Theses and Dissertations | 1. Thesis and Dissertation Collection, all items

2019-03

USING FISHER INFORMATION TO CREATE SELF-REFLECTION IN AUTONOMOUS SYSTEMS

Ji, Junghoon

Monterey, CA; Naval Postgraduate School

http://hdl.handle.net/10945/62259

NAVAL
POSTGRADUATE
SCHOOL

MONTEREY, CALIFORNIA

THESIS

**USING FISHER INFORMATION TO CREATE
SELF-REFLECTION IN AUTONOMOUS SYSTEMS**

by

Junghoon Ji

March 2019

Thesis Advisor:	Johannes O. Royset
Co-Advisor:	Suhwan Kim,
	Korea National Defense University
Second Reader:	Robert L. Bassett

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704–0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** March 2019	**3. REPORT TYPE AND DATES COVERED** Master's thesis	
4. TITLE AND SUBTITLE USING FISHER INFORMATION TO CREATE SELF-REFLECTION IN AUTONOMOUS SYSTEMS			**5. FUNDING NUMBERS**
6. AUTHOR(S) Junghoon Ji			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

In modern society, the number and popularity of autonomous systems are increasing, and it seems certain that their importance will grow in the future. As early as 2017, Amazon was already working with more than 100,000 warehouse robots, and many companies have begun shipping with drones or autonomous vehicles around the world. In the future, autonomous systems are likely to play a major role not only in the public sector but also in the defense sector. In fact, the Republic of Korea Army introduced a "drone-bot" force in 2018, for defense applications.

Nevertheless, the operation of autonomous systems poses several challenges. One is deciding how the autonomous system will make decisions in an uncertain situation. What if the collected data is scarce, contains extreme values, and follows an unknown distribution? In light of these uncertainties, a robust estimation method is needed. Autonomous systems should make judgments that lead to decisions that not only yield the good results but also, more importantly, avoid catastrophic outcomes.

In this thesis, we present two fast and conservative estimation methods using Fisher information that adapt to the quality and quantity of the data. We compare our two methods with parametric estimates and maximum likelihood estimation under normal, log-normal, and exponential distributions. Finally, we apply the two methods to predict whether an unmanned underwater vehicle can successfully perform a mission.

14. SUBJECT TERMS Fisher information, epi-spline, maximum likelihood estimation, robust PDF			**15. NUMBER OF PAGES** 79
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

**USING FISHER INFORMATION TO CREATE SELF-REFLECTION IN
AUTONOMOUS SYSTEMS**

Junghoon Ji
Captain, Republic of Korea Army
BS, Korea Military Academy, 2010

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

**NAVAL POSTGRADUATE SCHOOL
March 2019**

Approved by: Johannes O. Royset
Advisor

Suhwan Kim
Co-Advisor

Robert L. Bassett
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

iii

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

In modern society, the number and popularity of autonomous systems are increasing, and it seems certain that their importance will grow in the future. As early as 2017, Amazon was already working with more than 100,000 warehouse robots, and many companies have begun shipping with drones or autonomous vehicles around the world. In the future, autonomous systems are likely to play a major role not only in the public sector but also in the defense sector. In fact, the Republic of Korea Army introduced a "drone-bot" force in 2018, for defense applications.

Nevertheless, the operation of autonomous systems poses several challenges. One is deciding how the autonomous system will make decisions in an uncertain situation. What if the collected data is scarce, contains extreme values, and follows an unknown distribution? In light of these uncertainties, a robust estimation method is needed. Autonomous systems should make judgments that lead to decisions that not only yield the good results but also, more importantly, avoid catastrophic outcomes.

In this thesis, we present two fast and conservative estimation methods using Fisher information that adapt to the quality and quantity of the data. We compare our two methods with parametric estimates and maximum likelihood estimation under normal, log-normal, and exponential distributions. Finally, we apply the two methods to predict whether an unmanned underwater vehicle can successfully perform a mission.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AI	artificial intelligence
CVaR	Conditional Value at Risk
FP-MLE	Fisher penalized maximum likelihood estimation
MFI	minimizing Fisher information
MLE	maximum likelihood estimation
PDF	probability density function
ROKA	Republic of Korea Army
UUV	unmanned underwater vehicle

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The use of autonomous systems is likely to increase in the future, and the level of autonomy in these systems will also be more sophisticated than it is now. Increasing autonomy means that human intervention will be less important and that the autonomous system will be able to assess a situation and decide whether to initiate or continue a mission. Given collected information, it becomes necessary to calculate the risk of the mission and to reflect on whether it is too high. In particular, autonomous systems might want to make conservative assessments and decisions when the quantity and quality of the information collected is insufficient. In this thesis, we propose two methods to estimate robustly tail-dependent quantities such as superquantiles and demonstrate how they can be used in applications with autonomous systems.

The thesis is divided into three parts. In the first part, we develop two models (minimizing Fisher information, or MFI, and Fisher penalized maximum likelihood estimation, or FP-MLE) for estimating probability density functions robustly using a small amount of data. Both models find density functions with minimal Fisher information, but are distinct in that the first uses a constraint on the proximity to a benchmark density function and the second leverages a penalty formulation. Both are convex models that are routinely solved in less than one second, even with a data size of 100,000.

In the second part, we study benchmark problem instances with data from normal, log-normal, and exponential distributions and robustly estimate these density functions. When the sample size is five, we confirm that both models give more stable values than parametric estimation and nonparametric maximum likelihood estimation. Also, by adjusting the robustness parameter, the user can adjust the level of conservativeness as desired.

In the third part, we examine the possibility of using the models for decision making within autonomous systems. We estimate the probability that an underwater unmanned vehicle (UUV) can penetrate an area of interest undetected. Using as little as five, ten, and 30 data points, obtained, for example, through reconnaissance or intelligence, the models

yield conservative estimates of the probability that the UUV will be detected during the penetration. The study provides the insight that more information can reduce the risk associated with the mission, even when that information is "bad news" in the sense that the adversary has greater capability than originally expected.

The results of this thesis have the potential to guide autonomous systems to reflect more deeply on the risks associated with a mission, help them decide whether additional information is required, or whether the mission should simply be abandoned.

ACKNOWLEDGMENTS

THIS PAGE INTENTIONALLY LEFT BLANK

I.   BACKGROUND

The number of autonomous systems is increasing, and it seems certain that their importance will continue to grow in the future. As of 2017, Amazon was already working with more than 100,000 warehouse robots named "kiva," and autonomous vehicles have begun shipping to many countries around the world. In the future, autonomous systems will play a large role not only in the private sector but also in the defense sector. The Republic of Korea Army (ROKA) created the "drone-bot" force on September 28, 2018. As this compound word implies, the "drone-bot" is a combination of a drone and a robot, and this force belongs to the Territorial Intelligence Unit. In addition, from 2021 onward, from the corps to the battalion level, ROKA plans to organize a drone-bot force for the entire army [1].

There are challenges that need to be addressed in order to operate an autonomous system. One is about how the autonomous system will make decisions in uncertain situations. What if the data collected by the autonomous system is limited, contains extreme values, or follows an unknown distribution? Under these various uncertainties, a robust estimation method is needed. In this thesis, we present two fast and conservative estimation methods that adapt to the quality and quantity of the data using Fisher information. We apply the methodology in the context of autonomous systems attempting to decide whether a mission can be accomplished.

A.   DECISION UNDER UNCERTAINTY

There are hopes that autonomous systems will be better than humans at making decisions [2]. In general, it is difficult to make good decisions because the result of a decision is not only uncertain, but also the information on which the decision is based is also uncertain. In a setting with uncertainty, it is often prudent to be risk-averse. This holds true for an autonomous system, too. It needs to determine whether a mission is too risky or difficult.

Artificial intelligence (AI) has made several advances over the last decades. For example, in 2016 Alpha Go, an artificial intelligence developed by Google Deep Mind, won a game of GO against Lee Sedol and won the Ke jie in 2017. Alpha Go retired from

the game soon afterwards with the goal of mastering the even more complicated game of Starcraft. This represents an even greater challenge as the opponents' positions are hidden and a player needs to conduct reconnaissance in order to acquire information. Information then becomes uncertain, and it is necessary to make a judgment about whether there is sufficient knowledge to process with a task or pause and collect additional information. This assessment process is especially difficult for an AI as it requires experience and an ability to self-reflect.

Many methods for making decisions under uncertainty have been explored [3–5]. One possibility is to make decisions based on expected values, rather than on risks. If we repeat the decision, this may indeed be the best way. Nevertheless, regardless of whether or not we make a decision once or several times, it may be important to consider risk.

B.   SUPERQUANTILE RISK

Risk is inherently one-sided and tied to the likelihood and magnitude of outcomes for a random phenomenon. Standard deviations are not well suited for this purpose due to their symmetry; values above and below the mean are treated identically. Superquantiles give the average of the worst-cases and are therefore asymmetrical and better suited for decision making under risk [6–10]. Specifically, the superquantile at level α of a random variable is the average of the $(1-\alpha)100\%$ worst outcomes. Superquantiles were introduced in [11] with the purpose of having an application-independent terminology distinctive from the original name Conditional Value at Risk (CVaR) of [9].

There are many advantages to using superquantiles. First, they adapt to any level of risk-averseness by adjusting α. For example, if a decision maker wants to make decisions based on averages, $\alpha = 0$. If we want to make decisions based on the very worst outcome, $\alpha = 1$. Second, superquantiles lead to convex optimization problems with significant computational advantages. Third, superquantiles, in contrast to quantiles, account for the magnitude of high outcomes. Figure 1 shows two probability density functions (PDFs) with the same 0.9-quantile but rather different 0.9-superquantiles. The average after 0.9-quantile of the left PDF is greater than the average after 0.9-quantile of the right PDF because of

the different tail shapes. Thus, superquantiles are sensitive to the tail of a PDF; see [6] for further discussion and background.
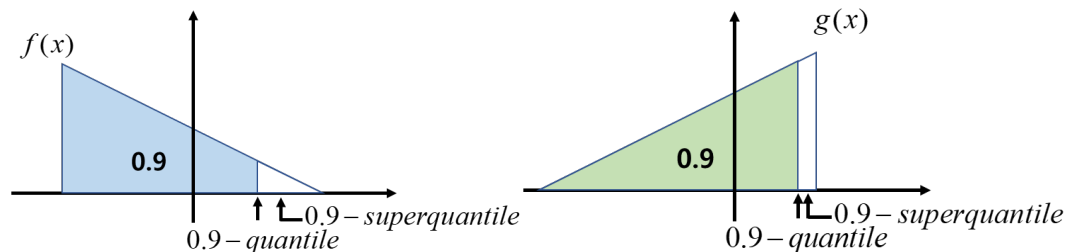
Figure 1. Two PDFs with the same quantile, but different superquantiles.

C. FISHER INFORMATION

In decision making for autonomous systems, it may be important to assess the information content in present and future data. A large amount of collected data might not have much vital information, and further data collection would be necessary. Conversely, if the collected data has a large amount of important information, additional data collection will not be required. Thus, the information content in the collected data is central to the operation of autonomous systems. Fisher information measures the amount of information that a random variable X carries about an unknown parameter θ of a distribution that models X and is expressed as

$$I(\theta) = -E\left(\frac{\partial^2 l(X;\theta)}{\partial \theta^2}\right),$$

where $l(X;\theta)$ is the log-likelihood function of X. Precision to which we can estimate θ is fundamentally limited by the Fisher information of the corresponding likelihood function through the Cramer-Rao bound

$$Var(\hat{\theta}) \geq \frac{1}{I(\theta)},$$

3

i.e., the variance of an unbiased estimator $\hat{\theta}$ of θ is always greater than or equal to the inverse of Fisher information; see [12, 13]. Fisher information is used in [14] to detect changing aviation threats; in [15] for selecting the best neighbor node, which can improve positioning accuracy; and in [16] to access stability in the performance of public transportation systems.

In this paper, we propose two models to control the robustness of nonparametric estimates of PDFs using Fisher information. The basis of both models is a nonparametric estimation method based on epi-splines proposed in [17]; see also [7, 18–21]. The first model is one in which the Fisher information is minimized subject to a constraint ensuring that the resulting PDF is "near" a maximum likelihood estimate or some other nominal estimate. In the second model, a maximum likelihood objective function is penalized with a weighted Fisher information term.

D.   ESTIMATION WITH SMALL DATA

There is an immense body of literature on statistical estimation for univariate PDFs. Parametric approaches are often viable when one can reasonably assume the form of the underlying PDF, with nonparametric alternatives being possible when a sufficient amount of data is available [22]. In contrast, if there is little data and limited information about the class of PDFs, then both parametric and classical nonparametric approaches are problematic. The nonparametric approach based on soft information and epi-splines described in [17] becomes interesting in such situations; however, the resulting estimates are necessarily associated with large uncertainty and they benefit from being made more robust.

E.   THESIS STRUCTURE

Following this chapter, we lay out the two models for robust estimation of PDFs (Chapter II). In the next section, we present empirical studies on benchmark problems (Chapter III). Finally, we examine an application related to autonomous systems (Chapter IV). Throughout, we concentrate on univariate PDFs.

II. FORMULATION OF ESTIMATION PROBLEMS

The goal of this study is to estimate a PDF robustly using little data. We develop two convex optimization models using epi-splines for this purpose. The maximum likelihood principle underlies both, and we start by recalling the standard maximum likelihood model [23].

A. MAXIMUM LIKELIHOOD USING EPI-SPLINES

A first-order epi-spline defined on a mesh, one segment, $m_0, m_1, \ldots, m_N$ is a piecewise affine function as illustrated in Figure 2. Let a_k and b_k be the slope and intercept coefficients of the affine function that defines the epi-spline in the kth segment (m_{k-1}, m_k) and let $a = (a_1, a_2, \ldots, a_N)$ and $b = (b_1, b_2, \ldots, b_N)$. Let N be the number of segments and S the number of data points. Then, the maximum likelihood estimation (MLE) problem given data $x_1, x_2, \ldots, x_S$ can be formulated as an optimization model over the vectors a and b as stated next; see [17, 19, 24].
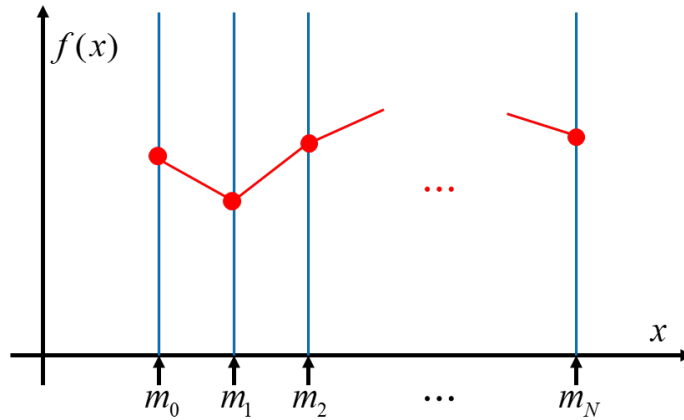
Figure 2.   Example of first-order epi-spline.

(MLE)
$$\min_{a,b} -\frac{1}{S}\sum_{i=1}^{S}\log(a_{k_i}x^i + b_{k_i}),$$

subject to

$$a_1 m_0 + b_1 \geq 0 \tag{1}$$

$$a_k m_k + b_k \geq 0, \qquad\qquad k = 1, \ldots, N \tag{2}$$

$$\sum_{k=1}^{N} b_k (m_k - m_{k-1}) + \sum_{k=1}^{N} \frac{a_k}{2}(m_k^2 - m_{k-1}^2) = 1 \tag{3}$$

$$a_k m_k + b_k = a_{k+1}m_k + b_{k+1}, \qquad k = 1, \ldots, N-1 \tag{4}$$

$$a_k \leq a_{k+1} + \delta, \qquad\qquad k = 1, \ldots, N-1 \tag{5}$$

$$a_k \geq a_{k+1} - \delta, \qquad\qquad k = 1, \ldots, N-1 \tag{6}$$

The objective function minimizes the negative log-likelihood and k_i is the segment in which x^i is located. Constraints (1) to (3) ensure that the resulting PDF is nonnegative and integrates to 1. Constraint (4) ensures it is continuous. Constraints (5) to (6) ensure that the PDF is somewhat "smooth," with δ being a parameter that can be adjusted; see [19, 25–29]. This optimization problem is convex.

B. FORMULATIONS USING FISHER INFORMATION

The model (MLE) tends to produce estimates that underestimate tail probabilities. To compensate for this, we formulate two models based on Fisher information. In the absence of any constraints, minimizing Fisher information on a bounded interval gives the uniform PDF on that interval. Thus, we need to include additional constraints as discussed in the following sections.

1. Minimizing Fisher Information (MFI)

We assumed that the data are independent with the common distribution function $f_\theta(x) = f(x - \theta)$. We recall that the Fisher information of PDF f is given by

$$I(f) = -E\left(\frac{\partial^2 l(X;\theta)}{\partial\theta^2}\right) = -\int_{-\infty}^{\infty} (\log f(x))'' f(x)dx.$$

When f is a first-order epi-spline of the form $f(x) = a_k x + b_k$ for $x \in (m_{k-1}, m_k)$, then the expression simplifies to

$$I(f) = -\sum_{k=1}^{N} \int_{m_{k-1}}^{m_k} (\ln(a_k x + b_k))''(a_k x + b_k) dx = \sum_{k=1}^{N} \int_{m_{k-1}}^{m_k} \frac{a_k^2}{(a_k x + b_k)} dx.$$

The integral is best carried out numerically, and we adopt a Gaussian quadrature rule for this purpose. We refer the reader to the Appendix for details. Using the standard argument from Chapter 7 in [30], we can establish that the error in the optimal solution is due to numerical integration, which can be made arbitrarily small. Using a Gaussian quadrature rule, we obtain that

$$\sum_{k=1}^{N} \int_{m_{k-1}}^{m_k} \frac{a_k^2}{(a_k x + b_k)} dx \approx \sum_{k=1}^{N} \sum_{j=1}^{J} \frac{w_j a_k^2}{(a_k x + b_k)},$$

where w_j and x_j are weights given by the quadrature rule and J is the number of integration points in each segment; see the Appendix.

The Fisher information minimization problem then becomes

(MIF)
$$\min_{a, b} \sum_{k=1}^{N} \sum_{j=1}^{J} \frac{w_j a_k^2}{(a_k x + b_k)},$$

subject to (1) to (6) and also

$$\left| (a_k m_k + b_k) - f_0(m_k) \right| \le \varepsilon \qquad\qquad k = 1, \ldots, N \qquad (7)$$

where f_0 is a candidate PDF. Without constraint (7), the resulting PDF tends to be uniform on $[m_0, m_N]$. Typically, we let f_0 be the estimate obtained by (MLE), but any other PDF can also be used. The constraint (7) easily transforms into two sets of linear constraints $(a_k m_k + b_k) - f_0(m_k) \le \varepsilon$ and $f_0(m_k) - (a_k m_k + b_k) \le \varepsilon$. This optimization problem is also a convex optimization problem. The size of the problem is $2N$ decision variables and $(6N - 1)$ constraints.

2. Fisher-Penalized Maximum Likelihood (FP-MLE)

The second model obtains a robust PDF by applying the addition of a penalty term to the likelihood expression. The resulting model minimizes a weighted sum of log-likelihood and Fisher information:

$$\text{(FP-MLE)} \qquad \min_{a,b} -\frac{1}{S}\sum_{i=1}^{S}\log(a_{k_i}x^i+b_{k_i})+\rho\sum_{k=1}^{N}\sum_{j=1}^{J}\frac{w_j a_k^2}{(a_k x+b_k)},$$

subject to constraints (1) to (6).

In the formulation, ρ is a parameter that determines how much penalty is applied. As the value of ρ increases, a more robust PDF is obtained. The additional penalty term is convex. Therefore, this optimization problem is also a convex optimization problem. The size of the problem is $2N$ decision variables and $(4N-1)$ constraints.

III. EMPIRICAL RESULTS ON BENCHMARK PROBLEMS

In this section, we compare the robustness of estimates from MLE, MFI, and FP-MLE on randomly generated data sets consisting of only five observations from standard normal, log-normal, and exponential distributions.

We solve all optimization problems on an Acer Swift SF314-52 laptop with

- Intel Core i5-7200U CPU @ 2.50GHz 2.71GHz processor

- 8GB RAM

- Windows 10 Home 64-bit operating system

- Pyomo and solver Ipopt [31] in Python

The MFI and FP-MLE solve in usually less than one second for up to 100,000 data points.

A. STANDARD NORMAL DISTRIBUTION

The first benchmark tests are based on data collected from a standard normal distribution using five observations. Figure 3 illustrates the data, the actual PDF, and the standard parametric estimate.
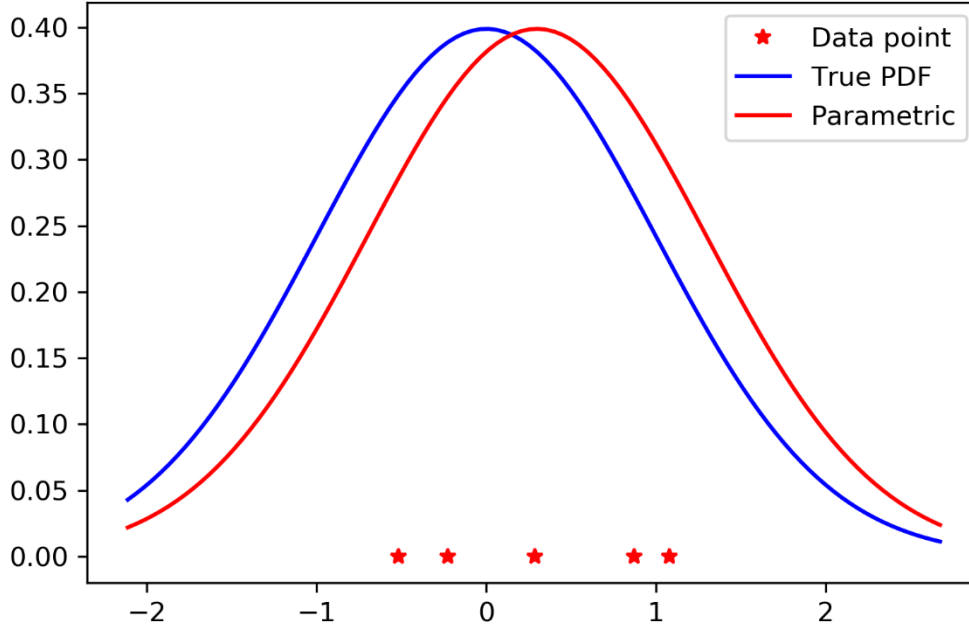
Figure 3.   True PDF (blue) and parametric estimate (red).

1.   USING THE MLE MODEL

The MLE model requires a mesh $m_0, m_1, \ldots, m_N$. The lower bound of the mesh is set to the minimum value of the data minus the difference between the maximum and minimum values. Similarly, the upper bound of the mesh is defined as the maximum value of the data plus the difference between the maximum and minimum values. The number of segments $N = 150$. Using the smoothing parameter $\delta = 0.01$, we obtain the black line in Figure 4.
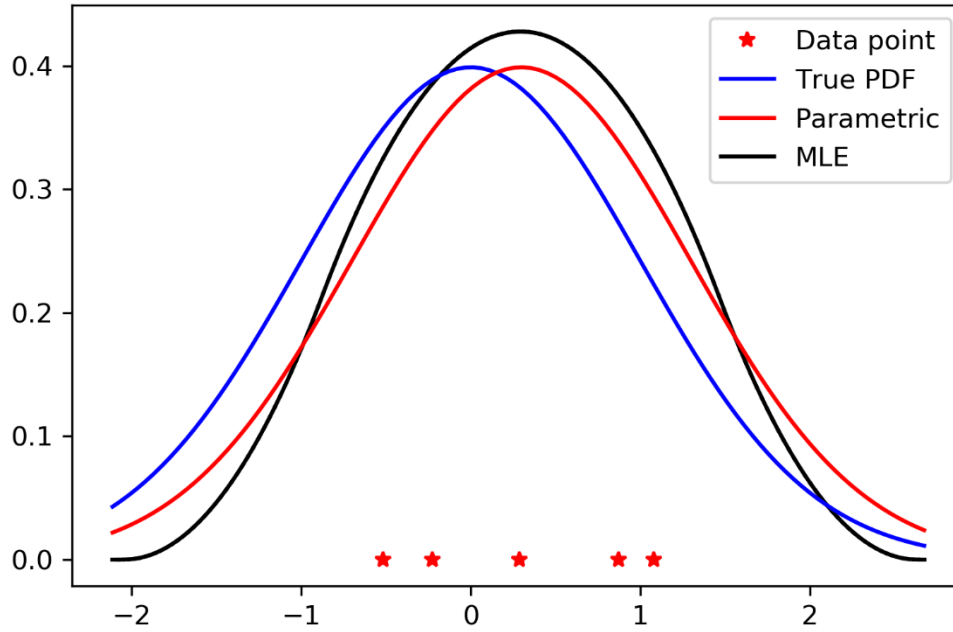
Figure 4.   True PDF (blue), parametric estimate (red), and MLE (black).

Figures 5 and 6 show how the estimated graph changes as δ and the number of segments vary. A low number of segments and the low values of the smoothing parameter induces a seemingly smooth PDF. Conversely, increasing the number of segments or increasing the smoothing parameter causes "spikes" in the estimates. This is an indication of overfitting to the five observations. Overfitting causes lighter tail than true PDF tail and decreases superquantile value. This means that we underestimate the risk. Therefore, we used δ is 0.01 and the number of segments is 150 in this thesis.
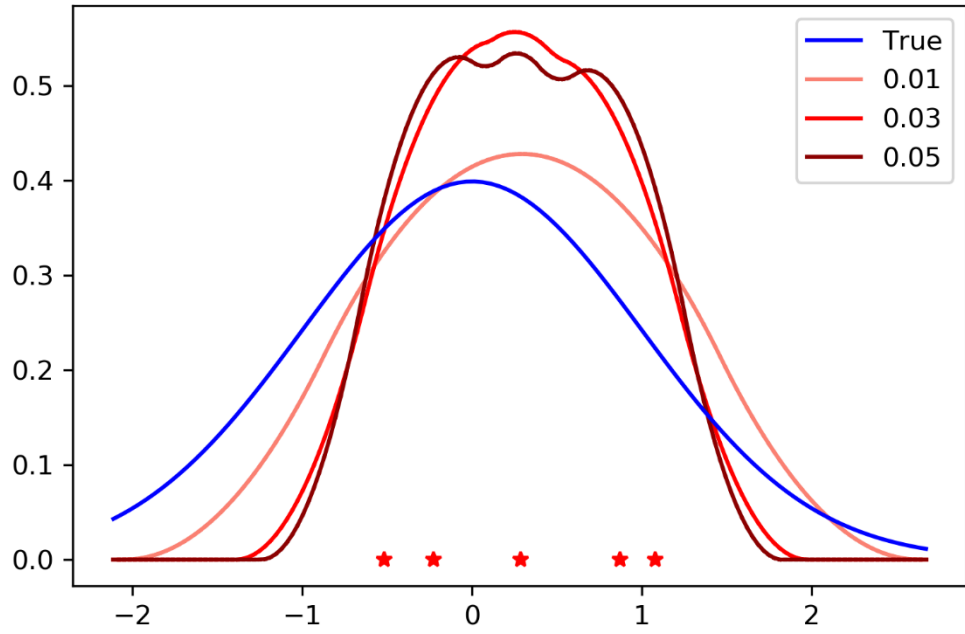
Figure 5.   MLE using $\delta = 0.01, 0.03, 0.05,$ with 150 segments.
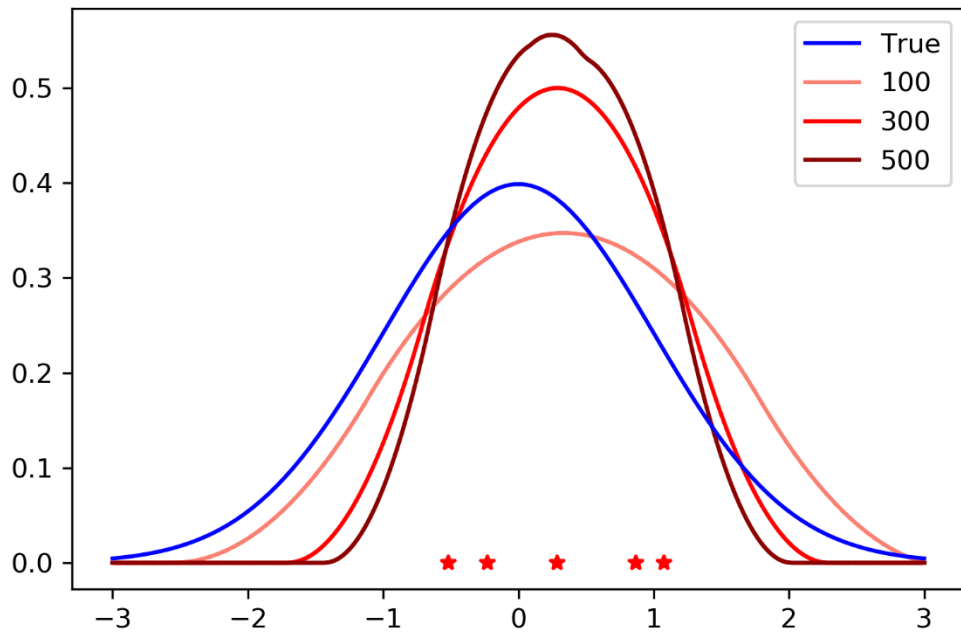
Figure 6.   MLE using $N = 100, 300, 500,$ with $\delta = 0.01$.

2.　　USING THE MFI AND FP-MLE MODELS

We next examine the MFI and FP-MLE models. Figure 7 shows the resulting estimates using $\varepsilon = 0.01$ and $\rho = 0.001$. The models yield similar results, with the resulting PDFs having heavier tails than that of MLE. This is an indication of robustness. Specifically, if the 0.9-superquantile values of the estimates from the MFI and FP-MLE are 1.79 and 1.74, respectively, this is a little higher than the 0.9-superquantile of the MLE estimate, which is 1.72. In comparison, the 0.9-superquantile value of the standard normal PDF is 1.75. Thus, when concerned about conservatively estimating high values of a random quantity, it seems more prudent to use the MFI and FP-MLE rather than MLE. In this case, the MLE produces a PDF that underestimates the probability in the tails. On the other hand, the MFI and FP-MLE furnish more robust estimates.
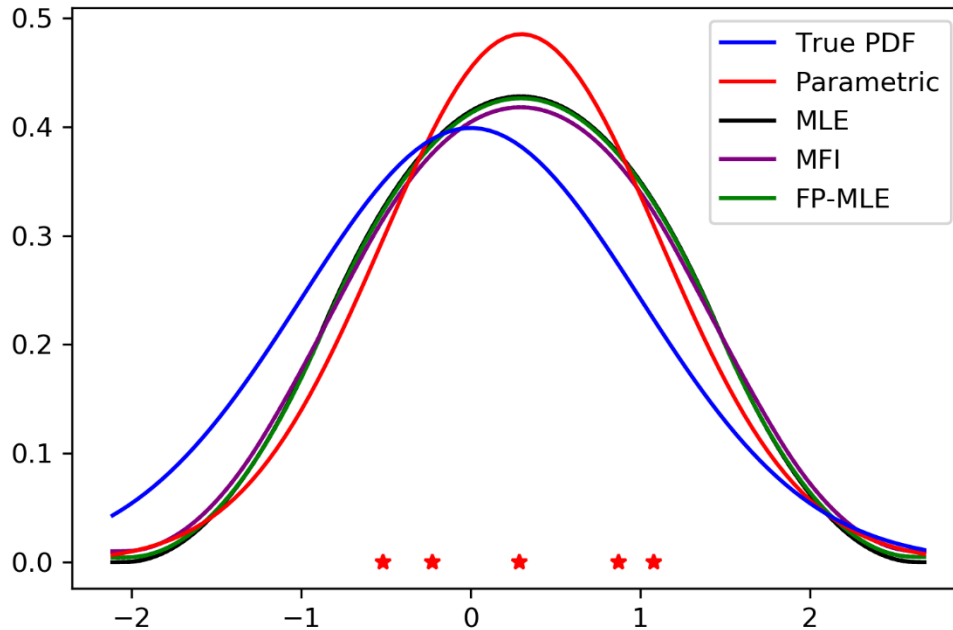
Figure 7.　MFI with $\varepsilon = 0.01$ and FP-MLE with $\rho = 0.001$.

The effect of ε is depicted in Figure 8. In constraint (7) in Chapter II, ε allows that estimates depart from candidate PDF f_0 to converge in a uniform distribution. When we increase ε, MFI produces estimates with increasingly heavy tails and in the limit tends

to a uniform distribution. Table 1 shows the values of a 0.9-superquantile when ε is increased from 0.01 to 0.30. In all these cases, the MFI produces estimates with heavier tails than the true PDF, as quantified by the 0.9-superquantile values.
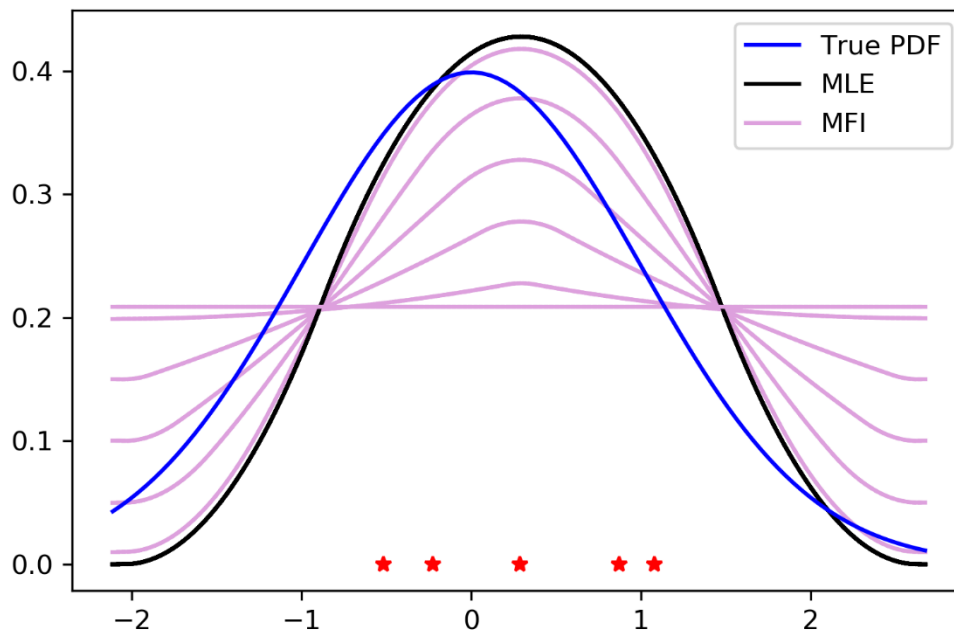
Figure 8.   MFI estimates using $\varepsilon = 0.01, 0.05, 0.10, \ldots, 0.30$.

Table 1.   0.9-superquantiles for MFI estimates using $\varepsilon = 0.01, 0.05, \ldots, 0.30$. The true PDF has 0.9-superquantile of 1.75.

ε	0.01	0.05	0.10	0.15	0.20	0.25	0.30
0.9-superquantile	1.79	2.06	2.26	2.38	2.45	2.47	2.47

Figure 9 illustrates changes in ρ. When we increase ρ, the FP-MLE becomes more robust and produces PDFs with heavier tails. Table 2 shows the value of 0.9-superquantiles when ρ is increased from 0.001 to 0.01. Again, for all values of ρ examined, FP-MLE produces estimates with heavier tails than the true PDF.

Looking at the MFI and FP-MLE, when the robust parameter ε and ρ were increased, we observe that the PDFs are more uniform. Yet, we also find a difference between the two methods. MFI was evenly spaced as the changed ε allowing from the PDF obtained by MLE to converge uniform distribution and converging speed was fast. On the other hand, early on FP-MLE showed a rapid converging speed, but later the converging speed was slow. In addition, FP-MLE tried to maximize the probabilities at the data points.
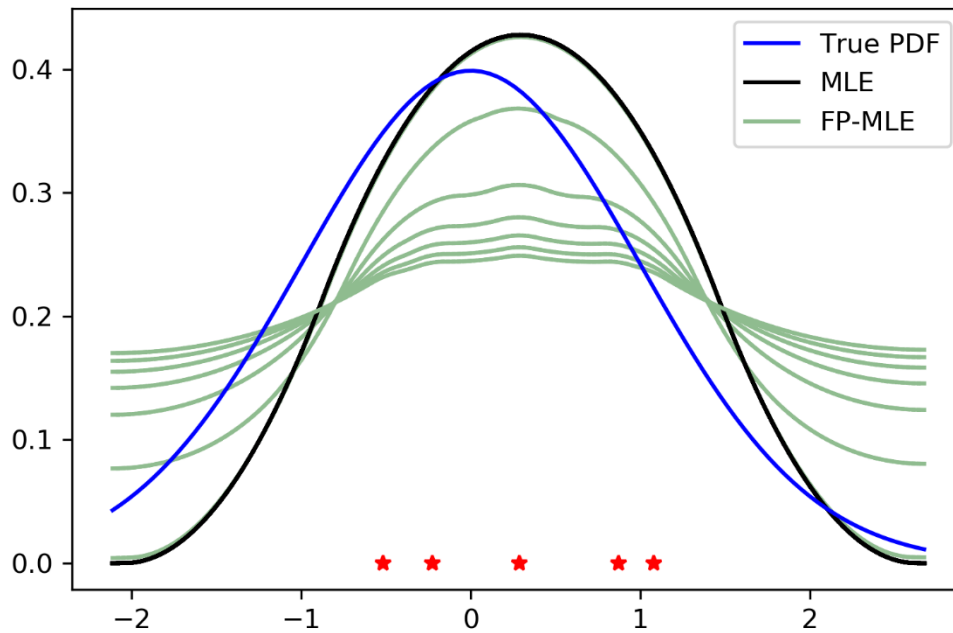
Figure 9.   FP-MLE estimates using $\rho = 0.001, 0.01, 0.02, \ldots, 0.06$.

Table 2.   0.9-superquantiles for FP-MLE using $\rho = 0.001, 0.01, 0.02, \ldots, 0.06$.
The true PDF has 0.9-superquantile of 1.75.

ρ	0.001	0.010	0.020	0.030	0.040	0.050	0.060
0.9-superquantile	1.74	2.16	2.32	2.37	2.39	2.41	2.42

The previous results rely on a single sample size of five. Table 3 shows the aggregated results across 1,000 replications of a sample size of five. The first column shows the mean and the standard deviation. When the robust parameters were increased, the mean value of the 0.9-superquantiles of MFI and FP-MLE increased gradually. The second column shows how many failures occurred in 1,000 cases. Failure is defined as a prediction that is smaller than the actual 0.9-superquantile value 1.75. In other words, it tells the probability that an underestimation error will occur. We should try to avoid underestimation error. For example, if we underestimate our budget, we would become bankrupt. In the cases of the parametric and MLE, the probabilities of underestimation error are 58.4% and 56.0%, respectively. On the other hand, the probabilities of underestimation error of MFI and FP-MLE are 45.8% and 53.4%, respectively. We are able to reduce the probabilities of underestimation error by increasing the robust parameters. There is, however, a trade-off relationship between the increasing mean of 0.9-superquantiles and the decreasing underestimation error.

In Figure 10, we confirm that the highest points of parametric estimates are higher than others. This means that the tails of parametric estimates are lower than others, and both MFI and FP-MLE show heavy tails as well. FP-MLE, however, has more peak shapes and more rely on MLE in comparison to MFI. Because of this, the means of the 0.9-superquantile of FP-MLE increase less than those of MFI when the robust parameters are increased. In comparison MFI (ε is 0.03) and FP-MLE (ρ is 0.01), we confirm that the mean of FP-MLE is smaller than mean of MFI even though the proportion of failure of FP-MLE is less than the proportion of failure of MFI. In conclusion, FP-MLE has better characteristics despite this trade-off relationship. The reason why some PDFs in Figure 10 appear abnormally shaped in the black circle is that we have estimated the sample space from five samples each time, which occasionally leads to especially poor estimates.

Table 3. The means and standard deviations of 0.9-superquantile values and proportion of failure.

Method	$\varepsilon \mid \rho$	Mean ± standard deviation	Proportion of failure
True	-	1.75	-
Parametric	-	1.61 ± 0.56	58.4%
MLE	-	1.69 ± 0.63	56.0%
MFI	0.01	1.81 ± 0.45	45.8%
	0.03	2.25 ± 0.93	31.5%
	0.05	2.53 ± 1.09	24.4%
FP-MLE	0.001	1.73 ± 0.53	53.4%
	0.005	1.87 ± 0.61	42.1%
	0.010	2.15 ± 0.66	25.9%
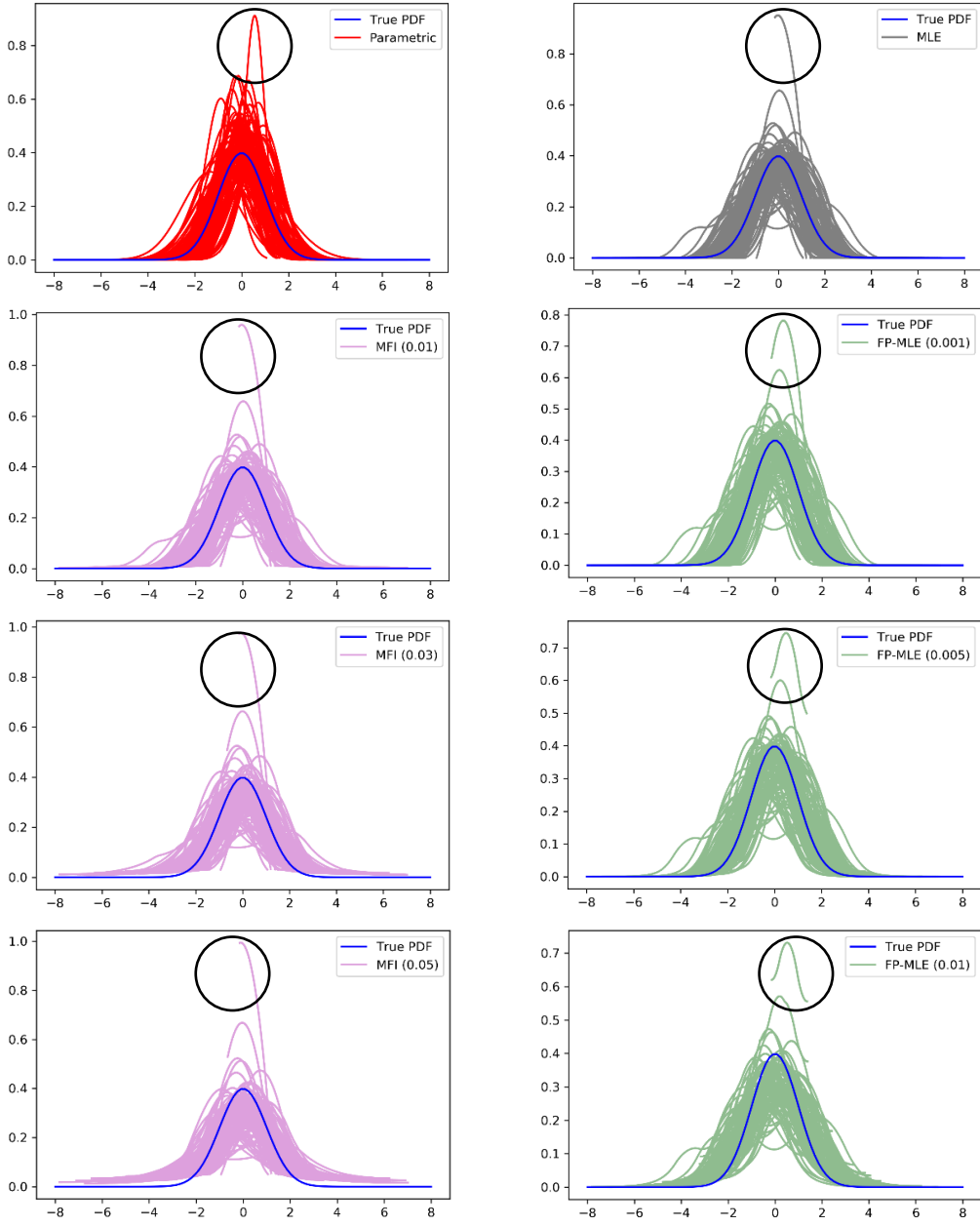
Figure 10. One hundred replications of estimates using five data points.

B.   LOG-NORMAL DISTRIBUTION

For parametric estimation, the probability density function is obtained by substituting the average obtained from the data into the PDF of the known log-normal distribution. Figure 11 illustrates the data, the actual PDF, and the standard parametric estimate.
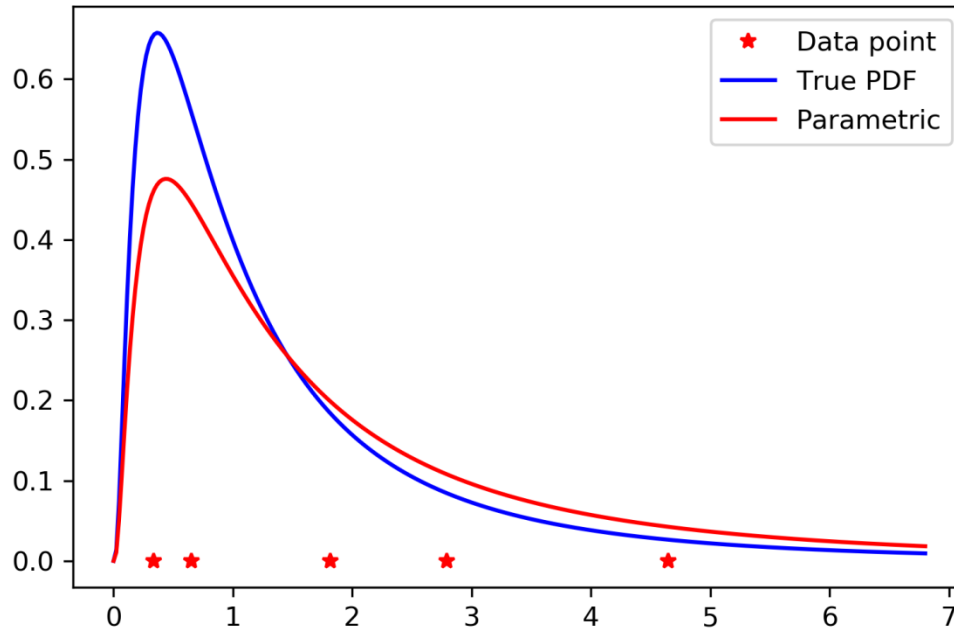
Figure 11.   True PDF (blue) and parametric estimate (red).

1.   USING THE MLE MODEL

Nonparametric estimates were obtained using the Epi-spline method, with a δ of 0.01, 150 segments. The lower bound was defined as zero. The upper bound was defined as the maximum value plus the difference between the maximum and minimum values. The black line in Figure 12 is the probability density function obtained from the MLE.
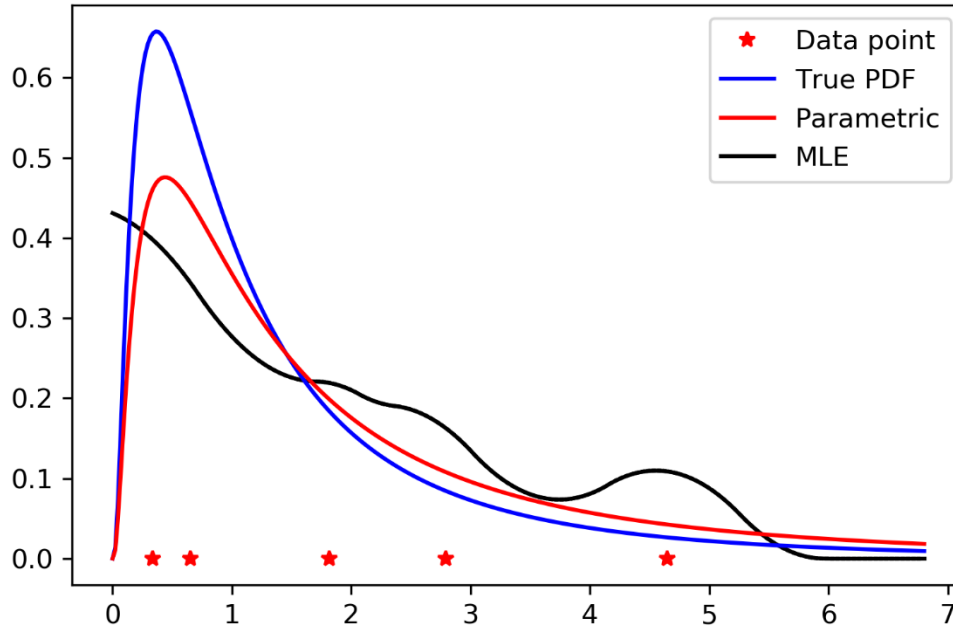
Figure 12.   True PDF (blue), parametric estimate (red), and MLE
(black).

Figures 13 and 14 show how the estimated graph changes as δ and the number of segments vary. The low number of segments and the low values of the smoothing parameter induce a seemingly smooth PDF. Conversely, increasing the number of segments or increasing the smoothing parameter causes "spikes" in the estimates. This is an indication of overfitting to the five observations. This is similar to the previous normal distribution analysis [A.1]. We assumed that the true PDF is a smooth continuous function. Because of this, we think the spike shape reflects noise or error. Therefore, we do not use large segments and δ, but use proper 150 segments and δ is 0.01.
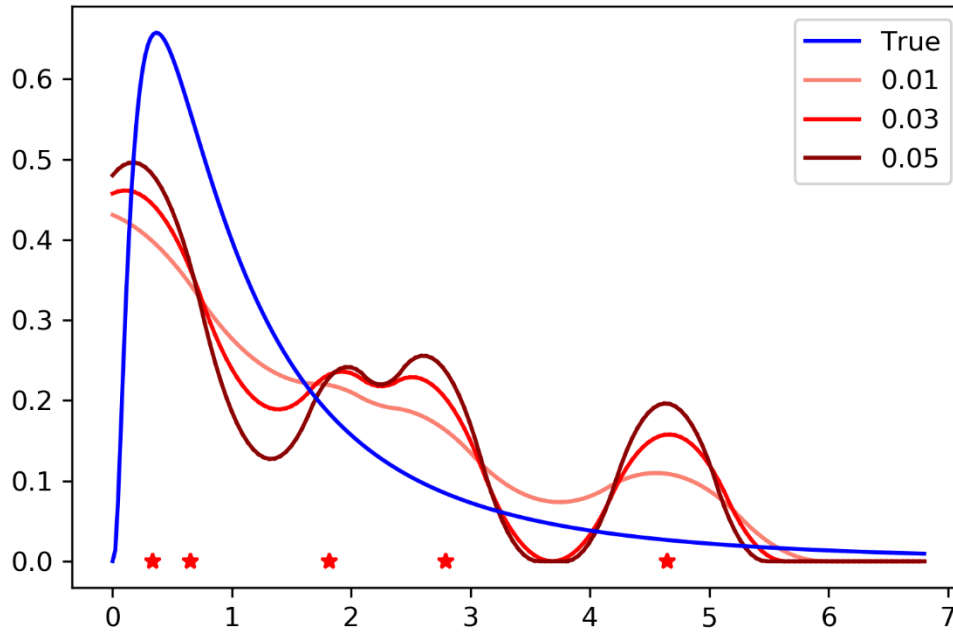
Figure 13.   MLE using $\delta = 0.01, 0.03, 0.05$, with 150 segments.
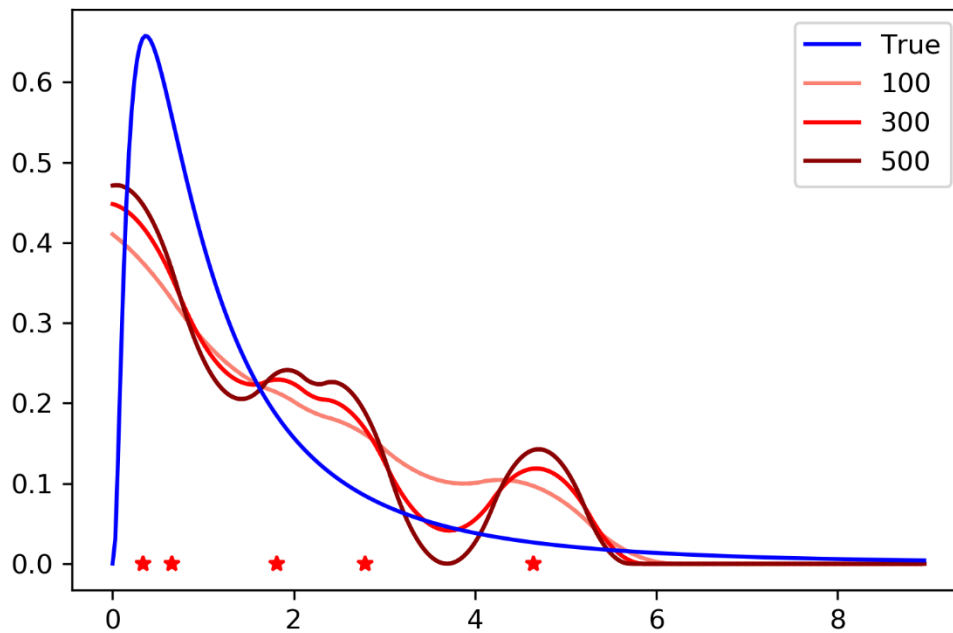
Figure 14.   MLE using $N = 100, 300, 500$, with $\delta = 0.01$.

21

2. USING THE MFI AND FP-MLE MODELS

We next examine the MFI and FP-MLE models. Figure 15 shows the resulting estimates using $\varepsilon = 0.01$ and $\rho = 0.01$. The models yield similar results, with the resulting PDFs having heavier tails than those of the MLE. This is an indication of robustness. Specifically, if the 0.9-superquantile values of the estimates from the MFI and FP-MLE are 5.26 and 5.10, respectively, this is quite a bit higher than the 0.9-superquantile of the MLE estimate, which is 4.90. The two 0.9-superquantile values are not greater than the true 0.9-superquantile value 6.42. This implies that using the MFI and FP-MLE for estimating the PDF, which is following log-normal distribution, could underestimate outcomes. We can handle this issue by increasing robust parameters ε and ρ. On the other hand, the value of the 0.9-superquantile of parametric is 10.28. This is larger than the true 0.9-superquantile value. In fact, the parametric could overestimate outcomes.
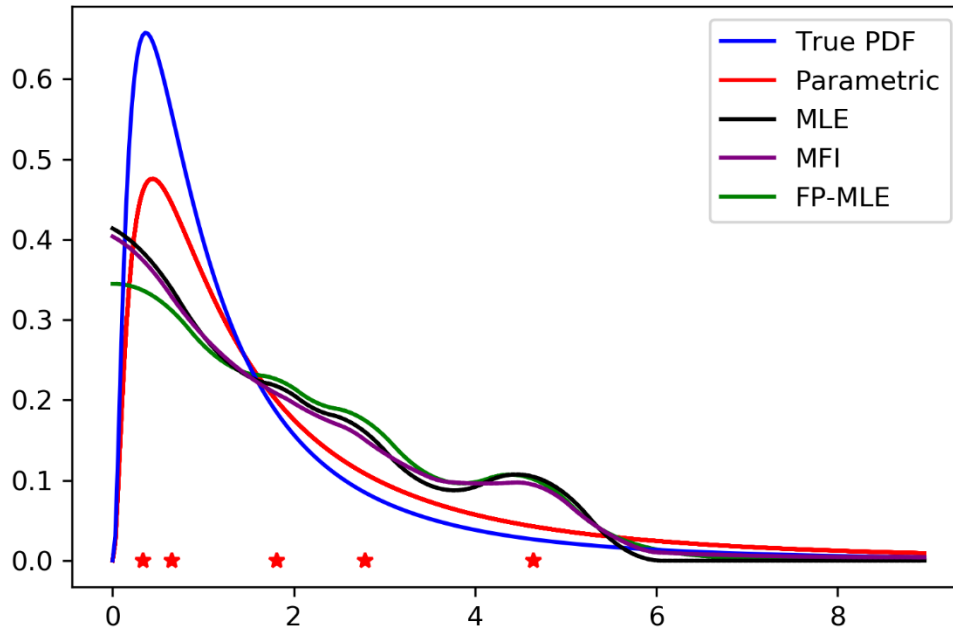
Figure 15. MFI with $\varepsilon = 0.01$ and FP-MLE with $\rho = 0.01$.

As shown in Figure 16, we can change ε to control robustness. As ε increases, the MFI approximates the uniform distribution. Table 4 shows the values of the 0.9-

22

superquantile when ε is increased from 0.01 to 0.30. The value of the 0.9-superquantile increased from 5.26 to 8.57. The value when we used ε is 0.01, however, was lower than the true value. This implies that the MFI result depends on how close the MLE to the true PDF. We need to think about the uncertainty of distribution. If we know the distribution of data, we should use a parametric estimate because this method is the best way to estimate. We could not, however, be sure which distribution is right considering the small amount of data. If the amount of data is sufficient, the MLE is close enough to the true PDF, and our methods are good enough for estimating robust superquantile values.
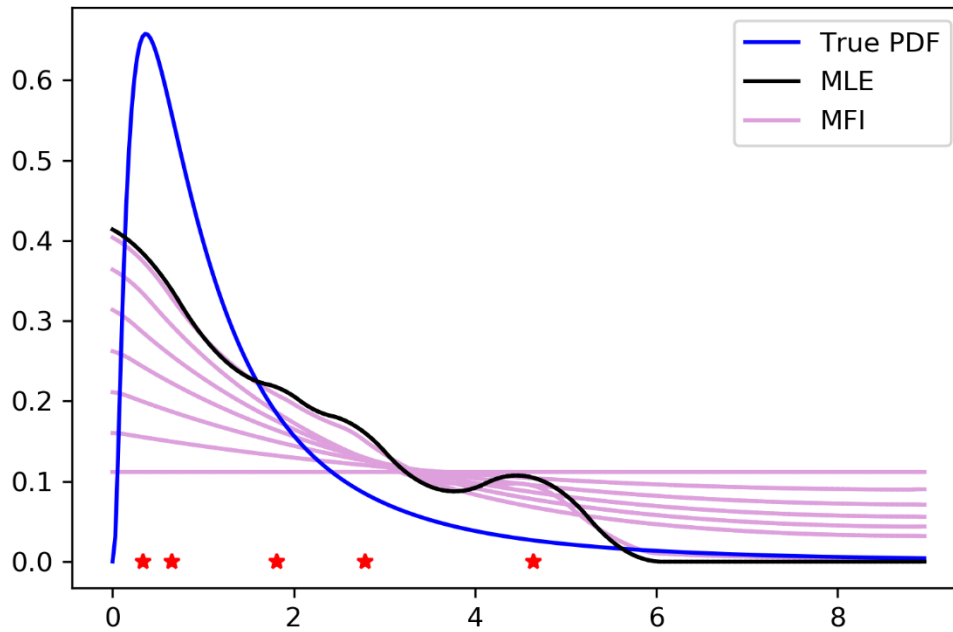
Figure 16.   MFI estimates using $\varepsilon = 0.01, 0.05, 0.10, \ldots, 0.30$.

Table 4.   0.9-superquantiles for MFI estimates using $\varepsilon = 0.01, 0.05, 0.10, \ldots, 0.30$.
The true PDF has 0.9-superquantile of 6.42.

ε	0.01	0.05	0.10	0.15	0.20	0.25	0.30
0.9-superquantile	5.26	7.55	7.90	8.13	8.32	8.46	8.57

Table 5 shows the values of the 0.9-superquantile when ρ is increased from 0.01 to 0.30. The value of the 0.9-superquantile gradually increased from 5.10 to 8.31. When we choose ρ, which is larger than 0.05, the FP-MLE 0.9-superquantile value is greater than the true 0.9-superquantile value 6.42. This is similar to the MFI, but the value of the FP-MLE is much smaller than the value of the MFI. This implies that the FP-MLE not only depends on the MLE but also depends on data. Comparing Figure 16 with Figure 17, we observe the FP-MLE tail shape is sharper than the MFI tail shape. This illustrates that the amount of data has a greater effect on the FP-MLE than it does on the MFI.

Table 5. 0.9-superquantiles for FP-MLE using $\rho = 0.01, 0.05, 0.10, \ldots, 0.30$. The true PDF has 0.9-superquantile of 6.42.

ρ	0.01	0.05	0.10	0.15	0.20	0.25	0.30
0.9-superquantile	5.10	6.73	7.66	8.00	8.16	8.25	8.31
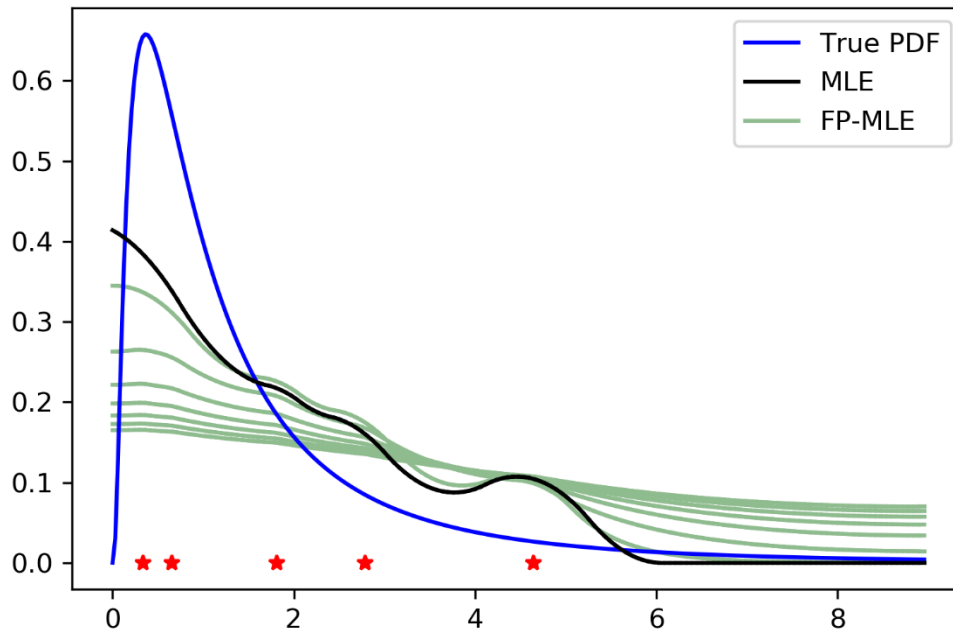
Figure 17. FP-MLE estimates using $\rho = 0.01, 0.05, 0.10, \ldots, 0.30$.

Table 6 shows the mean and standard deviation of the 0.9-superquantile values obtained by each method. One hundred simulations were conducted. For these simulations the number of segments was 150, δ was 0.01, and both ε and ρ were 0.01, 0.05, and 0.10, respectively. Although the parametric 0.9-superquantile value is greater than others, the standard deviation of the parametric is 13.08. Other methods' standard deviations, however, range from 3.73 to 4.60. This implies that even though the parametric shows a small number of failures, the parametric is too risky with small amounts of data because the parametric highly depends on sufficient data. On the other hand, the MFI and FP-MLE are less dependent on the amount of data, and they give us more stable results than the parametric result. In particular, the difference of failure between the parametric and the FP-MLE using $\rho = 0.10$ is only 2. In contrast, the difference of the mean between the parametric and the FP-MLE is 3.86. In fact, we can achieve a lower proportion of failure even though the mean of the 0.9-superquantile value is lower than the mean of the 0.9-superquantile value of parametric.

Table 6.   The means and standard deviations of 0.9-superquantile values.

Method	$\varepsilon \,|\, \rho$	Mean $\pm$ standard deviation	Proportion of failure
True	-	6.42	-
Parametric	-	10.71 ± 13.08	55%
MLE	-	5.18 ± 4.52	78%
MFI	0.01	5.39 ± 4.60	75%
	0.05	5.95 ± 4.02	69%
	0.10	5.90 ± 3.73	70%
FP-MLE	0.01	5.53 ± 4.35	78%
	0.05	6.39 ± 4.35	63%
	0.10	6.85 ± 4.59	57%

C. EXPONENTIAL DISTRIBUTION

For the parametric estimation, the probability density function is obtained by substituting the average obtained from the data into the PDF of a known exponential distribution. Figure 18 illustrates the data, the actual PDF, and the standard parametric estimate.
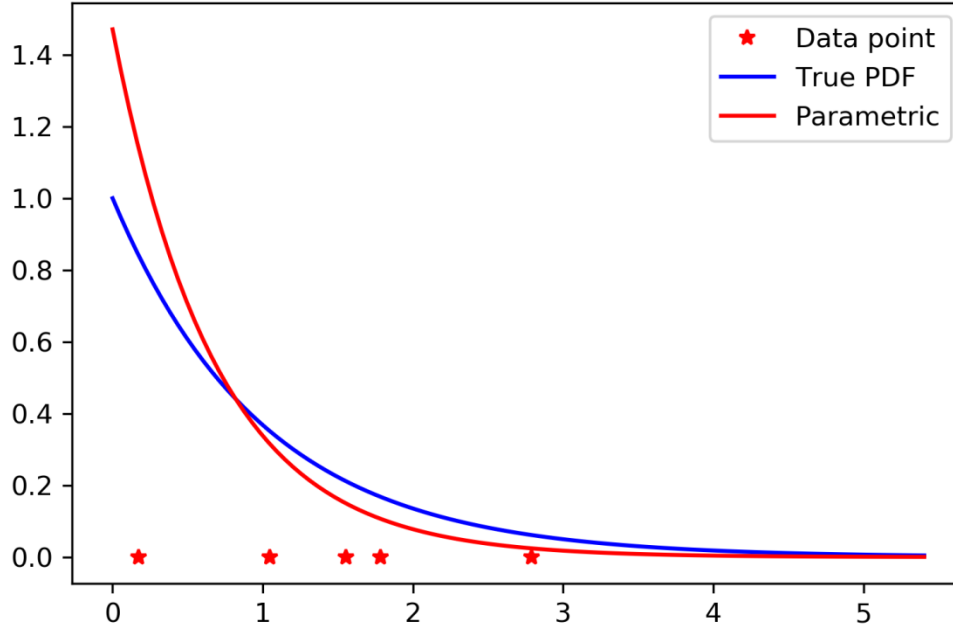
Figure 18.   True PDF (blue) and parametric estimate (red).

1. USING THE MLE MODEL

The MLE model requires a mesh $m_0, m_1, \ldots, m_N$. The lower bound of the mesh is set to zero and minimum values. Similarly, the upper bound of the mesh is defined as the maximum value of the data plus the difference between the maximum and minimum values. The number of segments $N = 150$. Using a smoothing parameter $\delta = 0.01$, we obtain the black line in Figure 19.

Figures 20 and 21 show how the estimated graph changes as δ and the number of segments are changed. The results are similar to the previous two distributions' results. In particular, changing δ has a more powerful effect on the shape than changing the number

of segments, and high complexity tends to bring the overfitting problem. This means that the fitting is good at the currently used data, but when the input data was changed, the fitting is worse than when we used low complexity. Therefore, we use δ is 0.01 and 150 segments.
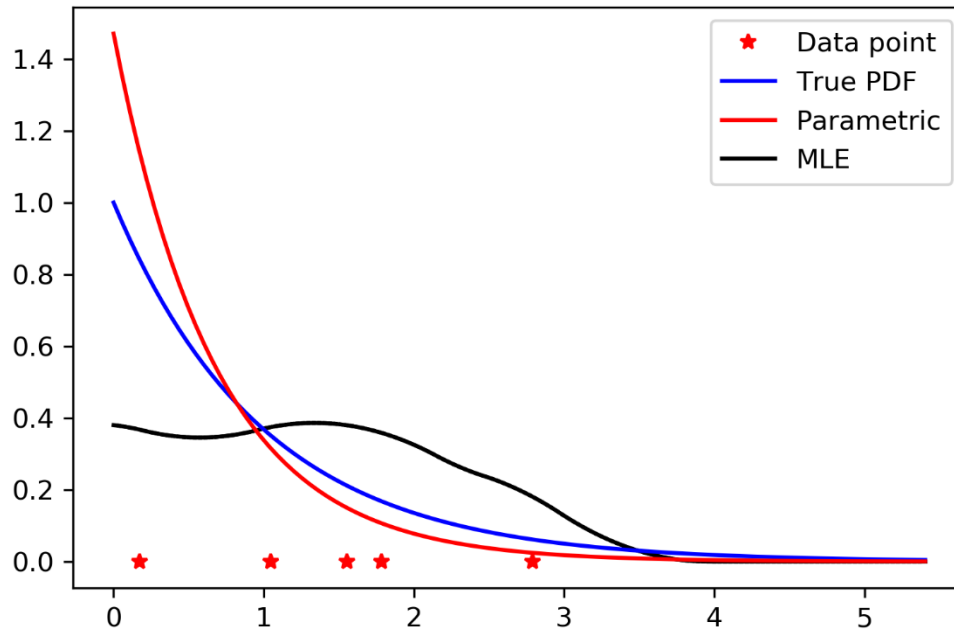
Figure 19.   True PDF (blue), parametric estimate (red), and MLE (black).
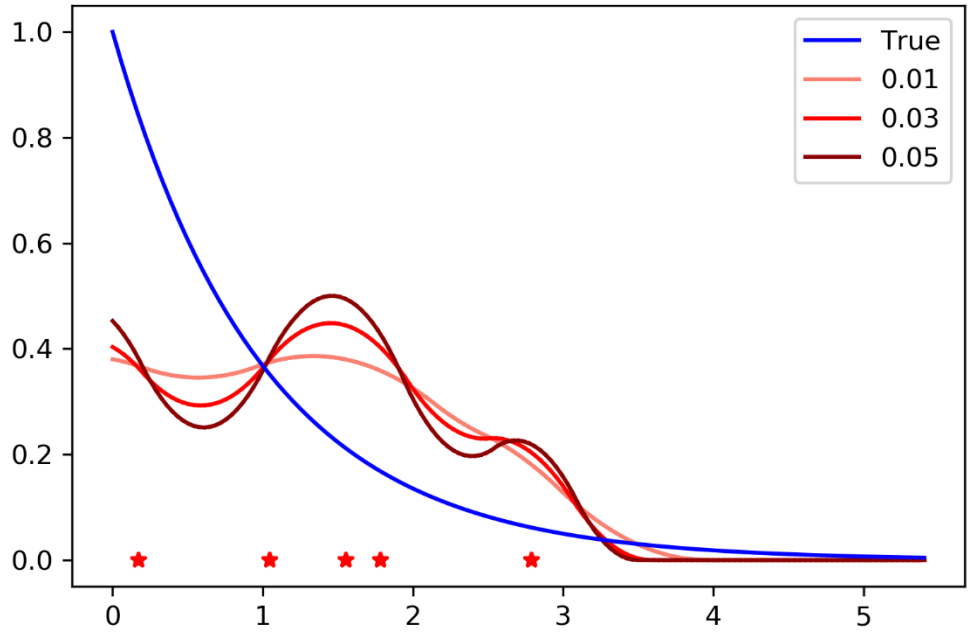
Figure 20. MLE using $\delta = 0.01, 0.03, 0.05,$ with 150 segments.
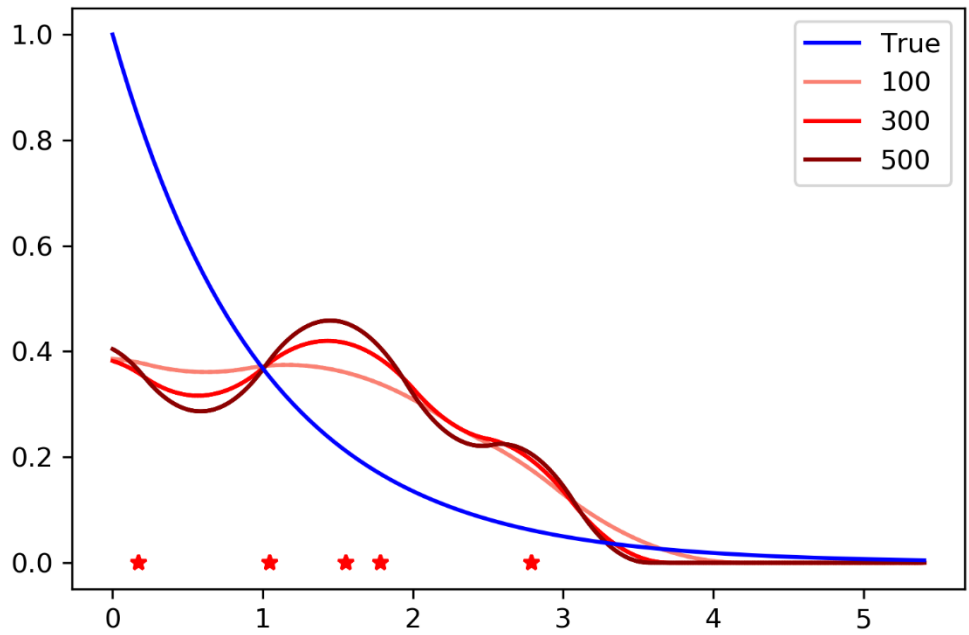
Figure 21. MLE using $N = 100, 300, 500,$ with $\delta = 0.01$.

2. USING THE MFI AND FP-MLE MODELS

We next examine the MFI and the FP-MLE models. Figure 22 shows the resulting estimates using $\varepsilon = 0.01$ and $\rho = 0.001$. The models yield similar results, with the resulting PDFs having heavier tails than those of the MLE. This is an indication of robustness. Specifically, if the 0.9-superquantile values of the estimates from the MFI and the FP-MLE are 3.21 and 3.03, respectively, this is quite a bit higher than the 0.9-superquantile of the MLE estimate, which is 3.01. The two 0.9-superquantile values are not greater than the true 0.9-superquantile value 3.30. This implies that using MFI and FP-MLE for estimating the PDF, which is following exponential distribution, could underestimate outcomes with these robust parameters. However, we can handle this issue by increasing robust parameters ε and ρ. In addition, this is still better estimates than MLE result and parametric result, which is 2.45.
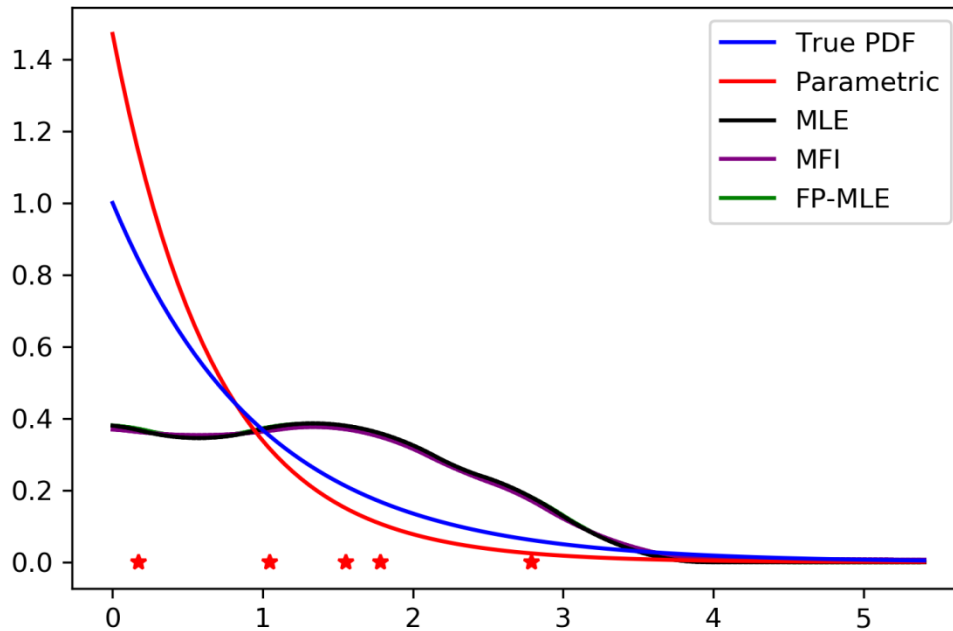
Figure 22. MFI with $\varepsilon = 0.01$ and FP-MLE with $\rho = 0.001$.

The effect of ε is examined in Figure 23. When we increase ε, the MFI produces estimates with increasingly heavy tails, and data in the limit tends toward a uniform

distribution. Table 7 shows the values of the 0.9-superquantile when ε is increased from 0.01 to 0.25. The MFI produces estimates with heavier tails than the true PDF when ε was greater than 0.05.
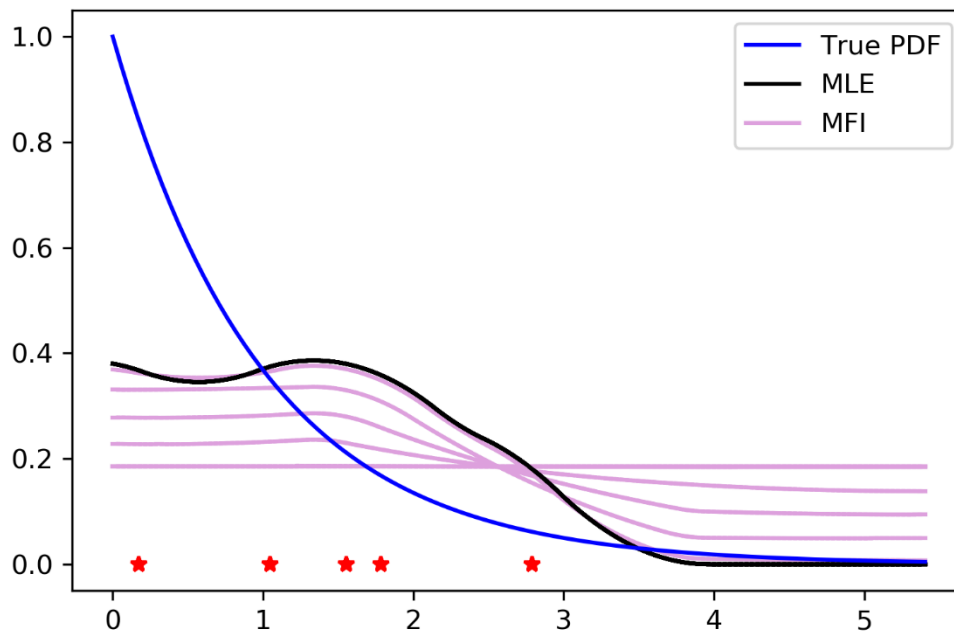
Figure 23.   MFI estimates using $\varepsilon = 0.01, 0.05, 0.10, \ldots, 0.25$.

Table 7.   0.9-superquantiles for MFI estimates using $\varepsilon = 0.01, 0.05, 0.10, \ldots, 0.25$.
The true PDF has 0.9-superquantile of 3.30.

ε	0.01	0.05	0.10	0.15	0.20	0.25
0.9-superquantile	3.21	4.43	4.91	5.08	5.17	5.17

Figure 24 examines changes in ρ. When we increase ρ, the FP-MLE becomes more robust and produces PDFs with heavier tails. Table 8 shows the 0.9-superquantiles when ρ is increased from 0.001 to 0.06. Again, for all values of ρ examined, the FP-MLE produces estimates with heavier tails than the true PDF when ρ was greater than 0.010.
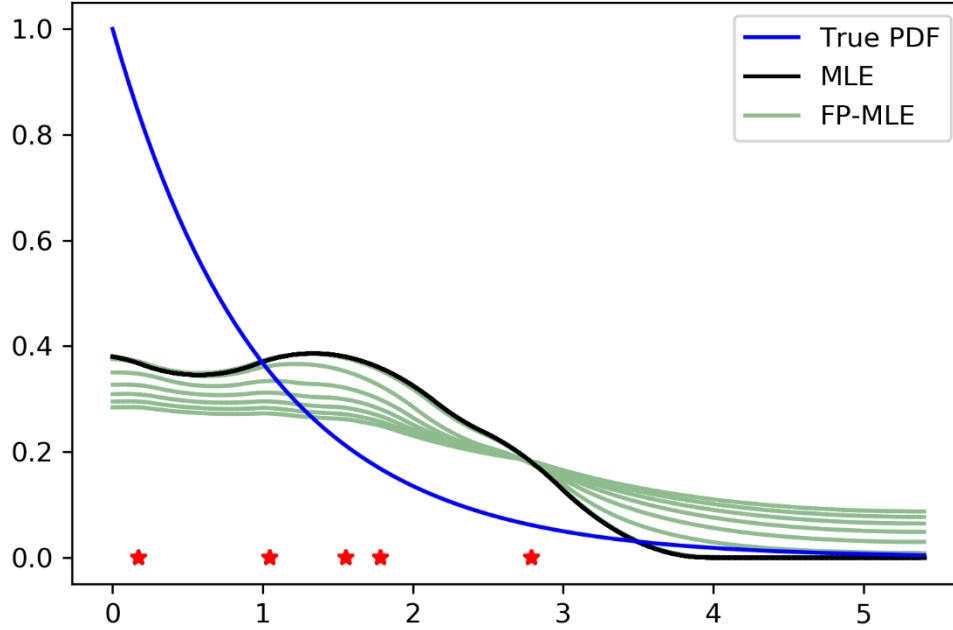
Figure 24.   FP-MLE estimates using
$\rho = 0.001, 0.010, 0.020, \ldots, 0.060$.

Table 8.   0.9-superquantiles for FP-MLE using $\rho = 0.001, 0.010, 0.020, \ldots, 0.060$.
The true PDF has 0.9-superquantile of 3.30.

ρ	0.001	0.010	0.020	0.030	0.040	0.050	0.060
0.9-superquantile	3.03	3.57	4.21	4.53	4.71	4.81	4.88

Table 9 shows the mean and standard deviation of the 0.9-superquantile values obtained by each method and the number of failures—less than the true 0.9-superquantile value 3.30. One thousand simulations were conducted. When we compare the proportion of failures of the parametric estimate with MFI (ε is 0.20) and FP-MLE (ρ is 0.030), MFI and FP-MLE results are less than parametric result. Despite this, 4.11, the mean of the 0.9-superquantile of the parametric estimate, is greater than 3.74, the mean of the 0.9-superquantile of the MFI, and 3.59, the mean of the 0.9-superquantile of the FP-MLE. Because the standard deviation of parametric estimate is higher than those of MFI and FP-MLE. This implies that the MFI and FP-MLE give us better results than parametric estimate.

Table 9. The means and standard deviations of 0.9-superquantile values.

Method	$\varepsilon \mid \rho$	Mean ± standard deviation	Proportion of failure
True	-	3.30	-
Parametric	-	4.11 ± 2.33	43.7%
MLE	-	2.57 ± 1.23	79.9%
MFI	0.01	2.68 ± 1.28	76.3%
	0.05	3.23 ± 1.54	56.3%
	0.10	3.45 ± 1.56	48.5%
	0.15	3.62 ± 1.66	44.5%
	0.20	3.74 ± 1.75	43.3%
	0.25	3.83 ± 1.84	42.7%
	0.30	3.91 ± 1.94	41.9%
FP-MLE	0.001	2.60 ± 1.20	79.7%
	0.005	2.82 ± 1.13	76.1%
	0.010	3.09 ± 1.15	61.9%
	0.015	3.27 ± 1.22	50.0%
	0.020	3.41 ± 1.30	46.2%
	0.025	3.51 ± 1.37	44.2%
	0.030	3.59 ± 1.44	42.9%

IV.   ROBUST ESTIMATION IN AUTONOMY

A.   SCENARIO

Suppose that a commander must decide whether to task an infiltrating unmanned underwater vehicle (UUV) to attack an enemy's aircraft carrier. If a UUV is detected and detained by the enemy during penetration, it would be a big loss for friendly forces because the enemy could learn some things about the friendly UUV's technology. Therefore, the UUV will have to penetrate where it is least likely to be detected by the enemy, and it is better not to infiltrate at all if the probability of being detected by the enemy is high. As a matter of course, a friendly force conducts various reconnaissance and intelligence operations to determine the enemy's defense capabilities relative to their aircraft carrier allowing friendly force to grasp these activities to some extent. This information, however, is usually highly uncertain, and a UUV receiving this information will have to make an assessment whether infiltration is simply too risky.

For robust path planning, two aspects must be considered. First, the UUV must utilize an optimization method to calculate the minimum risk path. Second, it must use a robust estimation method that takes into account the uncertainty of the collected data to assess whether the infiltration mission is viable. The first aspect, the study on the shortest path or the minimum risk path, has been performed extensively; see for example [32, 33]. Thus, we here concentrate on the second aspect using the model that this thesis has presented.

In the search theory [34], it is common to assume that the probability of detecting an object in an area of size A is given by

$$P[\text{being detected}] = 1 - e^{\frac{-VWt}{A}}, \tag{8}$$

where V is the speed of the searcher, W is the detection range of the searcher, and t is the search time. We use this simple model to illustrate the methodology for robust estimation, but any other model is just as easily incorporated.

B. TWO-STEP ANALYSIS

The analysis process that we propose is largely divided into two steps. Step A makes decisions using past data or simulation data. Step B acquires current information to make a decision. Step B can be dropped if Step A is deemed sufficient. It is possible to determine whether to perform Step B by considering the risk of a mission as assessed by Step A and the cost and time to obtain current information.

1. Step A

In most cases, there will be no exact information about the four variables (speed, detection range, search time, and search area) in (8). If correct information were available for the four variables, it is possible to obtain the probability that the UUV will be detected by substituting the four values into (8). In this scenario, however, the information on all four variables is uncertain or partially uncertain. Suppose that the UUV has the following information: the speed of the enemy searcher is 150–250 [knots/hour], its detection range is 20–40 [nm], the area of search is 10,000–50,000 [nm^2], and the time it takes for UUV to pass through the search area is 1–3 [hours]. Using these ranges, Monte Carlo simulations provide a set of *possible* probabilities of detection. We use a sample size of 1,000 in these simulations, and thus generate 1,000 possible probabilities of detection.

These simulations can be viewed as data drawn from an unknown PDF, which we can estimate robustly by using the two models MFI and FP-MLE of Chapter III. The UUV would select the parameters ε and ρ based on prior experience and level of risk averseness. We use 0.1 (low), 0.3 (mid), 0.5 (high) as the value of ε and 0.001 (low), 0.005 (mid), 0.01 (high) as the value of ρ in the following. Figure 25 shows the estimated PDF of the probability of detection using the 1,000 data points generated through simulations; the black, purple, and green lines correspond to MLE, MFI $(\varepsilon = 0.1)$, and FP-MLE $(\rho = 0.001)$, respectively. (MLE uses simulation data points with slope change parameter $\delta = 1$ and 150 segments throughout.) We note that the MFI and FP-MLE smooth out the humps seen in the estimate using the MLE. Moreover, the MFI and FP-MLE result in heavier tails.
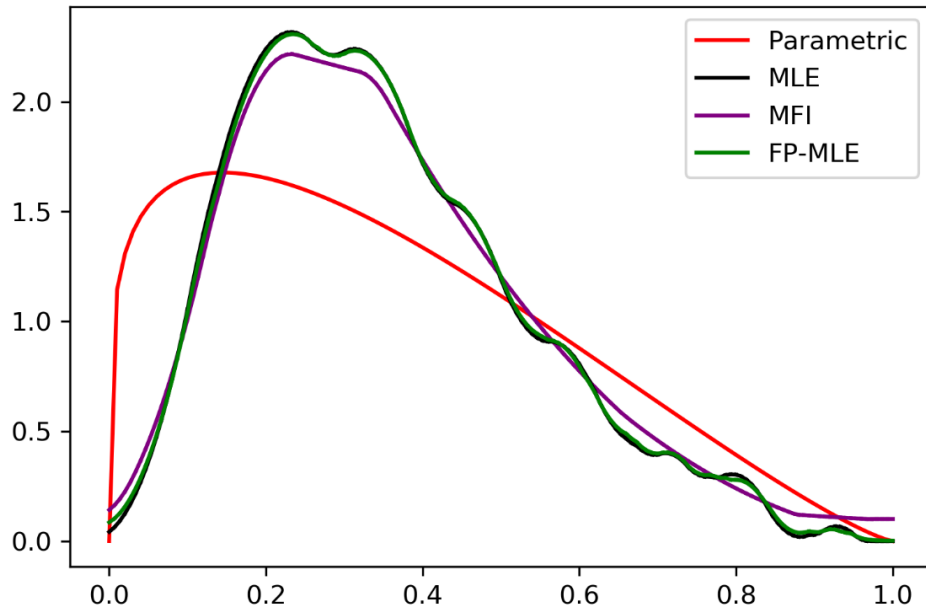
Figure 25.   PDF representing the UUV's belief about probability of detection during infiltration mission. Parametric (red), MLE (black), MFI (purple), and FP-MLE (green).

Table 10 quantifies this effect in terms of superquantiles. Even though we do not know what the true value is, the MLE is close to the true value because we use the data of 1,000 simulations. We can see that parametric estimate using beta distribution have the same mean value as the MLE, but possess higher superquantile values than the superquantile values of the MLE. This means that parametric estimate overestimates the superquantile values because of the heavier tail, as shown in Figure 25. On the other hand, the superquantile values of the FP-MLE are less than the superquantile values of the parametric and close to the MLE results.

This problem combines three optimization problems; MLE, MFI, and FP-MLE. The runtime to solve this problem is approximately 1.3 seconds.

Table 10. The probability of UUV detected α -superquantile values.

Method	$\varepsilon \mid \rho$	α			
		0.00	0.70	0.90	0.95
Parametric	-	0.35	0.64	0.78	0.84
MLE	-	0.35	0.58	0.71	0.78
MFI	0.1	0.36	0.60	0.77	0.85
	0.3	0.39	0.68	0.86	0.93
	0.5	0.42	0.74	0.90	0.95
FP-MLE	0.001	0.35	0.58	0.71	0.78
	0.005	0.35	0.58	0.72	0.79
	0.010	0.35	0.58	0.73	0.80

2.   Step B

In Step B, we assumed that the probability value is calculated through the values of the variables obtained by actual intelligence operations, and the 0.9-superquantile value is estimated by using the probability value. We looked at how the superquantile values change when 5, 10, and 30 data are collected, and then how the superquantile values change when one of the four variables is obtained.

Figures 26–28 represent the estimated probability distribution according to the size of the collected data with low, medium, and high robust parameters. The interesting thing was that there was no change in the robust parameter ε and ρ values, but the FP-MLE showed larger fluctuations than the MFI depending on the amount of data. This implies that the FP-MLE method could be too conservative when compared to the MFI with a small amount of data in a small sample space. In Chapter III, the MFI always yielded more conservative results than the FP-MLE did. In this circumstance, however, the FP-MLE is always conservative until the sample size is large enough. In other words, we should consider both sample size and space when working to find the proper robust parameter ρ. In these results, the low conservative level is better than the other conservative levels for the FP-MLE. In addition, when the data size becomes 30, the value of the 0.9-superquantile with a low conservative level is the closest value among all results.
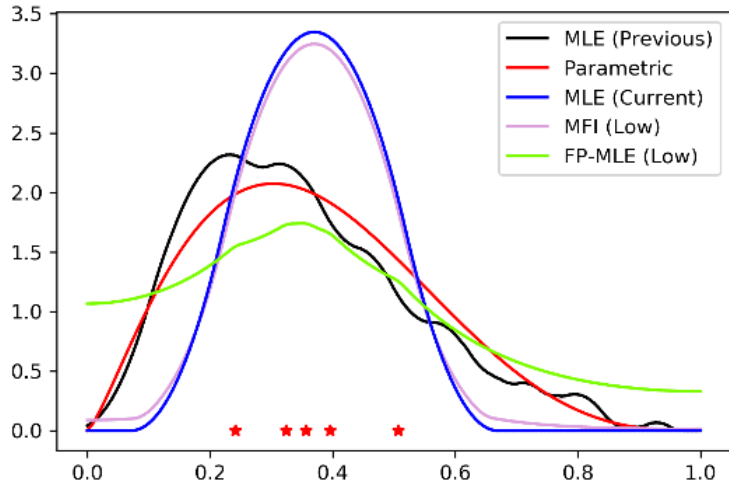
(a) Low conservative
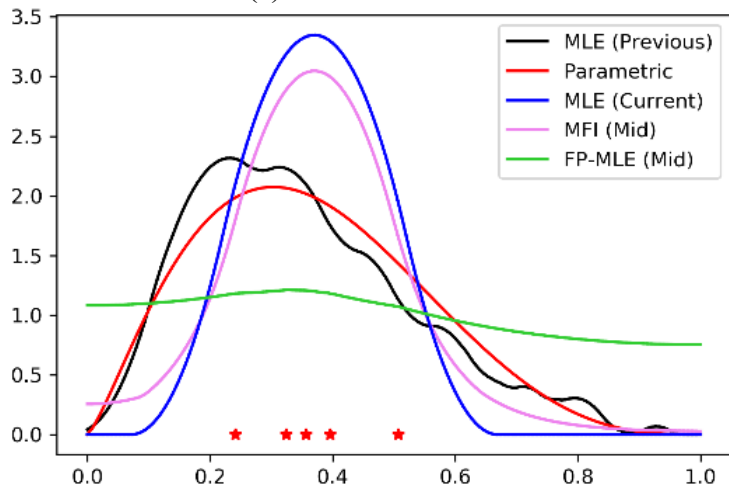
(a) Mid conservative
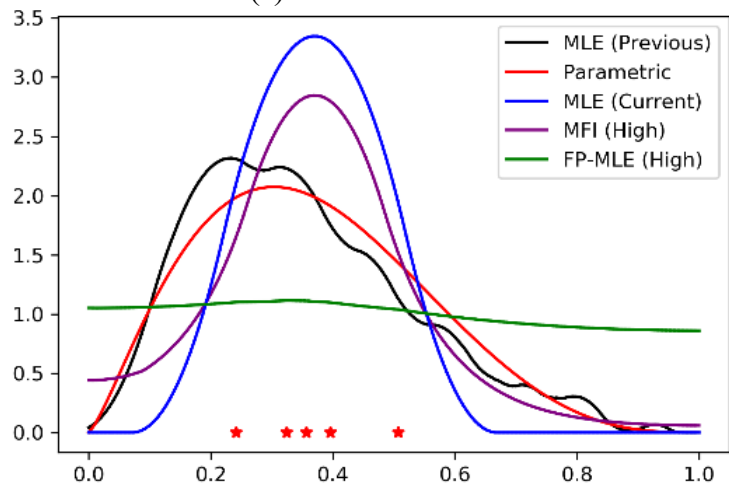
(c) High conservative

Figure 26.   Probability of UUV detected using five data points.
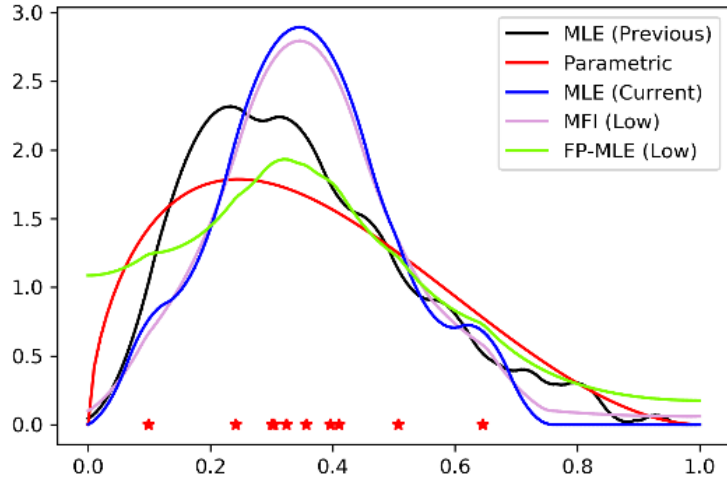
(a) Low conservative
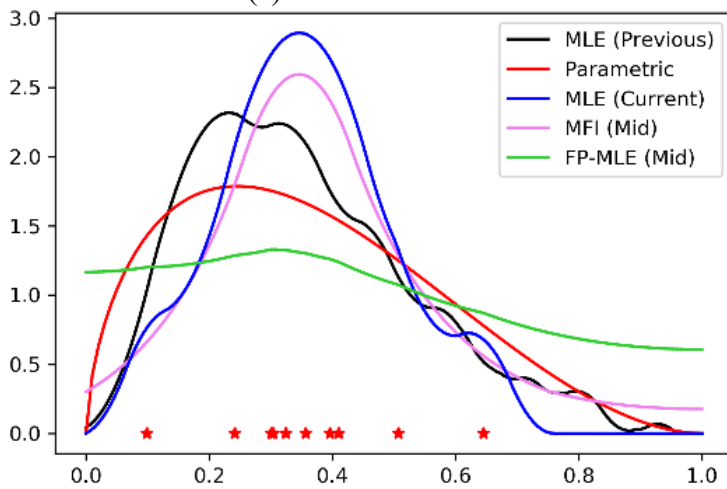
(b) Mid conservative

(c) High conservative

Figure 27.   Probability of UUV detected using 10 data points.
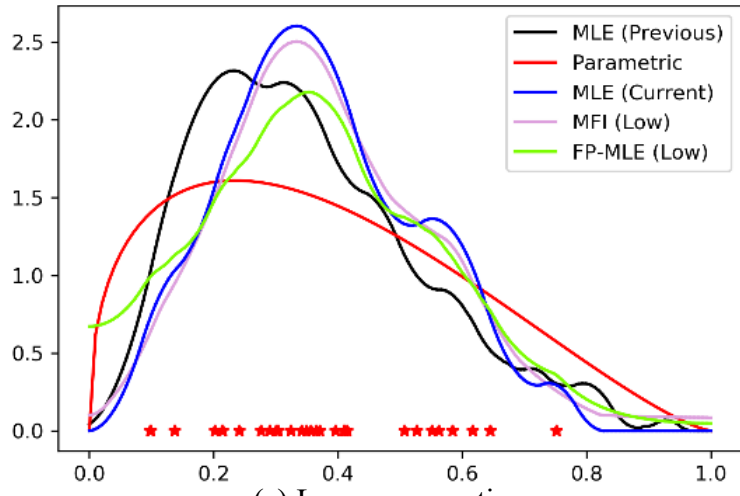
(a) Low conservative
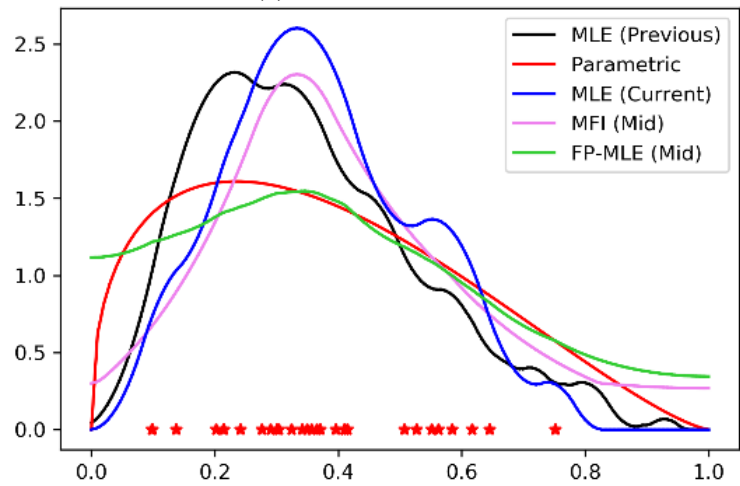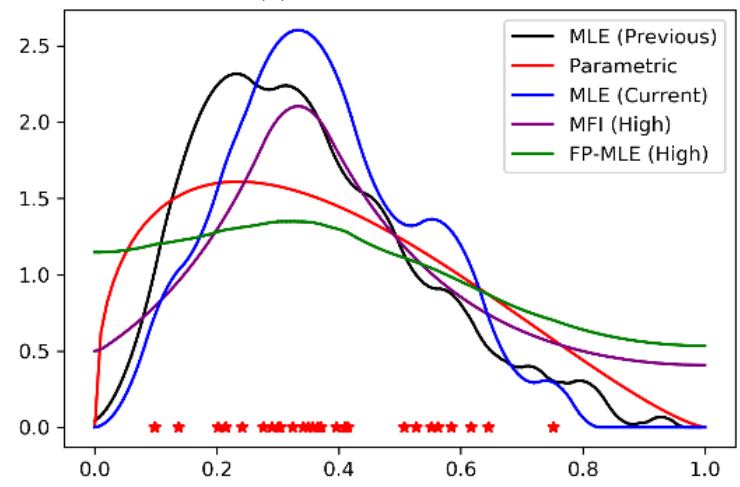
(b) Mid conservative

(c) High conservative

Figure 28.   Probability of UUV detected using 30 data points.

39

Table 11 shows the superquantile value according to the change of the α value with a different data size. Parametric estimate seems to relate to data quality. The 0.9-superquantile value of 10 data points 0.74, is closer to 0.71, MLE estimates with 1,000 data points, than the 0.9-superquantile value of 30 data points 0.49. On the other hand, when data points are 30, the mean and the 0.9-superquantile values of the MF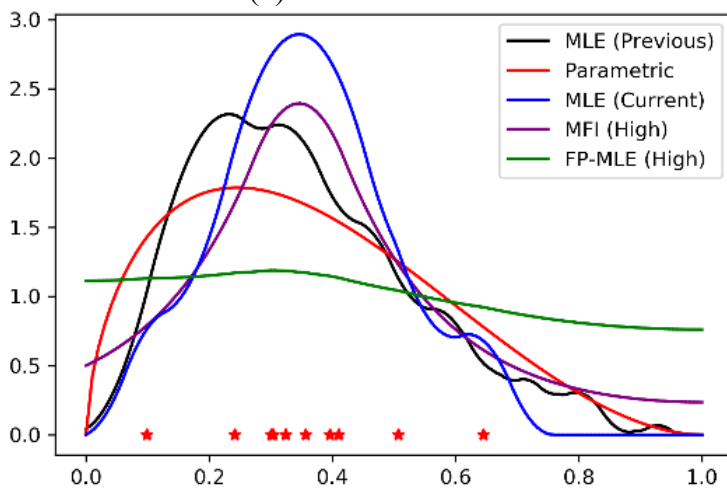I and FP-MLE are closer to the case with 1,000 data points. The superquantile value of MLE was always smaller than the superquantile values of MFI and FP-MLE. This means that both MFI and FP-MLE have heavier tails than the MLE does. In the case of 30 data points, the MLE overestimated the mean as 0.38, which is greater than 0.35 MLE estimates with 1,000 data points. By contrast, the MLE underestimated the 0.9-superquantile value as 0.67, which is smaller than 0.71 MLE estimates with 1,000 data points.

Table 11. The probability of UUV detected α -superquantile values using different data sizes.

Size	Method	$\varepsilon \mid \rho$	α			
			0.00	0.70	0.90	0.95
1000	MLE	-	0.35	0.58	0.71	0.78
5	Parametric	-	0.37	0.58	0.70	0.75
	MLE	-	0.38	0.50	0.56	0.58
	MFI	0.1	0.38	0.51	0.58	0.63
		0.3	0.38	0.53	0.64	0.71
		0.5	0.38	0.56	0.69	0.77
	FP-MLE	0.001	0.39	0.69	0.87	0.93
		0.005	0.47	0.81	0.94	0.97
		0.010	0.49	0.84	0.95	0.98
10	Parametric	-	0.36	0.61	0.74	0.80
	MLE	-	0.36	0.54	0.63	0.67
	MFI	0.1	0.37	0.55	0.67	0.76
		0.3	0.38	0.61	0.79	0.88
		0.5	0.39	0.65	0.83	0.91
	FP-MLE	0.001	0.36	0.63	0.80	0.88
		0.005	0.44	0.78	0.93	0.97
		0.010	0.47	0.82	0.94	0.97

Size	Method	$\varepsilon \mid \rho$	α			
			0.00	0.70	0.90	0.95
30	Parametric	-	0.38	0.65	0.79	0.84
	MLE	-	0.38	0.57	0.67	0.71
	MFI	0.1	0.39	0.59	0.72	0.80
		0.3	0.41	0.67	0.84	0.92
		0.5	0.43	0.72	0.89	0.95
	FP-MLE	0.001	0.38	0.60	0.73	0.80
		0.005	0.40	0.71	0.88	0.94

According to the change of sample size, Table 12 shows the mean and standard deviation of the 0.9-superquantile values obtained by each method and the number of failures—less than 0.71 which is the MLE estimates 0.9-superquantile value with 1,000 data points. One thousand simulations were conducted. When the data are small (5 and 10), the MFI and FP-MLE results are closer to 0.71. However, when the data becomes sufficiently large (30), the MLE estimate approaches 0.71 and the MFI and FP-MLE values become more conservative. In conclusion, the MFI and FP-MLE show stable results when collected data have less information about true value.

Table 12. The means and standard deviations of 0.9-superquantile values.

Size	Method	$\varepsilon \mid \rho$	Mean ± standard deviation	Proportion of failure
1000	MLE	-	0.71	-
5	Parametric	-	0.78 ± 0.15	31.9%
	MLE	-	0.64 ± 0.14	69.5%
	MFI	0.1	0.68 ± 0.13	60.9%
		0.3	0.77 ± 0.10	26.9%
		0.5	0.83 ± 0.09	16.7%
	FP-MLE	0.001	0.88 ± 0.04	00.0%
		0.005	0.94 ± 0.01	00.0%
		0.010	0.95 ± 0.003	00.0%
10	Parametric	-	0.78 ± 0.10	25.2%

Size	Method	$\varepsilon \mid \rho$	Mean ± standard deviation	Proportion of failure
	MLE	-	0.68 ± 0.11	63.3%
	MFI	0.1	0.71 ± 0.09	49.5%
		0.3	0.81 ± 0.07	10.8%
		0.5	0.85 ± 0.06	04.9%
	FP-MLE	0.001	0.82 ± 0.07	07.7%
		0.005	0.93 ± 0.01	00.0%
		0.010	0.94 ± 0.01	00.0%
30	Parametric	-	0.78 ± 0.06	15.5%
	MLE	-	0.71 ± 0.07	53.2%
	MFI	0.1	0.73 ± 0.05	34.9%
		0.3	0.82 ± 0.04	01.9%
		0.5	0.87 ± 0.04	00.3%
	FP-MLE	0.001	0.75 ± 0.07	30.5%
		0.005	0.87 ± 0.02	00.0%
		0.010	0.91 ± 0.01	00.0%

Suppose that in some of the four variables a deterministic value is obtained through an additional data collection activity to see how the value changes. First, the search area that had the largest variance is found to be 25,000 [nm^2]. As with Step A, the other values are calculated by 1,000 simulations and are shown in Table 13. Compared to the previous results, we can see that the mean increases from 0.35 to 0.37 or 0.38, and the value of 0.9-superquantile decreases from 0.72 to 0.64 in the MLE, to 0.71 in the MFI ($\varepsilon = 0.1$), and 0.64 in the FP-MLE ($\rho = 0.001$).

This result is interesting because the expected value of the search area is 30,000 [nm^2]. This means that 25,000 [nm^2] is worse than our expected value. The 0.9-superquantile value, however, is decreased because of the removal of uncertainty in the search area. This implies that not only we can estimate much about the data region, but also we can estimate more and more accurately about the tail region. In Figure 29, we can

confirm that the blue current MLE is moved to the right (worse), but the shape is more centered at the mean and the tail becomes lighter than the MLE tail.

Table 13. The probability of UUV detected α-superquantile values using search area is 25,000 [nm^2].

Method	$\varepsilon \mid \rho$	α			
		0.00	0.70	0.90	0.95
Parametric	-	0.37	0.63	0.76	0.82
MLE	-	0.37	0.55	0.64	0.68
MFI	0.1	0.38	0.58	0.71	0.79
	0.3	0.40	0.66	0.83	0.91
	0.5	0.42	0.72	0.89	0.95
FP-MLE	0.001	0.37	0.55	0.64	0.68
	0.005	0.37	0.55	0.65	0.69
	0.010	0.37	0.56	0.65	0.70
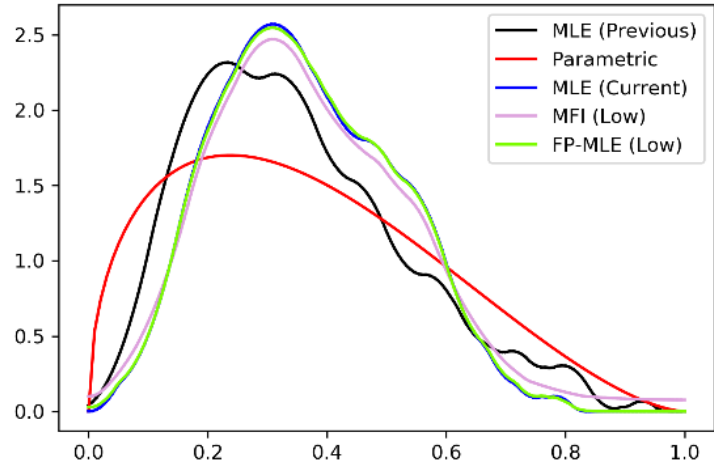
(a) Low conservative
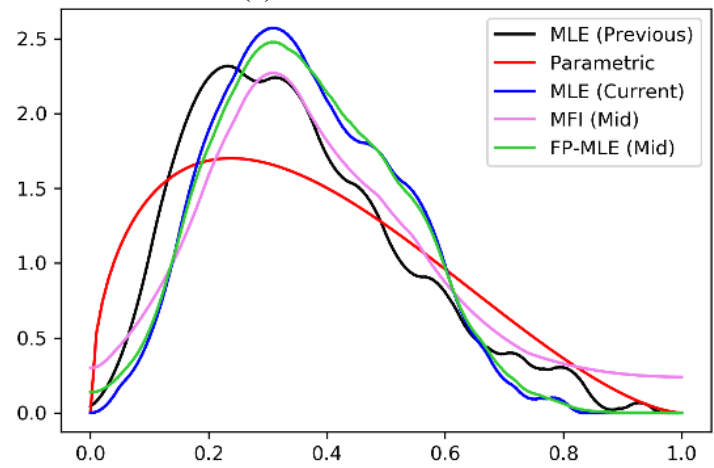
(b) Mid conservative

(c) High conservative

Figure 29.   Probability of UUV detected ($A = 25{,}000$ [nm^2]).

Table 14 shows the results obtained by repeating the previous procedure, assuming that the other values have also acquired deterministic values (expected value) one by one. In all three cases, we can see that the 0.9-superquantile values are not reduced, except for the V and W cases in the FP-MLE. Since the three cases used the expected value, this means it was a more beneficial to the friendly force, but the 0.9-superquantile value did not decrease significantly. In Figures 30–32, we can confirm that the blue current MLE is moved to the left (better) and more centered at the mean, but the tail is still heavy.

Comparing the amount of shrinkage, we observe that the uncertainty of the variables was reduced the most when the biggest variance value was found. In conclusion, the best solution to reduce the risk of detection is to obtain information on the enemy's search area where the highest degree of uncertainty exists.
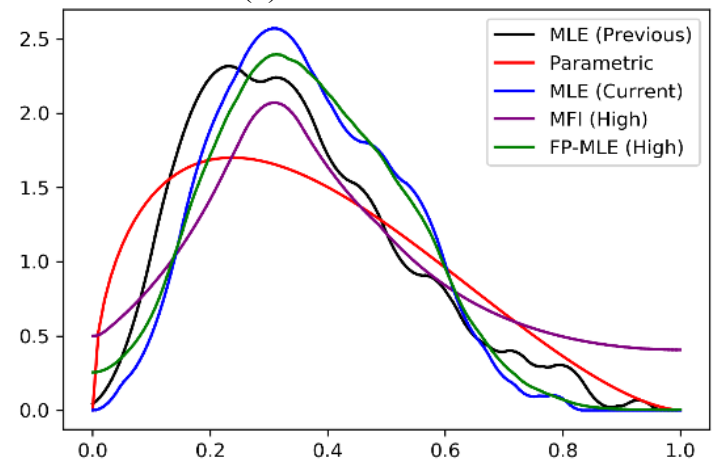
Table 14. The probability of UUV detected α-superquantile values using one deterministic value.

Variable	Method	$\varepsilon \mid \rho$	α			
			0.00	0.70	0.90	0.95
$V = 200$	Parametric	-	0.35	0.62	0.76	0.82
	MLE	-	0.35	0.56	0.70	0.76
	MFI	0.1	0.36	0.58	0.74	0.83
		0.3	0.38	0.66	0.85	0.92
		0.5	0.40	0.71	0.88	0.94
	FP-MLE	0.001	0.35	0.56	0.70	0.76
		0.005	0.35	0.57	0.70	0.77
		0.010	0.35	0.57	0.71	0.78
$W = 30$	Parametric	-	0.35	0.62	0.75	0.81
	MLE	-	0.36	0.55	0.68	0.74
	MFI	0.1	0.36	0.58	0.74	0.83
		0.3	0.38	0.64	0.83	0.91
		0.5	0.39	0.68	0.86	0.93
	FP-MLE	0.001	0.36	0.56	0.69	0.75

Variable	Method	$\varepsilon \mid \rho$	α			
			0.00	0.70	0.90	0.95
		0.005	0.36	0.56	0.69	0.75
		0.010	0.36	0.57	0.70	0.76
$t = 2$	Parametric	-	0.36	0.64	0.78	0.84
	MLE	-	0.36	0.58	0.72	0.78
	MFI	0.1	0.37	0.60	0.76	0.84
		0.3	0.38	0.66	0.85	0.92
		0.5	0.39	0.69	0.87	0.94
	FP-MLE	0.001	0.36	0.58	0.72	0.78
		0.005	0.37	0.59	0.72	0.78
		0.010	0.37	0.59	0.73	0.79
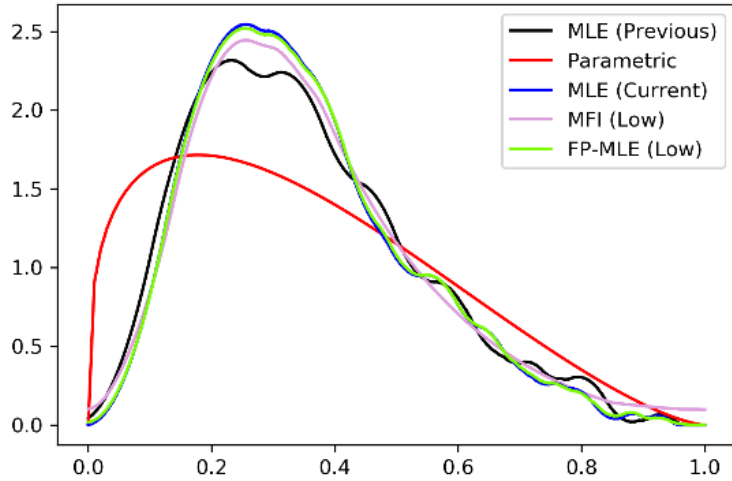
(a) Low conservative
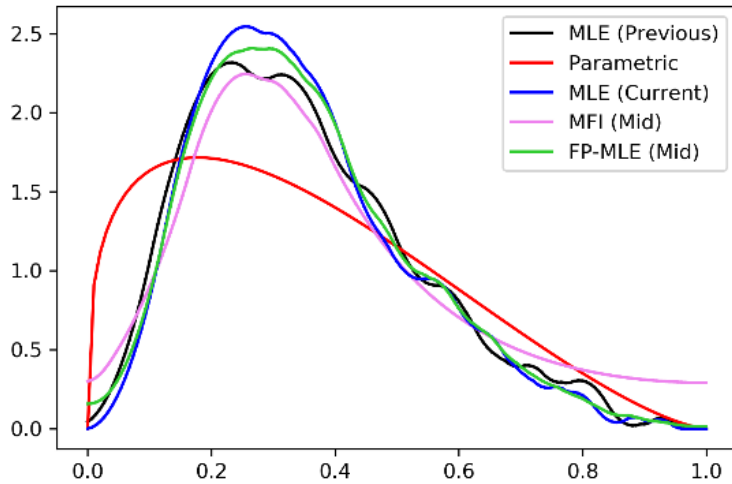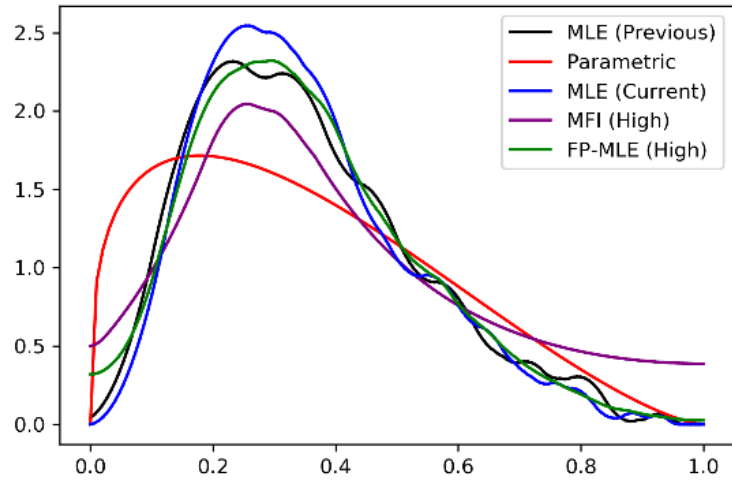
(b) Mid conservative

(c) High conservative

Figure 30.   Probability of UUV detected (V = 200 [knots/hour]).
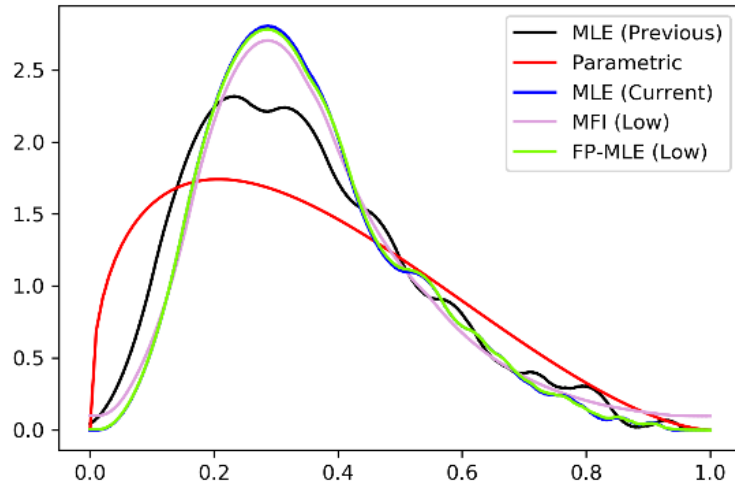
(a) Low conservative
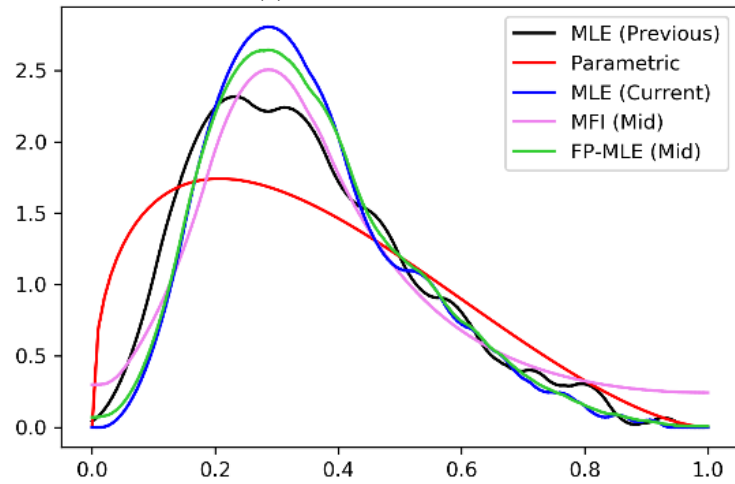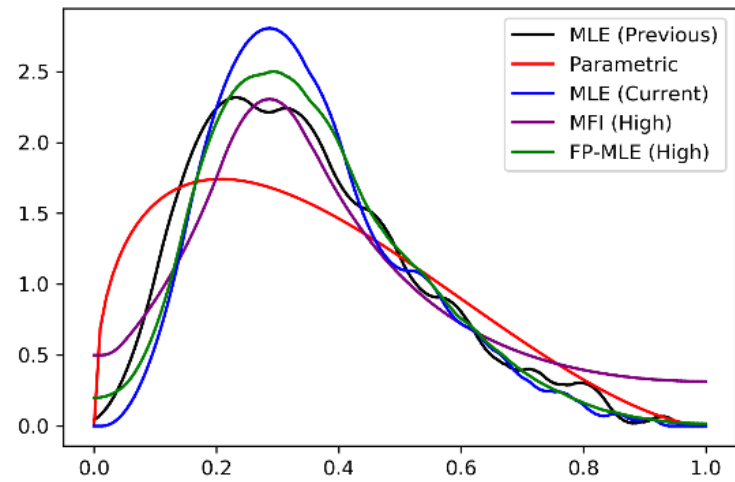
(b) Mid conservative

(c) High conservative

Figure 31. Probability of UUV detected (W = 30 [nm]).

48
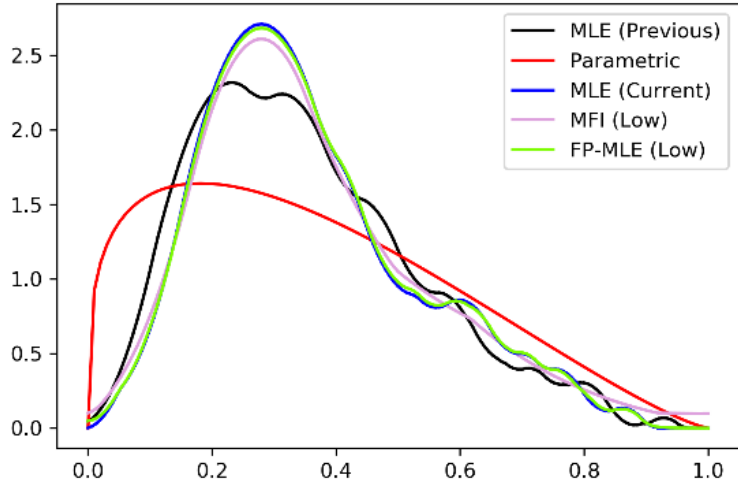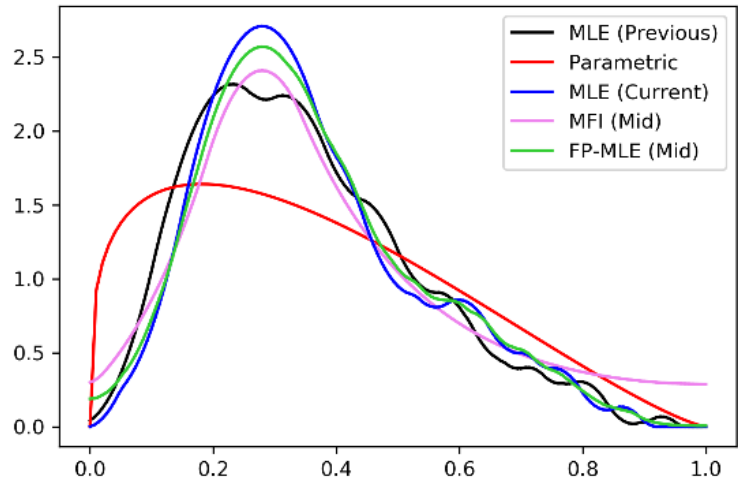
(a) Low conservative
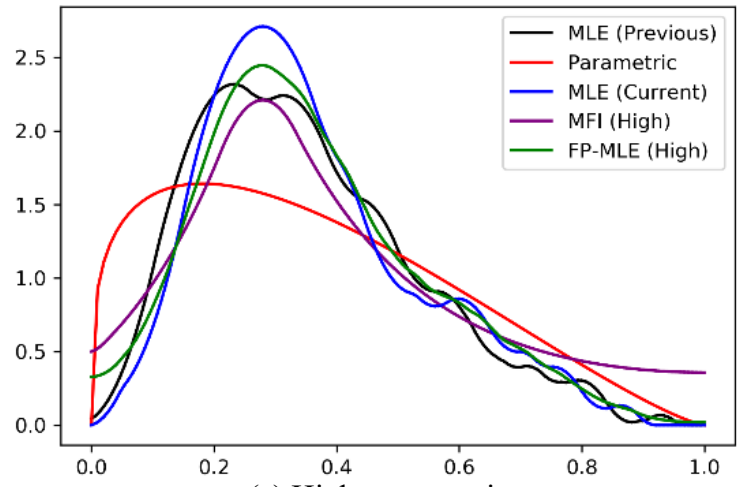
(b) Mid conservative

(c) High conservative

Figure 32.   Probability of UUV detected ($t = 2$ [hours]).

Applying this strategy to the defense decision-making procedure, we can confirm that the most volatile information among the variables predicted at present is the priority information. When we get the priority information, we can reduce the risk although the priority information might not be "good" news. In the game Starcraft, for example, it is much more beneficial to identify the opponent's position and initial strategy through reconnaissance rather than to have the worker, who is an important unit, capture resources early on. Even though the worker cannot collect resources after being killed by the enemy, the worker instead gives us vital information for reducing uncertainty.

V. CONCLUSIONS

The use of autonomous systems is likely to increase in the future, and the level of autonomy in these systems will also be more sophisticated than it is now. Increasing autonomy means that human intervention will be less important and that the autonomous system will be able to assess a situation and decide whether to continue a mission. When making decisions based on the collected information, it is necessary to calculate the risk and to reflect on whether the mission should be carried out. In particular, autonomous systems might want to make conservative decisions when the quantity and quality of the information collected is insufficient.

In the Chapter II, we develop two models (Minimizing Fisher Information, or MFI, and Fisher-Penalized Maximum likelihood, or FP-MLE) for estimating probability density functions robustly using a small amount of data. Both models rely on likelihoods and epi-splines and find density functions with minimal Fisher information. Both are convex models solved in less than one second, even with a data size of 100,000. However, they are distinct in that the first uses a constraint on the proximity to a candidate density function and the second leverages a penalty formulation.

In the Chapter III, we study benchmark problem instances with data from normal, log-normal, and exponential distributions and robustly estimate these density functions. Even though the sample size is only five, we confirm that both models give more stable estimates in a certain sense than parametric estimation or MLE. For example, the standard errors and the proportion of failure of the two models are overall better than those of parametric estimation or maximum likelihood estimation. Also, by adjusting the robustness parameter, the user can set the level of conservativeness as desired.

In the Chapter IV, we also examine the possibility of using the models for decision making of autonomous systems. We estimate the probability that an underwater unmanned vehicle (UUV) can penetrate an area of interest undetected. Using as little as five, 10, and 30 data points, obtained, for example, through reconnaissance or intelligence, the models yield conservative estimates of the probability that the UUV will be detected during the

penetration. The results of this study have the potential to guide autonomous systems to reflect more deeply on the risks associated with a mission, help them decide whether additional information is required, or whether the mission should simply be abandoned. The study also provides the insight that more information can reduce the risk associated with the mission, even when that information is "bad news" in the sense that the adversary has greater capability than originally expected. For example, if there is a high level of uncertainty about the number of enemy tanks in an operation, a plan should be established with an emphasis on the amount of corresponding power. On the other hand, if the high level of uncertainty is about the attack time of enemy tanks, a plan should be established with an emphasis on the reaction time.

Although both the MFI and the FP-MLE models show robust results, the choice between the MFI and the FP-MLE for robust estimating remains an ambiguous issue with no clear answer. This is an exploratory thesis; the best method will depend on the context in which it is applied. MFI tends to approximate the uniform distribution within the range allowed. Even though FP-MLE also tends to approach the uniform distribution, FP-MLE is more related to data size and support than for MFI. The FP-MLE model has a tendency to be overly conservative if there is little data or the support is small.

Determining the support and choosing parameters remains an issue throughout this thesis. In this thesis, we designate a reasonable range for the support and use it consistently. However, other possibilities could be examined. We also leave for future studies the extension of the approach to higher dimensions.

APPENDIX. GAUSSIAN QUADRATURE ROLE

Number of points	Points	Weights
1	0	2
2	$\pm\dfrac{1}{\sqrt{3}}$	1
3	0	$\dfrac{8}{9}$
	$\pm\sqrt{\dfrac{3}{5}}$	$\dfrac{5}{9}$
4	$\pm\sqrt{\dfrac{3}{7}-\dfrac{2}{7}\sqrt{\dfrac{6}{5}}}$	$\dfrac{18+\sqrt{30}}{36}$
	$\pm\sqrt{\dfrac{3}{7}+\dfrac{2}{7}\sqrt{\dfrac{6}{5}}}$	$\dfrac{18-\sqrt{30}}{36}$
5	0	$\dfrac{128}{225}$
	$\pm\dfrac{1}{3}\sqrt{5-2\sqrt{\dfrac{10}{7}}}$	$\dfrac{322+13\sqrt{70}}{900}$
	$\pm\dfrac{1}{3}\sqrt{5+2\sqrt{\dfrac{10}{7}}}$	$\dfrac{322-13\sqrt{70}}{900}$

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

[1]   H. Kim, "S. Korea's military creates drone combat unit," *Yonhap News*, Sept. 28, 2018. [Online]. Available: https://www.yna.co.kr/view/AKR20180928047300014?input=1195m

[2]   N. Kalra and D. G. Groves, "The enemy of good: Estimating the cost of waiting for nearly perfect automated vehicles," RAND Corp., Santa Monica, California, USA, Rep. RR-2150-RC, Jan. 18, 2019. [Online]. DOI: 10.7249/RR2150 or Available: https://www.rand.org/pubs/research_reports/RR2150.html

[3]   M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: Wiley-Interscience, 2005.

[4]   A. Shapiro and D. Dentcheva, *Lectures on Stochastic Programming: Modeling and Theory*, Second Edition. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2012.

[5]   W. Powell, *Approximate Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2013.

[6]   R. T. Rockafellar and J. O. Royset, "Engineering decisions under risk averseness," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, vol. 1, no. 2, p. 04015003, 2015.

[7]   J. J. Sabol III. "Dual approach to superquantile estimation and applications to density fitting," M.S. thesis, Dept. of Operations Research, NPS, Monterey, CA, USA, 2016.

[8]   S. Sarykalin, G. Serraino, and S. Uryasev, "Value-at-risk vs. conditional value-at-risk in risk management and optimization," Oct. 14, 2014. [Online]. Available: https://doi.org/10.1287/educ.1080.0052.

[9]   R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, no. 3, pp. 21–42, 2000.

[10]   R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.

[11]   R. T. Rockafellar and J. O. Royset, "On buffered failure probability in design and optimization of structures," *Reliability Engineering & System Safety*, vol. 95, no. 5, pp. 499–510, 2010.

[12]   M. H. DeGroot and M. J. Schervish, *Probability and Statistics*, 4th ed. London, UK: Pearson Education, 2011

[13]   A. Ly, M. Marsman, J. Verhagen, R. P. Grasman, and E.-J. Wagenmakers, "A Tutorial on Fisher information," *Journal of Mathematical Psychology*, vol. 80, pp. 40–55, 2017.

[14]   A. J. Lee and S. H. Jacobson, "Identifying changing aviation threat environments within an adaptive homeland security advisory System," *Risk Analysis*, vol. 32, no. 2, pp. 319–329, 2011.

[15]   F. Zhou and G. Wang, "Node selection algorithm based on Fisher information," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, 2016.

[16]   N. Ahmad, S. Derrible, and H. Cabezas, "Using Fisher information to assess stability in the performance of public transportation systems," *Royal Society Open Science*, vol. 4, no. 4, p. 160920, 2017.

[17]   J. O. Royset and R. J-B Wets, "Fusion of hard and soft information in nonparametric density estimation," *European Journal of Operational Research*, vol. 247, no. 2, pp. 532–547, 2015.

[18]   D. I. Singham, J. O. Royset, and R. J-B Wets, "Density estimation of simulation output using exponential epi-splines," *2013 Winter Simulations Conference (WSC)*, 2013. [Online]. doi: 10.1109/wsc.2013.6721468

[19]   J. O. Royset, 2018, "Slide for OA4021 APP3," class notes for nonlinear programming, Dept. of Operations Research, NPS, Monterey, CA, USA, summer 2018.

[20]   J. O. Royset and R. J-B Wets, "Multivariate Epi-splines and Evolving Function Identification Problems," *Set-Valued and Variational Analysis*, vol. 24, no. 4, pp. 517–545, 2015.

[21]   J. O. Royset and R. J-B Wets, "Erratum to: Multivariate Epi-splines and Evolving Function Identification Problems," *Set-Valued and Variational Analysis*, vol. 24, no. 4, pp. 547–549, 2016.

[22]   B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. London, UK: Chapman and Hall, 1986.

[23]   P. G. Hoel, *Introduction to Mathematical Statistics*. New York, NY, USA: John Wiley & Sons, 1964.

[24]   J. O. Royset and R. J-B Wets, "From data to assessments and decisions: Epi-spline technology," *Bridging Data and Decisions*, pp. 27–53, 2014.

[25]   J. O. Royset and R. J-B Wets, "Variational analysis of constrained M-estimators," arXiv preprint arXiv:1702.08109v4, 2018.

[26] I. C. Demetriou and M. J. D. Powell, "Least squares smoothing of univariate data to achieve piecewise monotonicity," *IMA Journal of Numerical Analysis*, vol. 11, no. 3, pp. 411–432, 1991.

[27] E. Mammen and C. Thomas-Agnan, "Smoothing splines and shape restrictions," *Scandinavian Journal of Statistics*, vol. 26, no. 2, pp. 239–252, 1999.

[28] M. C. Meyer, "Constrained penalized splines," *Canadian Journal of Statistics*, vol. 40, no. 1, pp. 190–206, 2012.

[29] J. O. Royset and R. J-B Wets, "On univariate function identification problems," *Mathematical Programming*, vol. 168, no. 1–2, pp. 449–474, 2018.

[30] R.T. Rockafellar and R. J-B Wets, *Variational Analysis*. Heidelberg, Germany: Springer, 1998.

[31] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, 2005.

[32] W. M. Carlyle, J. O. Royset, and R. K. Wood, "Lagrangian relaxation and enumeration for solving constrained shortest-path problems," *Networks*, vol. 52, no. 4, pp. 256–270, 2008.

[33] J. O. Royset, W. M. Carlyle, and R. K. Wood, "Routing military aircraft with a constrained shortest-path algorithm," *Military Operations Research*, vol. 14, no. 3, 2009.

[34] D. L. Stone, J. O. Royset, and, A. R. Washburn, *Optimal Search for Moving Targets*. Cham, Switzerland: Springer, 2016.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1.   Defense Technical Information Center
        Ft. Belvoir, Virginia

2.   Dudley Knox Library
        Naval Postgraduate School
        Monterey, California