Journal of Interdisciplinary Mathematics
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 28 (2025), No. 2, pp. 597–606
DOI : 10.47974/JIM-2103

A mathematical approach to model signal variation in an aquatic robot's environmental sensor using heavy-tailed distributions

Anuj Nain †
Department of Statistics
Central University of Rajasthan
Ajmer 305817
Rajasthan
India

Govind Murari Upadhyay *
Department of Computer Applications
Manipal University Jaipur
Jaipur 303007
Rajasthan
India

Navdeep Bohra §
Department of Computer Science and Engineering
Maharaj Surajmal Institute of Technology
Janak Puri 110058
Delhi
India

Pramod Kumar Soni ‡
Linesh Raja @
Department of Computer Applications
Manipal University Jaipur
Jaipur 303007
Rajasthan
India

† *E-mail:* anuj.nain.stats@gmail.com
* *E-mail:* govind.upadhyay@jaipur.manipal.edu (Corresponding Author)
§ *E-mail:* navdeepbohra@msit.in
‡ *E-mail:* pramod.soni@jaipur.manipal.edu
@ *E-mail:* linesh.raja@jaipur.manipal.edu

© Taru Publications

Abstract

In aquatic environments, autonomous robots equipped with environmental sensors play a pivotal role in gathering critical data to monitor and understand complex ecosystems. This research paper explores the modeling of signal variations generated by an aquatic robot submerged in water, considering factors such as temperature, pressure, salt content, and the number of living species in the surroundings. The primary focus is on characterizing and quantifying the random error inherent in the signal, with an emphasis on heavy-tailed probability distributions to account for extreme events and outliers. Various heavy-tailed distributions are examined in the context of their suitability to represent the distribution of errors in the robot's signal. The research is conducted through extensive simulations enabling controlled experimentation and analysis of signal characteristics. AIC, BIC, K-S test for goodness of fit have been used to identify the best fitting distribution. This interdisciplinary research uses Mathematical Statistics to provide useful insights into the error dynamics of aquatic robots thus helping to improve their quality, consequently proving beneficial for various stakeholders which include, Engineers, Environmentalists, mathematicians, statisticians, and AI developers.

Subject Classification: *76B75.*

Keywords: *Aquatic robots, Error dynamics, Heavy-tailed distributions, Distribution fitting, Mathematical statistics, Simulation, Good health and well-being.*

1. Introduction

In the realm of aquatic robotics, accurate data acquisition is vital for understanding and managing complex underwater ecosystems. This research delves into the intricacies of signal variation exhibited by an environmental sensor installed on an aquatic robot. The study investigates the effects of temperature, pressure, salt content, and the presence of living species in the aquatic environment on the sensor's signal. Emphasizing the need to account for extreme events and outliers, this research employs heavy-tailed probability distributions to model the inherent random error in the signal.

To achieve its objectives, this study employs a simulation-based approach conducted within the confines of an Automation Lab. By simulating a range of scenarios that mimic diverse aquatic conditions, the research dissects the complexities of signal variation in a controlled manner. The focus of the investigation lies in assessing the suitability of heavy-tailed distributions in capturing the unique characteristics of signal errors. A probability distribution is called a heavy-tailed if it gives more weightage to larger values of the random variable. Details of heavy-tailed distribution are discussed in section 2.

This paper presents interdisciplinary research that uses Mathematics and Statistics to provide an application in robotics. The study is not only confined to theoretical exploration; it seeks to bridge theory with practical application. By utilizing simulation data generated in the Automation Lab, the study offers insights into the behavior of heavy-tailed distributions in a real-world context. Collaborative efforts between robotics, statistics, and environmental science contribute to a holistic approach to understanding signal variation.

The outcomes of this study hold substantial significance in the field of aquatic robotics and environmental monitoring. By enhancing the precision of error modeling, this research informs the development of anomaly detection algorithms, data filtering techniques, and adaptive decision-making processes. Furthermore, the research underscores the importance of interdisciplinary collaboration in tackling the challenges posed by underwater data acquisition and analysis. In the following sections, we will delve into our pragmatic methodology for comprehending error dynamics within aquatic robot systems.

2. Methodology

To investigate the error dynamics in an aquatic robot's environmental sensor, a structured approach is employed, encompassing data collection, simulation, and statistical analysis. The methodology involves generating a dataset of 2000 observations to comprehensively explore the dynamics of errors made by the robots.

2.1 *Experiment and Simulation*

Controlled simulations have been conducted within the Automation Lab, Department of Applied Sciences, Meerut Institute of Engineering and Technology, to ensure precise and reproducible experimentation. Combinations of Temperature, Pressure, Salt content, and the number of species are selected as input parameters. These parameters are systematically varied to create a diverse range of scenarios. For each combination, the initial values of the input parameters are specified. The aquatic robot's environmental sensor generates signals in response to the input parameters, producing the observed outputs. The difference between the expected output (signal generated based on input parameters) and the observed output constitutes the error for each observation. Simulated scenarios emulate a wide spectrum of aquatic conditions, allowing for controlled manipulation of input parameters. The simulation process yields a dataset of 2,000 observations, each containing input parameter values, observed outputs, and corresponding errors.

2.2 *Heavy-tailed distributions and Statistical Analysis*

For non-negative data, heavy-tailed distributions are those whose right-tailed probabilities are greater than the exponential one mathematically and can be represented as:

$$\lim_{X\to\infty} \frac{e^{-\lambda X}}{1-F(X)} = 0 \quad \text{for any } \lambda > 0 \tag{1}$$

Where F(X) is the cumulative distribution function of X. For details on heavy-tailed distributions, one may refer to [1]–[3]. Note here that the various heavy-tailed distributions mentioned in the paper are characterized by their constants called as parameters and the word "parameter" here is not to be confused with the input parameters of the experiment. By fitting these distributions to the data, we mean estimating the value of these parameters [3].

A range of heavy-tailed probability distributions are considered for modeling the error distribution, given the inherent variability and potential for extreme events. These distributions are fitted onto the dataset of errors using R statistical software. Goodness-of-fit tests, such as the Kolmogorov-Smirnov test is performed to assess the adequacy of each distribution in representing the observed errors. Parameters of the fitted distributions are estimated to quantify the characteristics of the error distribution under different scenarios [4], [5]. Below we mention various Heavy-tailed distributions used in the paper along with their Pdf's.

2.2.1. The Burr distribution

The pdf of Burr distribution used by R software [6] and given in Eq. (2)

$$f(x|\alpha,\beta,\gamma) = \frac{\alpha\gamma\left(\frac{x}{\beta}\right)^{\gamma}}{x\left[1+\left(\frac{x}{\beta}\right)^{\gamma}\right]^{\alpha+1}} \text{ For } x>0,\ \alpha>0, \beta>0 \text{ and } \gamma>0 \tag{2}$$

2.2.2. The Gamma Distribution

The pdf of Gamma distribution used by R software [7] can be defined as

$$f(x|\alpha,\sigma)=\frac{x^{\alpha-1}e^{-x/\sigma}}{\sigma^{\alpha}\,\tilde{A}\alpha}\ \text{ For } x>0,\ \alpha>0 \text{ and } \sigma>0 \tag{3}$$

2.2.3. The Gaussian Distribution

The pdf of Gaussian distribution used by R software which is as follows.

$$f(x|\mu,\sigma)=\frac{e^{-(x-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma}} \tag{4}$$

For $-\infty<x<\infty,\ -\infty<\mu\langle\infty \text{and } \sigma\rangle 0$

2.2.4. The Weibull Distribution

The pdf of Weibull distribution used by R software [8].

$$f(x|\alpha,\beta)=\left(\frac{\alpha}{\beta}\right)\left(\frac{x}{\beta}\right)^{\alpha-1}\mathrm{e}^{-\left(\frac{x}{\beta}\right)^{\alpha}}\ \text{ For } x>0,\ \alpha>0,\ \beta>0 \tag{5}$$

2.2.5 The Log-normal distribution

The pdf of Log normal distribution used by R software [9] and can be represented as.

$$f(x|\mu,\sigma)=\frac{e^{-(\log(x)-\mu)^2/2\sigma^2}}{\sqrt{2\pi\sigma x}} \tag{6}$$

For $-\infty<x<\infty,\ -\infty<\mu\langle\infty \text{and } \sigma\rangle 0$

2.2.5. The Inverse Burr distribution

The pdf of Inverse Burr distribution used by R software [10].

$$f(x|\alpha,\beta,\gamma)=\frac{\alpha\gamma\left(\frac{x}{\beta}\right)^{\gamma\alpha}}{x\left[1+\left(\frac{x}{\beta}\right)^{\gamma}\right]^{\alpha+1}}\ \text{ For } x>0,\ \alpha>0,\ \beta>0 \text{ and } \gamma>0 \tag{7}$$

2.2.6. The Inverse Gamma Distribution

The pdf of Inverse Gamma distribution used by R software [11].

$$f(x|\alpha,\beta)=\frac{\alpha^{\beta}e^{-\alpha/x}}{x^{1+\beta}\,\text{Ãâ}} \text{ For } x>0,\ \alpha>0 \text{ and } \beta>0 \tag{8}$$

2.2.7 The Inverse Gaussian Distribution

The pdf of Inverse Gaussian distribution used by R software [12].

$$f(x|\mu,\sigma)=(2\pi\phi x^{3})^{-\left(\frac{1}{2}\right)}e^{-(x-\mu)^{2}/2\phi\mu^{2}x} \tag{9}$$

For $-\infty < x < \infty, \mu > 0$ and $\phi > 0$

The analysis provides insights into the suitability of various heavy-tailed distributions for modeling error dynamics in the aquatic robot's environmental sensor. Distributional characteristics such as tail behavior, central tendency, and dispersion are examined to identify patterns and anomalies. Findings shed light on the impact of input parameters (temperature, pressure, salt content, species count) on error propagation. Conclusions are drawn regarding the distribution that best captures the error dynamics and the potential implications for the robot's performance in different aquatic environments.

By following this methodology, the study unveils the intricate relationship between input factors and error dynamics in the aquatic robot's environmental sensor. Through systematic data generation, controlled simulation, and rigorous statistical analysis, the research contributes to a comprehensive understanding of how the robot responds to varying environmental conditions.

2.3. *Introducing new risk measures*

Error in the response delineates a shortfall, which is quantified as a random variable to gauge the extent of this shortfall, risk metrics can be established, by leveraging a designated probability distribution. We introduced and elucidated two critical risk measures:

Partial Robotic Deficiency PRD_P : Let X denote the error random variable and Xp is the 100p quantile of the distribution of X. Then Partial Robotic Deficiency, denoted as PRDp is same as Xp and satisfies.

$$P(X>X_{p})=1-p \text{ For } F_{X}(X_{p})=p \tag{10}$$

Total Robotic Deficiency TRD_P: It calculated as the mean of all PRD's surpassing the predetermined threshold.

$$\text{TRD}_p(X_p) = E[X \mid X > X_p] = \frac{\int_{X_p}^{1} xf(x)dx}{1-F_X(X_p)} = \frac{\int_{X_p}^{1} xf(x)dx}{1-P} \tag{11}$$

3. Result

All the computational tasks were executed using the R software. In Figure 1, the graph illustrates the Empirical density plot, accompanied by empirical evaluations of PRD_p and TRD_p. The data reveals a prevalence of positive heavy-tailed errors. To align with our analysis, we selected a value of p=0.90, resulting in empirical approximations of $\text{PRD}_{.90}$ = 11.784 and $\text{TRD}_{.90}$ = 17.152. The summarization of various heavy tailed distributions is outlined in Table 1, showcasing the Weibull distribution as the optimal fit. This determination is substantiated by its minimal AIC (387.174) and BIC (425.495), also the computed $\text{PRD}_{.90}$ = 11.775 and $\text{TRD}_{.90}$ = 17.041 values, closely resemble the empirical counterparts.

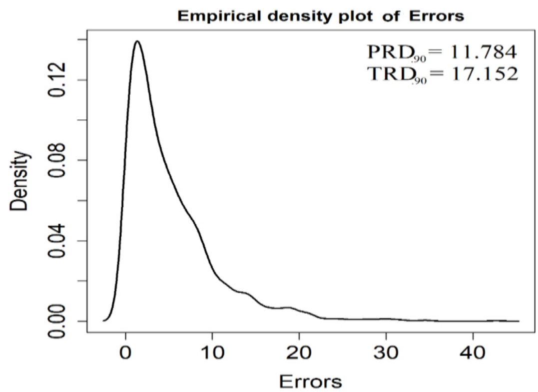

Figure 1

Empirical Density plot of the Errors

Figure 2 visually depicts the congruence between the fitted Weibull distribution and the empirical density, accompanied by corresponding parameter estimates $(\alpha = 1.012,\ \beta = 5.238)$. The R software uses inbuilt function to obtain Maximum Likelihood Estimates (MLE) of the parameters. To get more details on the technique one may refer [5] [4]. Meanwhile, Figure 3 presents a Q-Q Plot alongside the K-S goodness of fit test results. The KS test is conducted at 0.05 level of significance. The null hypotheses for the KS test state that, "The Weibull distribution with

Table 1

Summary of Various Distributions

Distribution	AIC	BIC	PRD.90	TRD.90
Burr	465.235	592.546	11.283	16.891
Gamma	887.453	958.875	15.283	25.234
Gaussian	398.485	465.945	11.354	18.984
Weibull	387.174	425.495	11.775	17.041
Inverse Burr	767.452	886.956	12.265	19.287
Inverse Gamma	576.564	634.772	13.456	23.286
Inverse Gaussian	598.387	656.753	13.654	26.329
Log-normal	773.387	826.656	14.276	21.453
Log-logistic	489.367	546.531	12.984	19.967

parameter estimates $\alpha = 1.012,\ \beta = 5.238,$ fits well for the data". Since the KS test statistic (0.001) is less than the critical value (0.03), the Null hypotheses is accepted. Also, from the QQ-Plot it is visible that the sample quantiles lie along the straight line supporting the fit of Weibull distribution further.

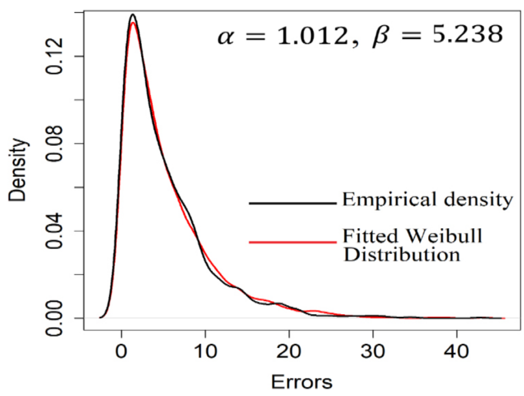

Figure 2

Fitted Weibull Distribution V/S Empirical density.

4. Conclusions

This is an interdisciplinary work that yields valuable insights into the appropriateness of various heavy-tailed distributions for effectively

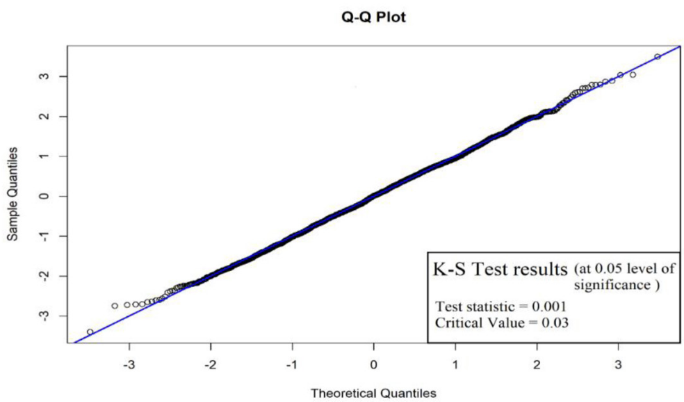

Figure 3

Q-Q Plot

modeling error dynamics in the aquatic robot's environmental sensor. In our knowledge we are amongst the pioneers to advocate this type of research, no further work has been done in this domain so far. Understanding the intricacies of error dynamics in aquatic robots can enhance signal quality, thereby benefiting environmentalists, AI developers, statisticians, and related stakeholders.

The comprehension of error dynamics within aquatic robots holds the potential to revolutionize real-time environmental applications, offering an array of benefits for environmentalists. By infusing these robots with an acute understanding of how errors propagate, environmentalists gain a powerful tool that can be harnessed to tackle urgent ecological challenges with unprecedented precision and timeliness.

In parallel, this understanding also presents a remarkable opportunity for AI developers to elevate the quality of robotic systems without incurring substantial costs. Armed with insights into error dynamics, developers can fine-tune and optimize the robots' responses, leading to more reliable and efficient operations. The result is an innovation-driven synergy that empowers these robots to seamlessly navigate complex aquatic environments, aiding environmentalists in their mission to monitor and protect fragile ecosystems. Simultaneously, the field of statistics stands to gain from this research endeavor. The intricate interplay between input factors and error propagation introduces a fresh avenue for statisticians to explore applying heavy-tailed distributions to this unique domain.

References

[1] J. Beirlant, G. Matthys, and G. Dierckx, "Heavy-Tailed Distributions and Rating," *ASTIN Bulletin: The Journal of the IAA*, vol. 31, no. 1, pp. 37–58 (2001), doi: 10.2143/AST.31.1.993.

[2] A. J. McNeil, "Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory," *ASTIN Bulletin: The Journal of the IAA*, vol. 27, no. 1, pp. 117–137 (1997), doi: 10.2143/AST.27.1.563210.

[3] S. I. Resnick, "Discussion of the Danish Data on Large Fire Insurance Losses," *ASTIN Bulletin: The Journal of the IAA*, vol. 27, no. 1, pp. 139–151 (1997), doi: 10.2143/AST.27.1.563211.

[4] C. Biswas, T. Biswas, and R. Biswas, "Relative (p, q)-φ order and relative (p, q)-φ type oriented some growth investigations of composite p-adic entire functions," *Journal of Interdisciplinary Mathematics*, vol. 25, no. 2, pp. 233–246 (2022).

[5] M. R. Spiegel, J. J. Schiller, and R. A. Srinivasan, *Schaum's Outline of Probability and Statistics*, McGraw-Hill Education (2013).

[6] A. R. Hakim, I. Fithriani, and M. Novita, "Properties of Burr distribution and its application to heavy-tailed survival time data," *Journal of Physics: Conference Series*, vol. 1747, no. 1, 012034, IOP Publishing Ltd, Jan. (2021).

[7] I. Slimane and Z. Dahmani, "Normalized fractional inequalities for continuous random variables," *Journal of Interdisciplinary Mathematics*, vol. 25, no. 2, pp. 335–349 (2022).

[8] W. J. Padgett, "Weibull Distribution," in *International Encyclopedia of Statistical Science*, M. Lovric, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1800–1804 (2011).

[9] E. L. Crow and K. Shimizu, *Lognormal Distributions*, New York: Marcel Dekker (1987).

[10] E. Chiodo and P. De Falco, "Inverse Burr distribution for extreme wind speed prediction: Genesis, identification and estimation," *Electric Power Systems Research*, vol. 141, pp. 549–561 (2016).

[11] A. Llera and C. F. Beckmann, "Estimating an inverse gamma distribution," *arXiv preprint* arXiv:1605.01019 (2016).

[12] J. L. Folks and R. S. Chhikara, "The inverse Gaussian distribution and its statistical application—A Review," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 40, no. 3, pp. 263–275 (1978).

Received February, 2024