r < n$ order statistics $X_{1:n}, \ldots, X_{r:n}$ is given by*

$$f_{1,\ldots,r}(x_{1:n}, \ldots, x_{r:n}) = \frac{n!}{(n-r)!} f(x_{1:n}) \ldots f(x_{r:n})(1 - F(x_{r:n}))^{n-r},$$

for $-\infty < x_{1:n} < \ldots < x_{r:n} < +\infty$.

In particular, Theorems 5 and 6 can be applied to the exponential distribution (see Definition 1) to determine the PDF and CDF for the k-th order statistic, as well as the joint density for r consecutive order statistics. One obtains the following:

$$f_{k:n}(x) = n\binom{n-1}{k-1} \lambda e^{-\lambda x(n+1-k)} (1 - e^{-\lambda x})^{k-1} \mathbf{1}_{[0,\infty)}(x),$$
$$F_{k:n}(x) = \sum_{i=k}^{n} \binom{n}{i} (1 - e^{-\lambda x})^i e^{-\lambda x(n-i)} \mathbf{1}_{[0,\infty)}(x),$$
$$f_{1,\ldots,r}(x_{1:n}, \ldots, x_{r:n}) = \frac{n!}{(n-r)!} \lambda^r \exp\left[-\lambda\left(\sum_{i=1}^{r} x_{i:n} + (n-r)x_{r:n}\right)\right]. \quad (8)$$

The properties of the exponential distribution enumerated in (8) are implemented to determine the explicit forms of three kinds of estimators suited for type II censored data: MVUE, MLE, and best linear unbiased estimator (BLUE).

5.1. Minimum-Variance Unbiased Estimator for Censored Data

The results presented in (8) along with the factorization theorem allow us to state that

$$T_r = \sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n} \quad (9)$$

is a sufficient statistic. Furthermore, we know that $-\lambda \in (-\infty, 0)$, which implies on the basis of Theorem 4 that T_r is a complete statistic. We formulate and prove the theorem.

Theorem 7. *For type II censored data, $X_{1:n}, \ldots, X_{r:n}$, selected from an exponential distribution, the rate parameter, λ, can be estimated with the MVUE given by*

$$\hat{\lambda}_{MVUE}(r) = \frac{r-1}{T_r} \quad (10)$$

Proof. Let us start by studying the distribution of the statistic T_r.

Suppose $Y_1, \ldots, Y_n$ are selected from the exponential distribution characterized by $\lambda = 1$. Then, one can prove that these figures are related to the order statistics by [34]

$$\begin{cases} Y_1 &= nX_{1:n}, \\ Y_2 &= (n-1)(X_{2:n} - X_{1:n}), \\ \vdots \\ Y_n &= X_{n:n} - X_{n-1:n} \end{cases} \Leftrightarrow \begin{cases} X_{1:n} = \frac{Y_1}{n} \\ X_{2:n} = \frac{Y_1}{n} + \frac{Y_2}{n-1} \\ \vdots \\ X_{n:n} = \frac{Y_1}{n} + \frac{Y_2}{n-1} + \ldots + \frac{Y_n}{1}. \end{cases}$$

We see that $X_{i:n} = \sum_{k=1}^{i} \frac{Y_k}{n-k+1}$ for $i = 1, \ldots, n$. Thus,

$$\begin{aligned} T_r &= \sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n} \\ &= \frac{Y_1}{n} + \frac{Y_1}{n} + \frac{Y_2}{n-1} + \ldots + \frac{Y_1}{n} + \frac{Y_2}{n-1} + \ldots + \frac{Y_r}{n-r+1} \\ &+ (n-r)\left(\frac{Y_1}{n} + \frac{Y_2}{n-1} + \ldots + \frac{Y_r}{n-r+1}\right) \\ &= Y_1\left(\frac{r}{n} + 1 - \frac{r}{n}\right) + Y_2\left(\frac{r-1}{n-1} + \frac{n}{n-1} - \frac{r}{n-1}\right) + \ldots \\ &+ Y_r\left(\frac{1}{n-r+1} + \frac{n}{n-r+1} - \frac{r}{n-r+1}\right) \\ &= Y_1 + Y_2 + \ldots + Y_r. \end{aligned}$$

Again, we use the fact that $\sum_{i=1}^{r} Y_i$, which is a sum of variables from the exponential distribution, has the distribution $\Gamma(r, n)$. Then, it follows that

$$\begin{aligned} E\left[\frac{r-1}{T_r}\right] &= (r-1)E\left(\frac{1}{T_r}\right) = (r-1)\int_0^{+\infty} \frac{1}{t}\frac{\lambda^r}{\Gamma(r)} t^{r-1} e^{-\lambda t} \mathrm{d}t \\ &= (r-1)\frac{\lambda}{r-1}\int_0^{+\infty} \frac{\lambda^{r-1}}{\Gamma(r-1)} t^{r-2} e^{-\lambda t} \mathrm{d}t = \lambda. \end{aligned}$$

The last integral gives 1 since it is computed over the entire distribution $\Gamma(r-1, \lambda)$. Finally, based on the Theorem 2, we can conclude that

$$\hat{\lambda}_{MVUE}(r) = \frac{r-1}{T_r} = \frac{r-1}{\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}}$$

satisfies the conditions for MVUE, which finishes the proof. □

In addition, one can notice that, for $r = n$, we have $\hat{\lambda}_{MVUE}(r) = \hat{\lambda}_{MVUE}$, which means that, without the censoring, our estimator agrees with the formula for the complete set of data; see (6).

One can also compute the variance of $\hat{\lambda}_{MVUE}(r)$ and obtain

$$\mathrm{Var}\left[\hat{\lambda}_{MVUE}(r)\right] = \frac{\lambda^2}{r-2}. \tag{11}$$

The proof of (11) shall be skipped since one can use the same reasoning as in the case of obtaining $\mathrm{Var}\left[\hat{\lambda}_{MVUE}\right]$; see (7). Here, T_r has the distribution $\Gamma(r, \lambda)$, which allows one to proceed analogously as in Section 4.3.

The result (11) informs us about the impact of the censoring on the reliability of estimation. Naturally, as we reduce the number of observations included in the model, we expect to increase the uncertainty of estimation. This feature is reflected in (11), which shows that, the less data we include, the more variance we obtain.

5.2. Maximum Likelihood Estimator for Censored Data

MLE is also obtainable with a set of censored data $X_{1:n}, \ldots, X_{r:n}$. We propose and prove a theorem.

Theorem 8. *Let $X_1, \ldots, X_n$ denote a sample of variables with the exponential distribution characterized by the rate parameter λ. The parameter λ can be estimated based on the first r order statistics, $X_{1:n}, \ldots, X_{r:n}$, from the following MLE:*

$$\hat{\lambda}_{MLE}(r) = \frac{r}{T_r}, \tag{12}$$

where the statistic T_r was defined in (9).

Proof. Since we know the statistics of r consecutive order statistics from the exponential distribution, we can formulate the maximum likelihood function for type II censored set of data

$$\begin{aligned} \mathcal{L}(\lambda) &= \frac{n!}{(n-r)!} \prod_{i=1}^{r} f(X_{i:n})(1 - F(X_{r:n}))^{n-r} \\ &= \frac{n!}{(n-r)!} \left(\prod_{i=1}^{r} \lambda e^{-\lambda X_{i:n}} \right) \left(e^{X_{r:n}} \right)^{n-r} \\ &= \frac{n!}{(n-r)!} \lambda^r \exp\left[-\lambda \left(\sum_{i=1}^{r} X_{i:n} + (n-r) X_{r:n} \right) \right]. \end{aligned}$$

In this approach, we search for the maximum value of $\mathcal{L}(\lambda)$. To achieve this, we compute the natural logarithm of $\mathcal{L}(\lambda)$ along with its first derivative:

$$\begin{aligned} l(\lambda) \equiv \ln \mathcal{L}(\lambda) &= \ln \frac{n!}{(n-r)!} + r \ln \lambda - \lambda \left(\sum_{i=1}^{r} X_{i:n} + (n-r) X_{r:n} \right), \\ \frac{\partial l(\lambda)}{\partial \lambda} &= \frac{r}{\lambda} - \left(\sum_{i=1}^{r} X_{i:n} + (n-r) X_{r:n} \right). \end{aligned}$$

Then, by solving $\frac{\partial l(\lambda)}{\partial \lambda} = 0$, we obtain the estimator:

$$\hat{\lambda}_{MLE}(r) = \frac{r}{\sum_{i=1}^{r} X_{i:n} + (n-r) X_{r:n}} = \frac{r}{T_r}. \tag{13}$$

Finally, we check that the second derivative is negative:

$$\frac{\partial^2 l(\lambda)}{\partial \lambda^2} = -\frac{r}{\lambda^2} < 0$$

for all λ, which implies that we maximize the likelihood function and that $\hat{\lambda}_{MLE}(r)$ is a legitimate estimator. □

Finally, let us notice that, for $r = n$, one obtains $\hat{\lambda}_{MLE}(r) = \hat{\lambda}_{MLE} \equiv \frac{1}{\bar{X}}$. This observation comes as no surprise since, without the censoring, both approaches are expected to deliver the same estimator.

5.3. Best Linear Unbiased Estimator

For order statistics, one can define linear estimators, i.e., linear combinations of the order statistics $\hat{V} = \sum_{i=1}^{n} a_i X_{i:n}$, where $a_i \in \mathbb{R}$ are some constant coefficients. Then, one can introduce the notion of best linear unbiased estimator (BLUE) [40].

Definition 10. *An unbiased linear estimator $\hat{V} = \sum_{i=1}^{n} a_i X_{i:n}$ of the parameter function $g(\theta)$ is called best linear unbiased estimator (BLUE) if and only if*

$$\forall \sum_{i=1}^{n} c_i X_{i:n} \sim \mathrm{UE}[g(\theta)] \;\; \forall \theta \in \Theta \;\; \mathbb{E}_\theta \left[\left(\sum_{i=1}^{n} a_i X_{i:n} - g(\theta) \right)^2 \right] \leqslant \mathbb{E}_\theta \left[\left(\sum_{i=1}^{n} c_i X_{i:n} - g(\theta) \right)^2 \right].$$

The goal of this section is to find a minimum-variance BLUE that is based on first r order statistics ($1 < r \leqslant n$), $X_{1:n}, \ldots, X_{r:n}$, to estimate the scale parameter $\beta \equiv \frac{1}{\lambda}$ of the exponential distribution. We formulate and prove a theorem.

Theorem 9. *Let $X_1, \ldots, X_n$ denote a sample of variables with the exponential distribution characterized by the rate parameter λ. The scale parameter β can be estimated based on the first r order statistics, $X_{1:n}, \ldots, X_{r:n}$, from the following BLUE:*

$$\hat{\beta}_{BLUE}(r) = \sum_{i=1}^{r} c_i X_{i:n} = \frac{1}{r} \left(\sum_{i=1}^{r} X_{i:n} + (n-r) X_{r:n} \right).$$

Proof. The estimator has the form:

$$\hat{\beta}_{BLUE}(r) = \sum_{i=1}^{r} c_i X_{i:n}, \tag{14}$$

where the coefficients c_i have to be properly determined to ensure that $\lambda_{BLUE}(r)$ is unbiased and features minimal variance. Let us notice that the random variables

$$\begin{aligned} Z_1 &= n X_{1:n} \\ Z_i &= (n-i+1)(X_{i:n} - X_{i-1:n}), \qquad i = 2, \ldots, r, \end{aligned}$$

are independent and have the exponential distribution with the rate parameter λ. One can use these variables to determine the coefficients c_i. First, let us note that the estimator (14) can be rewritten as

$$\begin{aligned} \hat{\beta}_{BLUE}(r) &= \sum_{i=1}^{r} c_i X_{i:n} \\ &= \sum_{i=1}^{r} c_i X_{1:n} - \sum_{i=1}^{r} (X_{i:n} - X_{i+1:n}) \sum_{j=i+1}^{r} c_j \\ &= X_{1:n} \sum_{i=1}^{r} c_i + \sum_{i=1}^{r} (X_{i+1:n} - X_{i:n}) \sum_{j=i+1}^{r} c_j. \end{aligned}$$

Then, if we denote

$$d_i \equiv \frac{1}{n-i+1} \sum_{j=i}^{r} c_j, \;\; i = 1, \ldots, r, \tag{15}$$

we obtain

$$\hat{\beta}_{BLUE}(r) = \sum_{i=1}^{r} c_i X_{i:n} = \sum_{i=1}^{r} d_i Z_{i:n}.$$

Next, we move on to computing the expected value of the estimator $\hat{\beta}_{BLUE}(r)$:

$$\mathbb{E}[\hat{\beta}_{BLUE}(r)] = \mathbb{E}\left[\sum_{i=1}^{r} d_i Z_{i:n}\right] = \sum_{i=1}^{r} d_i \mathbb{E}[Z_{i:n}] = \beta \sum_{i=1}^{r} d_i.$$

Since $\hat{\beta}_{BLUE}(r)$ needs to be unbiased, we obtain a necessary condition for BLUE:

$$\sum_{i=1}^{r} d_i = 1. \tag{16}$$

To ensure the minimal variance, we consider

$$\min_{d_1,\ldots,d_r} \mathbb{E}\left[\left(\sum_{i=1}^{r} d_i Z_{i:n} - \beta\right)^2\right], \tag{17}$$

which can be elaborated on by taking advantage of the necessary condition (16):

$$\begin{aligned}
\mathbb{E}\left[\left(\sum_{i=1}^{r} d_i Z_{i:n} - \beta\right)^2\right] &= \mathbb{E}\left[\left(\sum_{i=1}^{r} d_i (Z_{i:n} - \beta)\right)^2\right] \\
&= \sum_{i=1}^{r} d_i^2 \operatorname{Var} Z_i + \sum_{i\neq j}^{r} d_i d_j \operatorname{Cov} Z_i Z_j \\
&= \beta^2 \sum_{i=1}^{r} d_i^2 \\
&= \beta^2 \left[\sum_{i=1}^{r-1} d_i^2 + \left(1 - \sum_{i=1}^{r-1} d_i\right)^2\right],
\end{aligned}$$

where we used the fact that $d_r = 1 - \sum_{i=1}^{r-1} d_i$.

As a result, in order to determine the minimal value of (17), we need to consider the following:

$$\min_{d_1,\ldots,d_r} \left[\sum_{i=1}^{r-1} d_i^2 + \left(1 - \sum_{i=1}^{r-1} d_i\right)^2\right]. \tag{18}$$

By differentiating (18) with respect to d_i for $i = 1, \ldots, r-1$, i we obtain the conditions for optimization

$$\frac{\partial}{\partial d_i}\left[\sum_{j=1}^{r-1} d_j^2 + \left(1 - \sum_{j=1}^{r-1} d_j\right)^2\right] = 2d_i - 2\left(1 - \sum_{j=1}^{r-1} d_j\right) = 2\left(d_i - 1 - \sum_{j=1}^{r-1} d_j\right) = 0,$$

which leads to a system of equations:

$$d_i = 1 - \sum_{j=1}^{r-1} d_j, \qquad j = 1, \ldots, r-1.$$

We notice that all coefficients d_i must be equal for $i = 1, \ldots, r-1$. Let us, then, use one symbol $d \equiv d_i$, which gives

$$d = 1 - \sum_{i=1}^{r-1} d = 1 - (r-1)d$$

and, finally, we arrive at the result

$$d = d_i = \frac{1}{r}, \qquad i = 1, \ldots, r-1.$$

From the necessary condition (16), we obtain

$$d_r = 1 - \sum_{i=1}^{r-1} = \frac{1}{r}.$$

Since, for the BLUE, we need c_i, we calculate those from (15):

$$\frac{1}{n-i+1} \sum_{j=i}^{r} c_j = \frac{1}{r}$$

$$\sum_{j=i}^{r} c_j = \frac{n-i+1}{r}, \qquad i = 1, \ldots, r,$$

which leads to

$$c_r = \frac{n-r+1}{r},$$

$$c_i = \sum_{j=i}^{r} c_j - \sum_{j=i+1}^{r} c_j = \frac{n-i+1}{r} - \frac{n-(i+1)+1}{r} = \frac{1}{r} \quad \text{for} \quad i = 1, \ldots, r-1.$$

If we substitute the results into (14), we obtain the explicit formula for the BLUE:

$$\hat{\beta}_{BLUE}(r) = \sum_{i=1}^{r} c_i X_{i:n} = \frac{1}{r}\left(\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right),$$

which finishes the proof. □

6. Performance Analysis of the Estimators for Censored Data

In this section, we propose a numerical framework to evaluate the efficiency of the estimators that operate on censored data. In particular, we compare the performance of $\hat{\lambda}_{MVUE}(r)$ and $\hat{\lambda}_{MLE}(r)$ for different values of r.

First, let us consider a typical fiber with the attenuation coefficient $\alpha = 0.5$ dB/km. We know that the photon loss is characterized by an exponential distribution with the rate parameter $\lambda = \ln 10\, \alpha$. Thus, we can simulate an experimental scenario by randomly generating a sample of 10^7 observations such that $X \sim \text{Exp}(\lambda)$. Wolfram Mathematica 10.0 was used to generate random numbers and compute the estimators with order statistics. Each piece of data corresponds to the distance that a photon has covered before absorption (or scattering). To quantify the performance of the statistical estimators, we did not focus on specific results of estimation, but we computed the relative error

$$\delta_x := \frac{|\lambda - \hat{\lambda}_x|}{\lambda}\, 100\,\%, \tag{19}$$

where $\hat{\lambda}_x$ denotes any of the estimators considered. For convenience, we expressed the estimation error as a percentage of the actual rate parameter. This approach allowed us to investigate how the quality of estimation changes when we reduce the initial set of observations by leaving only the first r order statistics. The percent errors that depend on r are denoted by $\delta_{MVUE}(r)$ and $\delta_{MLE}(r)$ for the respective estimators $\hat{\lambda}_{MVUE}(r)$ and $\hat{\lambda}_{MLE}(r)$. The results of the numerical simulation are collected in Table 1.

Table 1. The percent errors of the rate parameter estimation. The results are rounded to two significant figures. The estimators $\hat{\lambda}_{MVUE}(r)$ (10) and $\hat{\lambda}_{MLE}(r)$ (12) were utilized.

r	$\delta_{MVUE}(r)$	$\delta_{MLE}(r)$
10^7	0.024%	0.024%
10^6	0.037%	0.036%
10^5	0.050%	0.051%
10^4	0.53%	0.52%
10^3	1.3%	1.2%
10^2	3.4%	4.5%
10	15%	5.9%

In the first place, one can notice that the first row with $r = 10^7$ corresponds to an estimation with the complete set of data. It was already explained that, in this case, $\hat{\lambda}_{MVUE}(r)$ and $\hat{\lambda}_{MLE}(r)$ are equivalent to standard estimators for the exponential model. We see that $r = 10^7$ observations are sufficient for a reliable estimation of the rate parameter and, consequently, determining the attenuation coefficient α. If we decrease r, which means that we incorporate only a portion containing r consecutive order statistics, we observe that the estimation error increases monotonically for both estimators. This tendency was anticipated since, by performing censoring, we reject a part of the highest-valued observation. This implies that a censored set of data is not representative for the distribution, which leads to a more significant error in parameter estimation. We see, however, that, up to $r = 10^5$, the growth of the error is not critical. By further censoring the data, we obtain results that are more distorted due to the incomplete representation of the distribution. Finally, if $r = 10$, which means that we implement only the lowest 10 observation, we obtain unreliable results.

In addition, we notice that the performance of $\hat{\lambda}_{MVUE}(r)$ for $r = 10$ and $r = 100$ differs considerably from $\hat{\lambda}_{MLE}(r)$, whereas, for the other values of r, both estimators deliver similar results. This discrepancy for a low number of statistics results from different figures in the numerators of $\hat{\lambda}_{MVUE}(r)$ and $\hat{\lambda}_{MLE}(r)$, which leads to a significant difference when only the lowest order statistics are included in the estimation. In conclusion, the feasibility study presented in this section allows one to verify how many observations are needed to guarantee reliable parameter estimation.

7. Discussion

In the paper, we considered applications of the estimation theory to the exponential distribution. The problem is formulated on the assumption that a measurable quantity is represented by a random variable X with a PDF $f_\theta(x)$ that depends on an unknown parameter θ. Then, on the basis of the observed sample values $X_1, \ldots, X_n$, it is desired to estimate the value of the unknown parameter θ or the value of some function of the parameter, $g(\theta)$. We followed the approach called the parametric point estimation with the goal of determining some statistic, say $\mathcal{S}(X_1, \ldots, X_n)$, to represent the unknown parameter.

The paper contains a brief revision of general concepts of estimation, included in Section 3. Then, in Section 4, we summarized the most common estimators for the exponential distribution, where the goal was to determine the rate parameter λ. We presented four types of estimators that operate on the complete set of data, i.e., $X_1, \ldots, X_n$.

In physics, the exponential distribution plays a central role in decay processes. In particular, the Beer–Lambert law (3) describes the attenuation that involves a decline in the intensity of light as the beam travels through a fiber. This macroscopic law has been implemented to develop a quantum picture for the photon loss. First of all, we have postulated that the distance that a photon covers before absorption (or scattering) has ah exponential distribution, i.e., $L \sim \text{Exp}(\lambda)$. In this context, the rate parameter λ is connected to the attenuation coefficient α that describes the properties of the optical link. Therefore, the estimators presented in Section 4 can be applied to characterize fibers by determining their attenuation coefficients. In addition, the estimation theory allows us to establish minimum-variance estimators that are particularly well-suited for physical applications. The variance, which is an indicator of statistical dispersion, conveys information about the reliability of the measurement. A higher variance means that the scheme provides different results under unchanged conditions. However, we are interested in consistent measurement procedures that feature only insignificant scatter within the delivered results. For this reason, the class of MVUE, which also guarantees validity through unbiasedness, appears very promising for further applications.

The key part of the paper relates to the estimators corresponding to type II censored data, which were introduced in Section 5. In such a case, operating with an incomplete set of measurements is assumed, i.e., we possess r consecutive order statistics: $X_{1:n}, \ldots, X_{r:n}$. This problem is intrinsically linked to the nature of physical measurement. For the exponential distribution, the random variable can be observed within the interval: $X \in [0; \infty)$. However, if X denotes a physical observable (e.g., fiber length or time), it cannot be tracked up to infinity. Consequently, the set of physically obtainable data is always truncated according to our inability to capture extremely large values of the random variable. This inherent limitation of measurement becomes a motivation to search for statistical estimators that can produce reliable results with realistic sets of data.

In this paper, three estimators for type II censored data have been presented. Two of them, i.e., MVUE and MLE, allow one to estimate the rate parameter (λ) of the exponential distribution, whereas the third (BLUE) provides a formula for the scale parameter (λ^{-1}). Any of the estimators can be implemented with a realistic set of data. Special attention, however, has been paid to MVUE, where an explicit formula for the variance can be given. Thanks to this, we can investigate how the reliability of the estimation declines as we reduce the set of data.

Numerical simulations have been performed to test the efficiency of the estimators (MVUE and MLE) with censored data. We assumed a specific value of the attenuation coefficient, which allowed us to generate a set of observations. To quantify the performance of the estimators, we computed the percent errors for different numbers of order statistics included in the estimation. The results demonstrate how the quality of estimation degenerates as we truncate the initial set of data. By following the method presented in this paper, one can conduct similar simulations with another attenuation coefficient or other numbers of statistics.

8. Conclusions

This paper provides statistical tools for analyzing the photon loss in fiber-optic communication. The attenuation coefficient can be determined based on the estimators presented in this work, including the methods intended for censored (truncated) sets of data. The results pave the way for the analytical treatment of the attenuation process. In the future, other types of distributions relevant to physics can be investigated to determine reliable estimators of the parameters.

Author Contributions: Conceptualization, A.C.; formal analysis, A.C. and K.C; investigation, A.C. and K.C.; methodology, A.C. and K.C.; project administration, A.C.; supervision, A.C.; validation, K.C.; writing—original draft, A.C.; writing—review and editing, A.C. and K.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data and code underlying the results presented in this paper can be obtained from the corresponding author upon a reasonable request.

Acknowledgments: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BLUE	best linear unbiased estimator
CCDF	complementary cumulative distribution function
CDF	cumulative distribution function
EDF	empirical distribution function
EM	estimator of moments
MGF	moment generating function
MLE	maximum likelihood estimator
MVUE	minimum-variance unbiased estimator
PDF	probability density function
QE	quantile estimator
UE	unbiased estimator

References

1. Okoshi, T. *Optical Fibers*; Academic Press Inc.: London, UK, 1982.
2. Ishio, H.; Minowa, J.; Nosu, K. Review and status of wavelength-division-multiplexing technology and its application. *J. Light. Technol.* **1984**, *2*, 448–463. [CrossRef]
3. Cordeiro, C.M.B.; Franco, M.A.R.; Chesini, G.; Barretto, E.C.S.; Lwin, R.; Brito Cruz, C.H.; Large, M.C.J. Microstructured-core optical fibre for evanescent sensing applications. *Opt. Express* **2006**, *14*, 13056. [CrossRef] [PubMed]
4. Maidi, A.M.; Yakasai, I.; Abas, P.E.; Nauman, M.M.; Apong, R.A.; Kaijage, S.; Begum, F. Design and Simulation of Photonic Crystal Fiber for Liquid Sensing. *Photonics* **2021**, *8*, 16. [CrossRef]
5. Matsuura, M. Recent Advancement in Power-over-Fiber Technologies. *Photonics* **2021**, *8*, 335. [CrossRef]
6. Oh, S.J.; Chun, W.; Riffat, S.B.; Jeon, Y.I.; Dutton, S.; Han, H.J. Computational analysis on the enhancement of daylight penetration into dimly lit spaces: Light tube vs. fiber optic dish concentrator. *Build. Environ.* **2013**, *59*, 261–274. [CrossRef]
7. Caravaca-Aguirre, A.M.; Piestun, R. Single multimode fiber endoscope. *Opt. Express* **2017**, *25*, 1656–1665. [CrossRef] [PubMed]
8. Krenn, M.; Malik, M.; Scheidl, T.; Ursin, R.; Zeilinger, A. Quantum communication with photons. In *Optics in Our Time*; Al-Amri, M.D., El-Gomati, M.M., Zubairym M.S., Eds.; Springer: Cham, Switzerland, 2016; pp. 455–482.
9. Pirandola, S.; Andersen, U.L.; Banchi, L.; Berta, M.; Bunandar, D.; Colbeck, R.; Englund, D.; Gehring, T.; Lupo, C.; Ottaviani, C.; et al. Advances in quantum cryptography. *Adv. Opt. Photon.* **2020**, *12*, 1012–1236. [CrossRef]
10. Bennett, C.H.; Brassard, G. Quantum Cryptography: Public key distribution and coin tossing. In Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India, 9–12 December 1984; pp. 175–179.
11. Ekert, A. K. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.* **1991**, *67*, 661–663. [CrossRef]
12. Boaron, A.; Boso, G.; Rusca, D.; Vulliez, C.; Autebert, C.; Caloz, M.; Perrenoud, M.; Gras, G.; Bussieres, F.; Li, M.-J.; et al. Secure quantum key distribution over 421 km of optical fiber. *Phys. Rev. Lett.* **2018**, *121*, 190502. [CrossRef]
13. Wei, J.L.; Sanchez, C.; Giddings, R.P; Hugues-Salas, E.; Tang, J.M. Significant improvements in optical power budgets of real-time optical OFDM PON systems. *Opt. Express* **2010**, *18*, 20732–20745. [CrossRef] [PubMed]
14. Grassl, M.; Beth, T.; Pellizzari, T. Codes for the quantum erasure channel. *Phys. Rev. A* **1997**, *56*, 33–38. [CrossRef]
15. Cory, D.G.; Price, M.D.; Maas, W.; Knill, E.; Laflamme, R.; Zurek, W.H.; Havel, T.F.; Somaroo, S.S. Experimental Quantum Error Correction. *Phys. Rev. Lett.* **1998**, *81*, 2152–2155. [CrossRef]
16. Lidar, D.A.; Brun, T.A. (Eds.) *Quantum Error Correction*; Cambridge University Press: New York, NY, USA, 2013.
17. Terhal, B.M. Quantum error correction for quantum memories. *Rev. Mod. Phys.* **2015**, *87*, 307–346. [CrossRef]
18. Chuang, I.L.; Leung, D.W.; Yamamoto, Y. Bosonic quantum codes for amplitude damping. *Phys. Rev. A* **1997**, *56*, 1114–1125. [CrossRef]
19. Gingrich, R.M.; Kok, P.; Lee, H.; Vatan, F.; Dowling, J.P. All Linear Optical Quantum Memory Based on Quantum Error Correction. *Phys. Rev. Lett.* **2003**, *91*, 217901. [CrossRef] [PubMed]
20. Wasilewski, W.; Banaszek, K. Protecting an optical qubit against photon loss. *Phys. Rev. A* **2007**, *75*, 042316. [CrossRef]
21. Fröhlich, B.; Dynes, J.; Lucamarini, M.; Sharpe, A.W.; Tam, S.W.-B.; Yuan, Z.; Shields, A.J. Quantum secured gigabit optical access networks. *Sci. Rep.* **2016**, *5*, 18121. [CrossRef] [PubMed]
22. Valivarthi, R.; Umesh, P.; John, C.; Owen, K.A.; Verma, V.B.; Nam, S.W.; Oblak, D.; Zhou, Q.; Tittel, W. Measurement-device-independent quantum key distribution coexisting with classical communication. *Quant. Sci. Technol.* **2019**, *4*, 045002. [CrossRef]
23. Li, Y.; Wang, N.; Wang, X.; Bai, Z. Influence of guided acoustic wave Brillouin scattering on excess noise in fiber-based continuous variable quantum key distribution. *J. Opt. Soc. Am. B* **2014**, *31*, 2379–2383. [CrossRef]
24. Chen, J.-P.; Zhang, C.; Liu, Y.; Jiang, C.; Zhang, W.; Hu, X.-L.; Guan, J.-Y.; Yu, Z.-W.; Xu, H.; Lin, J.; et al. Sending-or-not-sending with independent lasers: Secure twin-field quantum key distribution over 509 km. *Phys. Rev. Lett.* **2020**, *124*, 070501. [CrossRef] [PubMed]
25. James, D.F.V.; Kwiat, P.G.; Munro, W.J.; White, A.G. Measurement of qubits. *Phys. Rev. A* **2001**, *64*, 052312. [CrossRef]
26. Altepeter, J.B.; Jeffrey, E.R.; Kwiat, P.G. Photonic State Tomography. *Adv. At. Mol. Opt. Phys.* **2005**, *52*, 105–159.

27. Acharya, A.; Kypraios, T.; Guta, M. A comparative study of estimation methods in quantum tomography. *J. Phys. A Math. Theor.* **2019**, *52*, 234001. [CrossRef]
28. Paris, M.G.A. Quantum estimation for quantum technology. *Int. J. Quantum Inf.* **2009**, *7*, 125–137. [CrossRef]
29. Albarelli, F.; Demkowicz-Dobrzański, R. Probe Incompatibility in Multiparameter Noisy Quantum Metrology. *Phys. Rev. X* **2022**, *12*, 011039. [CrossRef]
30. Ross, S.M. *Introduction to Probability Models*, 9th ed.; Elsevier: London, UK, 2007; pp. 282–291.
31. Mood, A.M.; Graybill, F.A.; Boes, D.C. *Introduction to the Theory of Statistics*, McGraw-Hill: New York, NY, USA, 1974.
32. Czerwinski, A.; Szlachetka, J. Efficiency of photonic state tomography affected by fiber attenuation. *Phys. Rev. A* **2022**, *105*, 062437. [CrossRef]
33. Hogg, R.V.; Craig, A.T. *Introduction to Mathematical Statistics*, 4th ed.; Macmillan: New York, NY, USA, 1978.
34. David, H.A.; Nagaraja, H.N. *Order Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 2003.
35. Cohen, A. C. Simplified estimators for the normal distribution when samples are singlycensored or truncated. *Technometrics* **1959**, *1*, 217–237. [CrossRef]
36. Balakrishnan, N.; Cohen, A.C. *Order Statistics and Inference. Estimation Methods*, 1st ed.; Academic Press Inc.: London, UK, 1991.
37. Smith, P. J. *Analysis of Failure and Survival Data*; 1st ed.; Chapman and Hall: New York, NY, USA, 2002.
38. Mahmoud, M.R.; Sultan, K.S.; Saleh, K.M. Progressively censored data from the linear exponential distribution: Moments and estimation. *Int. J. Stat.* **2006**, *64*, 199–215.
39. Casella, G.; Berger, R.L. *Statistical Inference*, 2nd ed.; Thomson Learning Inc.: Duxbury, MA, USA, 2002.
40. Nelson, W.; Hahn, G.J. Linear Estimation of a Regression Relationship from Censored Data–Part II Best Linear Unbiased Estimation and Theory. *Technometrics* **1973**, *15*, 133–150.