NPS55-86-021

NAVAL POSTGRADUATE SCHOOL

Monterey, California

TECHNICAL

PROBLEMS OF IDENTIFICATION

DONALD P. GAVER
PATRICIA A. JACOBS
I. O'MUIRCHEARTAIGH
A. MELDRUM

NOVEMBER 1986

Prepared for:
National Security Agency
Fort Meade, MD 20755-6000

ffice of Naval Research
rlington, VA 22217

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

Rear Admiral R. C. Austin
Superintendent

D. A. Schrady
Provost

The work reported herein was supported in part by the National Security Agency and the Office of Naval Research.

This report was prepared by:

ICLASSIFIED
CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

SECURITY CLASSIFICATION ASSIFIED	1b. RESTRICTIVE MARKINGS
TY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited
SSIFICATION / DOWNGRADING SCHEDULE	
MING ORGANIZATION REPORT NUMBER(S) 55-86-021	5. MONITORING ORGANIZATION REPORT NUMBER(S)
OF PERFORMING ORGANIZATION al Postgraduate School / 6b. OFFICE SYMBOL (If applicable) Code 55	7a. NAME OF MONITORING ORGANIZATION
SS (City, State, and ZIP Code) terey, CA 93943-5000	7b. ADDRESS (City, State, and ZIP Code)
OF FUNDING / SPONSORING NIZATION National Security - Office of Naval Research / 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER H98230 R0860130 (NSA) N0001 486 WR 24083 (ONR)
SS (City, State, and ZIP Code) Fort Meade, MD 20755-6000 Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO: 61153N; PROJECT NO: RR014-05-01; TASK NO.: 4114531-01; WORK UNIT ACCESSION NO:

(Include Security Classification)
BLEMS OF IDENTIFICATION

ONAL AUTHOR(S)
, Donald P.; Jacobs, Patricia A.; O'Muircheartaigh, I. G.; Meldrum, A.

E OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
ical	FROM ____ TO ____	November 1986	42

LEMENTARY NOTATION

COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
D	GROUP	SUB-GROUP	robust estimation in a restricted parameter space, Cauchy and contaminated normal measurement errors, maximum likelihood and Bayes estimates, the discrimination problem

TRACT (Continue on reverse if necessary and identify by block number)

Several (I in number) observers cooperate to identify the member of a population of s (J in number) transmitting. The observers' measurements of basic item-identifying wn) parameters are error-prone; errors are frequently non-Gaussian (gross-error-prone). ous methods are introduced and studied for combining the observers' observations s to achieve as high a probability of correct identification as possible. Linear inations of observations are shown to be less than satisfactory in the environment ribed. Continuation and generalization of the work is in progress.

TRIBUTION / AVAILABILITY OF ABSTRACT NCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
AME OF RESPONSIBLE INDIVIDUAL d P. Gaver and Patricia A. Jacobs	22b. TELEPHONE (Include Area Code) (408)646-2605/646-2258 / 22c. OFFICE SYMBOL Code 55Gv / 55Jc

RM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

PROBLEMS OF IDENTIFICATION

D. P. GAVER
P. A. JACOBS
I. G. O'MUIRCHEARTAIGH
A. MELDRUM

1. Formulations: Background and Introductory Comments

Envision this abstract situation. There are J items, distinct from one another and bearing names. Think of them as BIRDS of different types. Each one is characterized by a vector of identifying components (you can possibly think of physical characteristics such as color, flight speed, characteristics of song, etc. as these parameters). In addition, it may (or may not) be known where the items are located geographically; they occasionally move, and may move together in suitable flocks.

Next, there is a group of individuals, called WATCHERS, who are sensitive to the parameters (physical characteristics) just mentioned, when the latter become evident. In effect, some bird may sing his song, and one or more of the WATCHERS will hear and record various features of the song, but with error; the same with other features or parameters. Observations are made "in the dark:" observing WATCHERS cannot see the BIRDS before the song and other qualities become evident. In fact, the objective of the group of individuals is to collectively identify the BIRDS in question as well as possible just by comparing notes on the parameter announcements, e.g. song and other feature characteristics.

Errors of various types are easily made, depending upon the operating characteristics of the WATCHERS and on the distribution of the parameters.

Perhaps a group of BIRDS will all sing, and two WATCHERS will confuse two or more BIRDS whose songs they hear: both will state their estimates of "the" song length of what they take to be a single BIRD, whereas in fact, they have focused on two different BIRDS with similar-enough songs.

If the individual assessments are error-prone (as they will be) and if the distribution of the vector parameters is unfortunate, being tightly concentrated around a point in p-space (parameter space) the advantage of the WATCHERS is minimal: they will be unable to accurately discern a particular BIRD's presence, much less how many BIRDS are singing. If several WATCHERS are responding to two different BIRDS, their composite single assessment of the parameter may fail to conform to anything real. The general problem is to identify singing BIRDS using error-prone and even gross error (outlier) prone observations.

With this as background we begin to formulate a variety of simple problems and to consider their implications.

2. The Single Item - One Parameter Case

To get started, focus on the information available from an announcement (song) by a single BIRD. Call the parameter value Θ, and focus first on estimating Θ from observations made by I WATCHERS on announcement of Θ. Now it may actually be known that if

$$\Theta = \mu_j \qquad j=1,2,\ldots J \tag{2.1}$$

the BIRD is the j^{th} of a group, and is named George; if the estimated Θ, $\hat{\Theta}$, actually is very near μ_j, then we announce confidently that we have heard an announcement by George the robin ("by George, I think I heard him"). Things might not be quite so simple: the actual parameter announced may be distributed somewhere near the value μ_j. Increasing the spread or variability of announced values of Θ around μ_j will be confusing to each WATCHER, and the announcement that we indeed heard George himself becomes less likely to be true.

2.1 A First Step: Likelihood

Suppose each individual, reacting to announcement (song), estimates or quotes a Θ-value X_i (for the i^{th}), and further it is known that

$$P\{X_i \in (dx_i) | \Theta\} = f_i(x_i;\Theta)dx_i. \tag{2.2}$$

An important special case is that errors are normally distributed (or Gaussian):

$$f_i(x_i;\Theta) = \exp\left[-\frac{1}{2}\left(\frac{x_i-\Theta}{\sigma_i}\right)^2\right] \frac{1}{\sqrt{2\pi}\,\sigma_i}, \tag{2.3}$$

so WATCHER i ($i=1,2,\ldots,I$) estimates the value of the parameter Θ with errors that are $N(\Theta,\sigma_i^2)$. For the moment, assume that there is just the one BIRD present. If all I WATCHERS independently estimate Θ and do so with independently distributed errors, then it makes sense to write down and examine the likelihood function

$$L(\Theta;\underline{x}) = \prod_{i=1}^{I} f_i(x_i;\Theta) \tag{2.4}$$

or, taking logs and concentrating on the Normal form,

$$\ell(\theta;\underline{x}) = \sum_{i=1}^{I} \ln f_i(x_i;\theta)$$

$$= k - \frac{1}{2} \sum_{i=1}^{I} (x_i - \theta)^2 \left(\frac{1}{\sigma_i^2}\right) \tag{2.5}$$

omitting irrelevant constants. Now with σ_i^2 known the unrestricted maximization of the likelihood produces the time-honored formula

$$\hat{\theta} = \frac{\sum_{i=1}^{I} x_i/\sigma_i^2}{\sum_{i=1}^{I} 1/\sigma_i^2} \tag{2.6}$$

i.e. the variance-weighted mean of the individual observations. The variance of the estimate is

$$\mathrm{Var}[\hat{\theta}] = \frac{1}{\sum_{i=1}^{I} 1/\sigma_i^2}. \tag{2.7}$$

If all the above assumptions hold true, then one presumably compares $\hat{\theta}$ to the "known" parameters of various BIRDS and picks BIRD j#, where $|\hat{\theta} - \mu_{j\#}| < |\hat{\theta} - \mu_j|$, $j \neq j\#$; call this the <u>nearest neighbor</u> strategy, NN. Because of symmetry, the solution $\hat{\theta}$ is also the mean of a Bayes posterior with non-informative (flat, improper) prior. It is also the best linear, unbiased (BLUE) estimator of θ, so it should be at least mildly satisfactory to most non-rabid statisticians of any faith or persuasion. In the important (oversimplified) case in which each individual bird sings precisely on key so the unknown $\Theta = \mu_j$ for some j then the procedure yields the maximum likelihood estimator of μ from the restricted parameter space $\{\mu_1, \mu_2, \ldots, \mu_J\}$. See Hammersley (1950) for an early discussion of this problem.

2.2 A More Robust Likelihood

While many (perhaps transformed) measurement errors of physical quantities are approximately Normal, especially "in the middle" of their distribution, there can well be occasional outliers, in this case possibly caused by individual mis-performance. In order to model this empirically observed feature, it is becoming conventional to extend the tails of the Normal in (2.3) in one of these ways

(a) continuous scale mixing, where σ_i^2 is taken to be a conveniently distributed (e.g. inverse Gamma) random variable.

(b) ε contamination, wherein

$$f_i(x_i,\Theta) = (1 - \varepsilon_i) \exp[-\tfrac{1}{2} (\frac{x_i-\Theta}{\sigma_{i1}})^2] \frac{1}{\sqrt{2\pi}\,\sigma_{i1}} + \varepsilon_i \exp[-\tfrac{1}{2} (\frac{x_i-\Theta}{\sigma_{i2}})^2] \frac{1}{\sqrt{2\pi}\,\sigma_{i2}} \qquad (2.8)$$

in which usually $1 - \varepsilon_i = \bar{\varepsilon}_i$ is close to one and σ_{i1} is relatively small ($\sigma_{i1} < \sigma_{i2}$), while $\varepsilon_i > 0$ is close to zero, but σ_{i2} is large, e.g. $\sigma_{i2} = 10\sigma_{i1}$. This model was utilized in a classical robustness context by Tukey, and also by Berger and Berliner (1983) for Bayesian robustness purposes.

Begin by discussing (a). This approach can (but need not) lead to replacement of the normal observation density by a Student t form:

$$f_i(x_i;\theta) = \frac{C(d_i)}{\left[1 + \left(\frac{x_i-\theta}{\sigma_i}\right)^2 \frac{1}{d_i}\right]^{(d_i+1)/2}}. \tag{2.9}$$

Here view d_i as a shape tuning parameter; $\mathrm{Var}[X_i] = \frac{d_i}{d_i-2}\sigma_i^2$ if $d_i > 2$, but kurtosis (fourth central moment scaled to be dimension-free) can induce very extended tails, simulating outlier occurrence. If $d_i = 1$, we get the centered and scaled Cauchy, with notoriously long, symmetric tails. The Cauchy tails are so long that neither mean nor variance -- nor any other moment -- exists. The likelihood obtained by combining individual measures is now

$$L(\theta;\underline{x}) = \prod_{i=1}^{I} \frac{C(d_i)}{\left[1 + \left(\frac{x_i-\theta}{\sigma_i}\right)^2 \frac{1}{d_i}\right]^{(d_i+1)/2}} \tag{2.10}$$

and so, up to irrelevant constants,

$$\ln L(\Theta;\underline{x}) = \ell(\Theta;\underline{x}) = \sum_{i=1}^{I} \left(\frac{d_i+1}{2}\right) \ln\left[1 + \left(\frac{x_i-\Theta}{\sigma_i}\right)^2 \frac{1}{d_i}\right].$$

Now differentiation with respect to Θ gives

$$\frac{\partial \ell}{\partial \Theta} = \sum_{i=1}^{I} \left(\frac{d_i+1}{2}\right) \frac{2(x_i-\Theta)\frac{1}{\sigma_i^2 d_i}}{1 + \left(\frac{x_i-\Theta}{\sigma_i}\right)^2 \frac{1}{d_i}} = 0 \tag{2.11}$$

as a condition for a maximizing Θ, denoted by $\hat{\Theta}$. In principle this equation could have more than one solution; Copas has discussed this situation.

To obtain a usually sensible (optimal) solution, proceed as follows:

<u>Iterative Reweighting</u>

Rewrite (2.11) as follows:

$$\frac{\partial \ell}{\partial \Theta} = 0: \quad \sum_{i=1}^{I} \left(x_i - \hat{\Theta}(r+1)\right)\frac{1}{\sigma_i^2} \cdot w_i(r) = 0 \tag{2.12}$$

or

$$\hat{\Theta}(r+1) = \frac{\sum_{i=1}^{I} \left(\frac{x_i}{\sigma_i^2}\right) w_i(r)}{\sum_{i=1}^{I} \left(\frac{1}{\sigma_i^2}\right) w_i(r)} \tag{2.13}$$

where the weight

$$w_i(r) = \frac{\frac{d_i+1}{d_i}}{1+\left(\frac{x_i-\hat{\Theta}(r)}{\sigma_i}\right)^2 \frac{1}{d_i}}. \tag{2.14}$$

One might start the iteration at

$$\hat{\Theta}(1) = \text{median } (x_1, x_2, \ldots, x_I). \tag{2.15}$$

and then compute the first weight

$$w_i(1) = \frac{\dfrac{d_i+1}{d_i}}{1+\left(\dfrac{x_i-\hat{\Theta}(1)}{\sigma_i}\right)^2 \dfrac{1}{d_i}} \tag{2.16}$$

and use this to find the second estimate $\hat{\Theta}(2)$. Even the first iteration, as described, will be quite successful at taming down individual widely discrepant values, or "outliers". The smaller is d_i ($d_i \geq 1$, presumably) the more effectively discrepant values are reduced in influence.

After obtaining convergence, one may apply the NN approach to identify $j^{\#}$, the name (number) of the BIRD actually singing.

The above procedure will usually work satisfactorily, but may err because of an unfortunate starting estimate. If each BIRD sings almost precisely on key, so the unknown $\Theta = \mu_j$ for some j, then a precise maximum likelihood solution can be obtained by simple enumeration: one simply evaluates (2.10) for $\Theta = \{\mu_1, \mu_2, \ldots, \mu_J\}$ and picks $j = j^{\#}$ that gives the maximum. For small J this is computationally feasible and provides the truly maximum likelihood solution given the problem formulation. On the other hand, the weights produced by the iterative solution provide a convenient index as to the relative importance to be attached to the data values by the iterative procedure, so it seems sensible to become a "weight watcher". The pattern of weights might suggest reasons for relative comfort or discomfort with an identification: for instance, a relatively uniform distribution of low weights perhaps gives discomfort, while mostly high weights with a few very low ones thrown in may give reason for comfort -- presumably with a consensus of the high-weighted values. One fact that should be noted is that the likelihood equation (2.10) may well have multiple peaks or modes, and the primary one is presumably usually found by the re-weighted iteration NN scheme suggested. In any case, the parameter space point-by-point enumeration is often feasible.

Next discuss the ε-contamination model (b). Unfortunately the likelihood is of an awkward form

$$L(\Theta;\underline{x}) = \prod_{i=1}^{I}\left[\bar{\varepsilon}_i \exp\left[-\frac{1}{2}\left(\frac{x_i-\Theta}{\sigma_{i1}}\right)^2\right]\frac{1}{\sqrt{2\pi}\,\sigma_{i1}} + \varepsilon_i \exp\left[-\frac{1}{2}\left(\frac{x_i-\Theta}{\sigma_{i2}}\right)^2\right]\frac{1}{\sqrt{2\pi}\,\sigma_{i2}}\right]. \tag{2.17}$$

Now it can be seen that if the multiplication is carried out and some re-arrangement is done, we can express the likelihood as

$$L(\Theta;\underline{x}) = \sum_{k=1}^{K} \pi_k(\underline{x}) \exp\left[-\frac{1}{2}\left(\frac{\theta-\mu_k(\underline{x})}{\sigma_k(\underline{x})}\right)^2\right]\frac{1}{\sqrt{2\pi}\,\sigma_k(\underline{x})} \tag{2.18}$$

where

$$\pi_k(\underline{x}) = \pi_k \exp\{-\frac{1}{2} R_k(\underline{x})\} \tag{2.19}$$

and $K = 2^I$. It turns out that $\mu_k(\underline{x})$ is a linearly weighted function of the individual observations, and $1/\sigma_k^2(\underline{x})$ is a corresponding sum of inverse variances; $R_k(\underline{x})$ measures the discrepancy of the individual observations from $\mu_k(\underline{x})$.

For illustration, suppose $I = 2$, so, up to multiplicative constants,

$$L(\Theta,\underline{x}) = \prod_{i=1}^{2} [\bar{\varepsilon}_i \frac{1}{\sqrt{2\pi}\,\sigma_{i1}} \exp\{-\frac{1}{2} (\frac{x_i-\Theta}{\sigma_{i1}})^2\} + \varepsilon_i \frac{1}{\sqrt{2\pi}\,\sigma_{i2}} \exp\{-\frac{1}{2} (\frac{x_i-\Theta}{\sigma_{i2}})^2\}]$$

$$= \sum_{k=1}^{4} \pi_k(\underline{x}) \exp[-\frac{1}{2} (\frac{\Theta-\mu_k(\underline{x})}{\sigma_k(\underline{x})})^2] \frac{1}{\sqrt{2\pi}\,\sigma_k(\underline{x})}$$

where $\{\pi_k\} = \{\bar{\varepsilon}_1\bar{\varepsilon}_2, \varepsilon_1\bar{\varepsilon}_2, \bar{\varepsilon}_1\varepsilon_2, \varepsilon_1\varepsilon_2\}$; the terms $\mu_k(\underline{x})$ and $\sigma_k(\underline{x})$ are obtained by completing the square in the exponent of each summand.

The form of (2.18) suggests that $L(\Theta;\underline{x})$ is a possibly multimodal function, as was true of the Student t likelihood. An iterative scheme can be set up as detailed below to estimate Θ and the NN approach can then be taken. If each BIRD sings almost precisely on key so the unknown $\Theta = \mu_j$ for some j, then a precise maximum likelihood solution can be obtained by simple enumeration as before.

Iterative Reweighting

Taking logarithms of (2.17) and differentiating with respect to Θ results in the equation

$$\frac{\partial \ell}{\partial \Theta} = 0: \sum_{i=1}^{I} (x_i - \hat{\Theta}(r+1)) \left[\left(\frac{1}{\sigma_{i2}}\right)^2 + w_i(r)\left[\left(\frac{1}{\sigma_{i1}}\right)^2 - \left(\frac{1}{\sigma_{i2}}\right)^2\right]\right] \quad (2.20)$$

where the weight

$$w_i(r) = \frac{A_i(r)}{\varepsilon_i + A_i(r)} \quad (2.21)$$

with

$$A_i(r) = (1-\varepsilon_i) \frac{\sigma_{i2}}{\sigma_{i1}} \exp\left\{-\frac{1}{2} (x_i - \Theta(r))^2 \left[\left(\frac{1}{\sigma_{i1}}\right)^2 - \left(\frac{1}{\sigma_{i2}}\right)^2\right]\right\}. \quad (2.22)$$

As for the Student t distribution, one might start the iteration at $\hat{\Theta}(1)$ = median $(x_1, x_2, \ldots x_I)$ and then compute the first weight; use the weights to find the second estimate; etc.

3. Bayesian Formulations: Everything Normal

In addition to the information X_i on Θ coming from individual i there may be information on Θ codable in the form of a probability density: $p_{\underset{\sim}{\Theta}}(\Theta)$. The latter may actually take the form of a series of near delta functions, one for each of the BIRDS in question. For the sake of a bit of generality write

$$p_{\underset{\sim}{\Theta}}(\Theta) = \sum_{j=1}^{J} P_j \exp\left[-\frac{1}{2}\left(\frac{\Theta-\mu_j}{\tau_j}\right)^2\right] \frac{1}{\sqrt{2\pi}\,\tau_j}. \tag{3.1}$$

Here possibly $P_j \approx \frac{1}{J}$, where J represents the number of BIRDS believed to be in the vicinity and of interest. If $\tau_j = 0$ then the above indeed represents a "Dirac comb" with teeth at the points μ_j, $j=1,2,\ldots,J$; the sharpness of the teeth dictated by τ_j: small τ_j means that the j^{th} tooth (density) is long (tall) and sharp.

Now by routine Bayes we get for the posterior density

$$p_{\underset{\sim}{\Theta}|\underline{x}}(\Theta) = K \prod_{i=1}^{I} f_{X_i}(x_i;\Theta)\, p_{\underset{\sim}{\Theta}}(\Theta) \tag{3.2}$$

and, if we adopt the Normal model,

$$p_{\underset{\sim}{\Theta}|\underline{x}}(\Theta) = K \prod_{i=1}^{I} f_{X_i}(x_i;\Theta)\, p_{\underset{\sim}{\Theta}}(\Theta)$$

$$= K \exp\left[-\frac{1}{2}\sum_{i=1}^{I}\left(\frac{x_i - \Theta}{\sigma_i}\right)^2\right] \cdot \sum_{j=1}^{J} P_j \exp\left[-\frac{1}{2}\left(\frac{\Theta - \mu_j}{\tau_j}\right)^2\right]. \tag{3.3}$$

This can be simplified: write

$$\left(\frac{\Theta - \mu_j(\underline{x})}{\tau_j(\underline{x})}\right)^2 + Q_j = \left(\frac{\Theta - \mu_j}{\tau_j}\right)^2 + \sum_{i=1}^{I}\left(\frac{x_i - \Theta}{\sigma_i}\right)^2 \tag{3.4}$$

where Q_j doesn't depend on Θ, and look for $\mu_j(\underline{x})$ and $\tau_j^2(\underline{x})$ in terms of the observations and parameters; to do so differentiate with respect to Θ to find

$$2\left(\frac{\Theta - \mu_j(\underline{x})}{\tau_j(\underline{x})^2}\right) = 2\left(\frac{\Theta - \mu_j}{\tau_j^2}\right) + 2\sum_{i=1}^{I}\left(\frac{\Theta - x_i}{\sigma_i^2}\right). \tag{3.5}$$

Since the coefficients of Θ and 1 on each side of the equation must match,

$$\frac{1}{\tau_j^2(\underline{x})} = \frac{1}{\tau_j^2} + \sum_{i=1}^{I}\frac{1}{\sigma_i^2} \tag{3.6}$$

and

$$\mu_j(\underline{x}) = \left[\frac{\mu_j}{\tau_j^2} + \sum_{i=1}^{I}\frac{x_i}{\sigma_i^2}\right]\tau_j^2(\underline{x}). \tag{3.7}$$

To identify Q_j let $\Theta = \mu_j(\underline{x})$ in (3.4); then

$$Q_j = \left(\frac{\mu_j(\underline{x}) - \mu_j}{\tau_j}\right)^2 + \sum_{i=1}^{I}\left(\frac{x_i - \mu_j(\underline{x})}{\sigma_i}\right)^2; \tag{3.8}$$

that is, Q_j is the scaled sum of squared deviations of (a) the j^{th} posterior mean from the j^{th} prior mean and (b) the j^{th} posterior mean from each individual observation. Now return to (3.3) and substitute:

$$P_{\underset{\sim}{\Theta}|\underline{x}}(\theta) = K \sum_{j=1}^{J} P_j \exp\left[-\frac{1}{2}\left(\frac{\theta-\mu_j(\underline{x})}{\tau_j(\underline{x})}\right)^2\right]\exp\left[-\frac{1}{2} Q_j\right]. \tag{3.9}$$

By normalization,

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} P_{\underset{\sim}{\Theta}|\underline{x}}(\theta)\, d\theta \\ &= K \sum_{j=1}^{J} P_j \exp\left[-\frac{1}{2} Q_j\right] \tau_j(\underline{x}) \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{\theta-\mu_j(\underline{x})}{\tau_j(\underline{x})}\right)^2\right] \cdot \frac{d\theta}{\tau_j(\underline{x})} \qquad (3.10) \\ &= K \sum_{j=1}^{J} P_j \exp\left[-\frac{1}{2} Q_j\right] \sqrt{2\pi}\, \tau_j(\underline{x}). \end{aligned}$$

Thus the posterior density is of the form

$$P_{\underset{\sim}{\Theta}|\underline{x}}(\theta) = \sum_{j=1}^{J} P_j^{\#}(\underline{x}) \exp\left[-\frac{1}{2}\left(\frac{\theta-\mu_j(\underline{x})}{\tau_j(\underline{x})}\right)^2\right] \frac{1}{\sqrt{2\pi}\, \tau_j(\underline{x})} \tag{3.11}$$

where

$$P_j^{\#}(x) = \frac{P_j e^{-\frac{1}{2} Q_j} \tau_j(\underline{x})}{\sum_{j=1}^{J} P_j e^{-\frac{1}{2} Q_j} \tau_j(\underline{x})} \tag{3.12}$$

In other words by completing the square in (3.4), the resulting form of the posterior density (3.12) and the prior density, (3.1), resemble each other closely. For the important special case in which $\tau_j^2 = 0$, $j=1,2,\ldots,J$, one obtains the discrete distribution concentrated at the values μ_j ($j=1,2,\ldots,J$) having probability mass function

$$P\{\underset{\sim}{\theta} = \mu_j | \underset{\sim}{x}\} = P_j^{\#}(\underset{\sim}{x}) = \frac{P_j \exp[-\frac{1}{2} \sum_{i=1}^{I} (x_i - \mu_j)^2/\sigma_i^2]}{\sum_{j=1}^{J} P_j \exp[-\frac{1}{2} \sum_{i=1}^{I} (x_i - \mu_j)^2/\sigma_i^2]} . \tag{3.13}$$

The component probabilities P_j are very simply modified in accordance with the observations, and can be easily updated as more observations become available. One could hope that after a set of observations has become available then, say,

$$P_3^{\#} = 1$$

and

$$P_j^{\#} = 0 \quad \text{for } j \neq 3,$$

which points strongly at BIRD 3 as being the one that is actually singing. If, on the other hand, all $P_j^{\#}$-values were to remain similar it might be thought that some individuals have focused on two or more different items, <u>or</u> that the noise is not well represented by the normal (Gaussian) model. This possibility is not included in the present model, however. Note, too, that if the $\{P_j = \frac{1}{J}, j = 1, 2, \ldots, J\}$, the discrete uniform distribution, then naming $j=j^{\#}$ by picking the maximum probability from (3.13) is exactly equivalent to maximizing the likelihood by direct enumeration.

3.1 Bayesian Formulations; (1) Student t Observations

Let the prior information on BIRDS be given by (3.1). Individual watchers independently observe Θ with errors distributed according to the Student t family:

$$f_{X_i}(x_i;\Theta) = \frac{C(d_i)}{\left[1 + \left(\frac{x_i - \Theta}{\sigma_i}\right)^2 \frac{1}{d_i}\right]^{(d_i+1)/2}} \cdot \frac{1}{\sigma_i} \tag{3.14}$$

Then the posterior density is of the form

$$p_{\Theta|\underline{x}}(\Theta) = K \prod_{i=1}^{I} f_{X_i}(x_i;\Theta) \sum_{j=1}^{J} p_j \exp\left[-\frac{1}{2}\left(\frac{\Theta-\mu_j}{\tau_j}\right)^2\right] \tag{3.15}$$

$$= K' \sum_{j=1}^{J} p_j \exp\left[-\frac{1}{2}\left(\frac{\Theta-\mu_j}{\tau_j}\right)^2\right] \exp\left[-\frac{1}{2}\sum_{i=1}^{I}(d_i+1)\ \ln\left[1 + \left(\frac{x_i-\Theta}{\sigma_i}\right)^2 \frac{1}{d_i}\right]\right]$$

In order to normalize this expression (determine K'), and to compute moments (for point estimates, the NN approach, etc.) it is necessary to integrate over all Θ-values; of course no simple analytic closed form expression exists. There are two practical options:

(a) numerical integration, using Gauss-Hermite integration, e.g., by adopting the program of Naylor and Smith (1982); or

(b) analytical approximation, using a variant of the Laplace method, see deBruijn (1958) or the equivalent; this classical approach has been invoked by Mosteller and Wallace (1964), Gaver (1985), Tierney and Kadane (1986), Lindley and Singpurwalla (1986), and by many others as well.

To apply Laplace, write

$$p_{\underset{\sim}{\theta}|\underline{x}}(\theta) = K \sum_{j=1}^{J} p_j e^{-\frac{1}{2} S_j} \tag{3.16}$$

where

$$S_j = \left(\frac{\theta-\mu_j}{\tau_j}\right)^2 + \sum_{i=1}^{I} (d_i + 1) \ln\left[1 + \left(\frac{x_i-\theta}{\sigma_i}\right)^2 \cdot \frac{1}{d_i}\right] . \tag{3.17}$$

The plan is to replace S_j by an approximating quadratic in θ, and thus to exhibit closed-form approximating expressions for the updated probabilities $p_j^{\#}(\underline{x})$ that are quite analogous to the "exact" formulas (3.12) obtainable under normal/Gaussian error specifications.

To proceed, assume that the exponent is of the form

$$S_j \approx \left(\frac{\theta-\mu_j(\underline{x})}{\tau_j(\underline{x})}\right)^2 + Q_j(\underline{x}) \tag{3.18}$$

where $Q_j(\underline{x})$ is at least nearly independent of θ. Now differentiate the two forms on θ:

$$2\left(\frac{\Theta-\mu_j(\underline{x})}{\tau_j(\underline{x})^2}\right) = 2\left(\frac{\Theta-\mu_j}{\tau_j^2}\right) + \sum_{i=1}^{I} 2\left(\frac{d_i+1}{d_i}\right)\frac{(\theta-x_i)}{1+\left(\frac{x_i-\Theta}{\sigma_i}\right)^2\frac{1}{d_i}} \cdot \frac{1}{\sigma_i^2} \tag{3.19}$$

Now identify the coefficients of Θ and 1 to see that

$$\frac{1}{\tau_j^2(\underline{x})} = \frac{1}{\tau_j^2} + \sum_{i=1}^{I}\frac{1}{\sigma_i^2} \cdot \frac{(d_i+1)/d_i}{\left[1+\left(\frac{x_i-\Theta}{\sigma_i}\right)^2\frac{1}{d_i}\right]} = \frac{1}{\tau_j^2} + \sum_{i=1}^{I}\frac{1}{\sigma_i^2} w_i \tag{3.20}$$

and

$$\mu_j(\underline{x}) = \tau_j^2(\underline{x})\left[\frac{\mu_j}{\tau_j^2} + \sum_{i=1}^{I}\frac{x_i}{\sigma_i^2} w_i\right] \tag{3.21}$$

where the weights are

$$w_i = \frac{(d_i+1)/d_i}{\left[1+\left(\frac{x_i-\Theta}{\sigma_i}\right)^2\frac{1}{d_i}\right]} . \tag{3.22}$$

In practice, it will be necessary to estimate Θ by an iterative re-weighting procedure, so approximate weights will be used:

$$\hat{w}_i = \frac{(d_i+1)/d_i}{\left[1+\left(\frac{x_i-\hat{\Theta}}{\sigma_i}\right)^2\frac{1}{d_i}\right]} \tag{3.23}$$

Now replace S_j in (3.16) by the quadratic approximation to find

$$P_{\underset{\sim}{\theta}|\underline{x}}(\theta) = K' \sum_{j=1}^{J} P_j \exp[-\tfrac{1}{2}\hat{Q}_j(\underline{x})]\exp[-\tfrac{1}{2}(\frac{\theta-\mu_j(\underline{x})}{\tau_j(\underline{x})})^2] \tag{3.24}$$

where the approximate prior probability revision factor is from (3.18)

$$\hat{Q}_j(\underline{x}) = (\frac{\mu_j-\mu_j(\underline{x})}{\tau_j})^2 + \sum_{i=1}^{I}(d_i+1)\ln[1+(\frac{x_i-\mu_j(\underline{x})}{\sigma_i})^2 \cdot \frac{1}{d_i}] \tag{3.25}$$

and so

$$\hat{P}_j^{\#}(\underline{x}) = \frac{P_j e^{-\frac{1}{2}\hat{Q}_j(\underline{x})}\tau_j(\underline{x})}{\sum_{j=1}^{J} P_j e^{-\frac{1}{2}\hat{Q}_j(\underline{x})}\tau_j(\underline{x})} \tag{3.26}$$

provides the approximate data-updated probability that BIRD j is singing. It is reasonable to designate $j=j^{\#}$ if $\hat{P}^{\#}_{j^{\#}}(\underline{x}) > \hat{P}^{\#}_j(\underline{x})$ for $j \neq j^{\#}$. The form of the Bayes posterior is of course quite analogous to that derived for the normal errors case. Here we have

$$P_{\underset{\sim}{\theta}|\underline{x}}(\theta) = \sum_{j=1}^{J} \hat{P}^{\#}_j(\underline{x}) \exp[-\tfrac{1}{2}(\frac{\theta-\mu_j(\underline{x})}{\tau_j(\underline{x})})^2]\frac{1}{\sqrt{2\pi}\,\tau_j(\underline{x})} \tag{3.27}$$

with the squared-error-minimizing point estimate, i.e., the expected value of $\underset{\sim}{\theta}$ given observations $\underline{x}$, is

$$\hat{\theta} = \sum_{j=1}^{J} P_j^{\#}(\underline{x})\, \mu_j(\underline{x}). \tag{3.28}$$

It is this value that should apparently appear in the weights, $\hat{w}_i$, in the course of the iterative calculation.

The behavior of $\hat{Q}_j(\underline{x})$, and hence of the prior to posterior revision $P_j \to \hat{P}_j^{\#}(\underline{x})$ seems intuitively appealing: first, one is lead to estimate the most likely characteristic of the song of the j^{th} BIRD, $\mu_j(\underline{x})$, by combining data x_i, $i = 1, 2, \ldots, I$ (using the knowledge that outliers will occur, so the estimate is made robustly) and prior information about the variability of BIRD j's song. Then this estimate is effectively compared to (1) the candidate true mean value of j^{th} BIRD's song, μ_j, and (2) the data obtained by the listeners; both (1) and (2) measured on an appropriate scale of variability. If the sum of these distances (squared and scaled) is small, the conditional probability of j being the songster is correspondingly increased; otherwise, it is reduced.

If $\tau_j^2 = 0$, $j=1,2,\ldots,J$, so the BIRDS always sing precisely on key, then the above density becomes a probability mass function:

$$P_j^{\#}(\underline{x}) = P\{\underline{\theta}=j\,|\,\underline{x}\} = K\, P_j \prod_{i=1}^{I} \frac{C(d_i)}{\left[1 + \left(\frac{x_i-\mu_j}{\sigma_i}\right)^2 \frac{1}{d_i}\right]^{(d_i+1)/2}} \tag{3.29}$$

where

$$1 = K \sum_{j=1}^{J} P_j \prod_{i=1}^{I} \frac{C(d_i)}{\left[1 + \left(\frac{x_i - \mu_j}{\sigma_i}\right)^2 \frac{1}{d_i}\right]^{(d_i+1)/2}} \tag{3.30}$$

determines K. To identify the optimal $j=j^{\#}$, simply locate the maximum $P_j^{\#}$.

4. Results of Simulation Experiments

This section reports some of the results of simulation experiments to study the performance of various methods of combining WATCHER observations to obtain an estimate of the parameter of the singing BIRD. All simulations were carried out on an IBM 3033AP at the Naval Postgraduate School. Random numbers were generated using IMSL routines. Some details and results of the simulations are given below; for more see Meldrum [1986].

4.1 BIRDS with Univariate Parameters

There are 5 BIRDS with parameters $\{\mu_j\}$ equal to 1, 2, 3, 4 and 5. The BIRD that sings has parameter μ_j with probability p_j.

The number of WATCHERS varies between 2 and 5. The observation of the i^{th} WATCHER is

$$X_i = \Theta + E_i \tag{4.1}$$

where Θ is the parameter of the BIRD that sings and E_i is the observational error. The distributions of observational error considered are:

1) the normal distribution with mean 0 and standard deviation $\sigma = 0.5$ (e.g. (2.3));

2) the ε-contaminated normal (2.8) with mean 0, standard deviations $\sigma_1 = 0.5$ and $\sigma_2 = 5$, and contamination probability $\varepsilon = 0.1$ and 0.25; CN(ε);

3) the Student t distribution (2.9) with $\sigma = 0.5$ and $d = 1$, which is the Cauchy distribution.

Each simulation case has 10,000 replications. In each replication the BIRD with parameter μ_j was drawn to sing with probability p_j and WATCHER observations were generated. The following estimates of the parameter of the singing BIRD $\hat{\theta}$ were computed:

1. the mean of the observations;

2. the median of the observations;

3. the iterative Student-t estimate (2.12)-(2.16) with assumed values $\sigma = 0.5$ and $d = 1$ and 10;

4. the iterative ε-contaminated normal estimate (2.20)-(2.22) with assumed parameter values $\sigma_1 = 0.5$, $\sigma_2 = 5.0$ and $\varepsilon = 0.1$ and 0.25.

In each case, the BIRD whose parameter was closest to the estimated $\hat{\theta}$ was estimated to be the BIRD that sang.

In each replication Bayes procedures for combining WATCHER observations were also considered. The prior probability of the BIRD with parameter μ_j singing was assumed to be 1/5 in each case; (equally likely prior). The assumed error distributions for the Bayes models were as follows:

1. normal with mean 0 and standard deviation $\sigma = 0.5$;

2. Student t with $\sigma = 0.5$ and degrees of freedom $d = 1$, and 10;

3. The ε-contaminated normal with $\sigma_1 = 0.5$, $\sigma_2 = 5$ and $\varepsilon = 0.1$ and 0.25. For each assumed error distribution the posterior probability of the singing BIRD having parameter μ_j was computed.

The BIRD whose parameter had the largest posterior probability was the estimate of the BIRD that sang.

In Table 1 proportions of correct identifications are given for a simulation in which BIRD with parameter μ_j sings with probability 1/5. The number of WATCHERS varies between 2 and 5. Observation errors were generated using the "true error distribution". Estimates of Θ were computed using both correct and incorrect assumptions given in the first column concerning the error distribution. The Bayes estimates assumed the equally likely prior and correct and incorrect assumptions given in the first column about the error distribution.

Here are some conclusions that may be made from the simulations. The BIRD estimates based on the mean and the normal Bayes model are the most sensitive to incorrect error distribution assumptions; note that in the case in which the true error distribution is ε-contaminated normal with $\varepsilon = 0.25$, increasing the number of WATCHERS actually <u>decreases</u> the proportion of correct identifications for these two procedures. This behavior results from the fact that, with small numbers of WATCHERS, increasing the number WATCHERS increases the chances of having one or more outlying observations. A more detailed explanation of this phenomenon can be found in the Appendix. All of the procedures do about the same when the true error distribution is normal. When the true error distribution is not normal, the Bayes estimates based on the correct prior of equally likely BIRDS and error distribution other than normal or Student t with 10 degrees of freedom tend to have the highest proportion of correct identifications.

In Table 2 the proportions of correct identifications are given for a simulation experiment in which the BIRD with parameter $\mu = 1$ always sings in each replication; this parameter is on an extreme of the parameter set. In

Table 3, the proportions are given for an experiment in which the BIRD with parameter $\mu = 3$ always sings; this parameter is in the middle of the parameter set. The other parameters in the simulations remain the same as those for the simulation in Table 1. In particular the Bayes estimation procedures use the (incorrect) prior of equally likely BIRDS.

Comparing Tables 1-3, the proportion of correct identifications is smallest (respectively largest) in the case in which BIRD 3 (respectively BIRD 1) always sings; that is, the position of the parameter of the singing BIRD within the parameter space can make correct identification easier or harder. Once again procedures based on the normal distribution (mean and normal Bayes) do well if the true error distribution is normal but tend to have smaller proportions of correct identifications if the true error distribution is not normal; this decrease in the proportion of correct identifications is greater than the decrease obtained by using procedures based on non-normality of the error distribution when in fact the observations have normal errors. In Tables 2 and 3 the effect of using incorrect prior distributions in the Bayes models has been to make their proportions of correct identifications closer to those of the parametric procedures. However, the Bayes procedure based on an error distribution Student-t with 1 degree of freedom appears to be quite robust to model assumptions particularly for the case of 2 WATCHERS.

4.2 BIRDS with Bivariate Parameters

In this subsection there are 5 BIRDS. Each BIRD has two parameters associated with it. Two configurations of the BIRDS' parameters were considered:

LINE: $\mu_1 = (1,1)$, $\mu_2 = (2,2)$, $\mu_3 = (3,3)$, $\mu_4 = (4,4)$, $\mu_5 = (5,5)$;

BOX: $\mu_1 = (2,2)$, $\mu_2 = (2,4)$, $\mu_3 = (3,3)$, $\mu_4 = (4,2)$, $\mu_5 = (4,4)$.

The observation errors have the following distributions

1. ε -contaminated bivariate normal with density function

$$f(x,y) = (1-\varepsilon)\frac{1}{2\pi\sigma_{1,1}\,\sigma_{1,2}\sqrt{1-\rho_1^2}}\exp\{-Q_1(x,y)\} \tag{4.2}$$

$$+\ \varepsilon\ \frac{1}{2\pi\,\sigma_{2,1}\,\sigma_{2,2}\sqrt{1-\rho_2^2}}\exp\{-Q_2(x,y)\}$$

where

$$2(1-\rho_i^2)\,Q_i(x,y) = \left(\frac{x}{\sigma_{i1}}\right)^2 - 2\rho_i\left(\frac{x}{\sigma_{i1}}\right)\left(\frac{y}{\sigma_{i2}}\right) + \left(\frac{y}{\sigma_{i2}}\right)^2.$$

The parameters used are $\sigma_{1,1} = \sigma_{1,2} = 0.5$; $\sigma_{2,1} = \sigma_{2,2} = 5$, $\rho_2 = \rho_1 = -0.5$, 0, 0.5, and $\varepsilon = 0$, 0.1, and 0.25.

Note when $\varepsilon = 0$, the error distribution is bivariate normal with $\sigma_1 = \sigma_2 = 0.5$ and $\rho = -0.5$, 0, 0.5.

2. A bivariate Student-t with density function

$$f(x,y) = \frac{c}{\sigma_1\sigma_2\sqrt{1-\rho^2}}\left[1 + \frac{1}{d}\,\frac{1}{2(1-\rho)^2}\left[\left(\frac{x}{\sigma_1}\right)^2 - 2\rho\left(\frac{x}{\sigma_1}\right)\left(\frac{y}{\sigma_2}\right) + \left(\frac{y}{\sigma_2}\right)^2\right]\right]^{-((d+2)/2)}$$

where c is a constant and d is the number of degrees of freedom. The parameters used are $\sigma_1 = 0.5$, $\sigma_2 = 5.0$, d=1, and $\rho = -0.5$, 0, 0.5.

Each simulation case has 2,000 replications. For each replication a BIRD with parameter μ_j is drawn to sing with probability p_j. An independent observational error is drawn for each WATCHER from one of the error distributions. The WATCHER observations are combined by using one of the procedures below to determine the BIRD that sang.

1. <u>Median</u>: Compute the medians of the WATCHER observations of the first parameter and of the second parameter. Compute the Euclidean distance of the median pair to each BIRD parameter pair. The BIRD whose parameter pair has the smallest distance is estimated to be the one that sang.

2. <u>MLE normal</u>: The observational errors are assumed to have a bivariate normal distribution with parameters $\sigma_1 = 0.5$ $\sigma_2 = 0.5$ and correlation coefficient ρ. The likelihood is calculated for each BIRD parameter pair and the BIRD having the largest likelihood is estimated to be the BIRD that sang.

3. <u>MLE CN</u>: The observational errors are assumed to have an ε - contaminated normal distribution function with parameters $\sigma_{i1} = \sigma_{i2} = 0.5$ and $\sigma_{21} = \sigma_{22} = 5$, ρ and ε. The procedure is the same as 2.

4. <u>MLE T</u>: The same as 2 and 3 except the observational errors are assumed to have a Student t distribution with parameters $\sigma_1 = \sigma_2 = 0.5$, ρ and d.

Table 4 shows the proportion of correct identifications for a case in which the BIRDS' parameters are in the LINE configuration. Each BIRD is equally likely to sing for each replication. The true error distributions simulated were the bivariate normal, the ε-contaminated normal with $\varepsilon = 0.1$, the ε-contaminated normal with $\varepsilon = 0.25$, and the Student t with 1 degree of freedom, (Cauchy); they are listed in the first column of the table; the correlation coefficients of the simulated errors were $\rho = 0.5$, 0, and -0.5 and are listed on the first row of the table. The number of WATCHERS varies between 2 and 5. The estimation procedures are listed in the second column of the Table and assumed $\rho = 0.5$.

The simulations of Table 5 used the same models and estimation procedures as those of Table 4 except that the maximum likelihood procedures assumed $\rho = -0.5$. A comparison of Tables 4 and 5 indicates that the value of the assumed ρ for the maximum likelihood procedures made little difference in the proportion of correct identifications.

In both Tables a comparison of the proportion of correct identifications when the correlation coefficient of the true error distribution is RHO = 0.5 with those when RHO = 0 or -0.5 indicates that it is more difficult to identify the correct singing BIRD when the errors of observation for the BIRD'S two parameters are positively correlated.

A comparison of the proportion of correct identifications when the normal maximum likelihood method is used, with the other methods indicates that the normal estimate is the most sensitive to incorrect assumptions concerning the error distribution. As was true in the uninvariate case, use of the normal estimate on data whose true error distribution has longer tails than normal can result in decreasing proportions of correct identifications as the number of WATCHERS increases.

Table 6 reports results of a simulation experiment with models and estimation procedures the same as in Table 4 except that BIRD whose parameter is (3,3) always sings. Comparison of the results of Tables 4 and

6 indicates that the position of the singing BIRD in a pattern can affect the chances of a correct identification.

Table 7 reports results of a simulation experiment in which the parameters and estimates are the same as those of Table 4 except that the BIRDS' parameters are in the BOX pattern instead of the LINE pattern. A comparison of Tables 4 and 7 indicates that, if the correlation coefficient of the true error distribution is $\rho = 0.5$ (respectively $\rho = -0.5$), then it is easier (respectively harder) to make a correct identification of the singing BIRD with the BIRDS' parameters in the BOX pattern.

4.3 Conclusions from the Simulation Study

1. Estimation and identification procedures based on assumptions of normal errors are sensitive to outlying observations.

2. Estimation procedures based on assumptions of a long-tailed error distribution are more robust to incorrect error distribution assumptions than normal estimation procedures.

3. Bayes estimation procedures are sensitive to incorrect specification of the prior distribution of which BIRD is singing.

4. The following attributes affect the ability to correctly identify the singing BIRD.

   a. If each BIRD has more than 1 parameter, correlation between the parameters' observation errors can influence the difficulty of identification of the correct BIRD.

   b. The configuration of the parameter space for the BIRDS can make correct identification more difficult, e.g. LINE, BOX.

c. The location of the parameter of the singing BIRD in the parameter space can make correct identification easier or harder, e.g. middle BIRD or end BIRD in the univariate parameter case.

REFERENCES

J. O. Berger and L. M. Berliner. Robust Bayes and empirical Bayes analysis with ε-contaminated priors. Ann. Stat., 14 (1986), pp. 461-486.

J. B. Copas. A Computer-Intensive Non-Parametric Estimate of a Mixing Distribution with Application to Poisson Failure Data. Technical Report, Naval Postgraduate School, Monterey, Calif., 1984.

N. G. deBruijn. Asymptotic Methods in Analysis, Wiley Interscience: New York, 1958.

D. P. Gaver. Discrepancy Tolerant Hierarchical Poisson Event Rate Analysis. Technical Report, Naval Postgraduate School, Monterey, Calif., 1985.

J. M. Hammersley. On estimating restricted parameters, J. Royal Stat. Soc. B 12 (1950), pp. 192-240.

D. V. Lindley and N. D. Singpurwalla. Reliability (and fault tree) analysis using experts opinions, J. Amer. Stat. Assoc., 81 (1986), pp. 87-90.

A. G. Meldrum. Transmitter Identification with a Small Number of Independent Observers, Masters Thesis, Naval Postgraduate School, Monterey, Calif. 1986.

F. Mosteller and D. L. Wallace. Inference and Disputed Authorship: The Federalist. Addison - Wesley: Reading, Ma, 1964.

J. C. Naylor and A. F. M. Smith. Applications of a Method for the Efficient Computation of Posterior Distributions. Applied Statistics, 31, (1982), pp. 214-225.

L. Tierney and J. B. Kadane. Accurate approximations for posterior moments and marginal densities. J. Amer. Stat. Assoc., 81 (1986), pp. 82-86.

APPENDIX

HIT PROBABILITY WHEN BIRDS ARE ON A LINEAR LATTICE, ERRORS ARE ε-CONTAMINATED, AND A LINEAR SUMMARY, NEAREST-NEIGHBOR ALGORITHM IS USED: LINEAR CONSENSUS PROCEDURES NEED NOT SHOW "SAFETY IN NUMBERS".

Suppose BIRD characteristics μ_j are at equal intervals: $\mu_j = 0, \pm 1, \pm 2, \ldots$ with no limit to number. Let the i^{th} of I WATCHERS have the ε-contaminated error density

$$f_i(x_i;\Theta) = \bar{\varepsilon} \exp\left[-\frac{1}{2}\left(\frac{x_i-\Theta}{\sigma_{1i}}\right)^2\right]\frac{1}{\sqrt{2\pi}\,\sigma_{1i}} + \varepsilon \exp\left[-\frac{1}{2}\left(\frac{x_i-\Theta}{\sigma_{2i}}\right)^2\right]\frac{1}{\sqrt{2\pi}\,\sigma_{2i}} \qquad (A-1)$$

where $\bar{\varepsilon} + \varepsilon = 1$, $\bar{\varepsilon} \geq 0$, $\varepsilon \geq 0$. It is known that a BLUE of Θ is

$$\hat{\Theta}_{BLUE} = \frac{\sum_{i=1}^{I} x_i/\sigma_i^2}{\sum_{i=1}^{I} 1/\sigma_i^2} \qquad (A-2)$$

where here

$$\sigma_i^2 = \bar{\varepsilon}\,\sigma_{1i}^2 + \varepsilon\,\sigma_{2i}^2 \qquad (A-3)$$

and that

$$\mathrm{Var}[\hat{\theta}_{BLUE}] = \frac{1}{\sum_{i=1}^{I} 1/\sigma_i^2} \quad . \tag{A-4}$$

Clearly there is a tendency for the above variance to decrease with I, so one might conclude that adding more WATCHERS improves hit probability. This conclusion is false. Perhaps more surprisingly, existence of theoretical population moments does not seem to govern the behavior of the linear estimate, $\hat{\theta}_{BLUE}$. Of course, if σ_i^2 doesn't exist then the above weighting can not be carried out, but if the error scale is the same for all WATCHERS then equal weighting is suggested. It can be seen analytically that the Student t with one d.f. (the Cauchy) error model implies that $\hat{\Theta}_{BLUE}$ has exactly the same distribution regardless of the number of WATCHERS, and this effect is plainly visible from simulations and numerical calculations. On the other hand, the ε-contaminated Normal/Gauss error model has all moments finite, and yet can exhibit a hit probability that <u>decreases</u> with I, later of course increasing as it must, eventually, by central limit theorem effects. For what is possibly a plausible example, the advantage of number becomes evident only after about a dozen WATCHERS are performing simultaneously!

Simplest Case: Statistically Identical WATCHERS

Let $\sigma_{1i}^2 = \sigma_1^2$, $\sigma_{2i}^2 = \sigma_2^2$, $\sigma_1^2 < \sigma_2^2$. To calculate hit probability, assume with no loss of generality that BIRD 0 is singing. Then the probability of a hit is the probability that the ordinary average of I errors lies between $-\frac{1}{2}$ and $+\frac{1}{2}$:

$$P\{HIT\} = P\{-\frac{1}{2} < \frac{X_1 + X_2 + \ldots + X_I}{I} < \frac{1}{2}\} \tag{A-5}$$

Condition on the error components involved: if G represents the number of "good" (small variance, σ_i^2) observations, and $B = I - G$ the number of "bad" (large variance, σ_2^2) then $G \sim \text{Binomial}\ (\bar{\varepsilon}, I)$ and, given G,

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_I}{I} \sim N(0, \frac{G\sigma_1^2 + (I-G)\sigma_2^2}{I^2}) \tag{A-6}$$

so

$$P\{HIT|G=g\} = 2\Phi\left(\frac{\frac{1}{2}}{\sqrt{(g(\sigma_1^2 - \sigma_2^2) + I\sigma_2^2)/I^2}}\right) - 1 = \Phi_{g,I} \tag{A-7}$$

Consequently, when the condition is removed,

$$P\{HIT\} = \sum_{g=0}^{I} \binom{I}{g} (\bar{\varepsilon})^g (\varepsilon)^{I-g}\, \Phi_{g,I} \tag{A-8}$$

Numerical Illustration

Suppose BIRDS occur as above and that WATCHERS are independent with ε-contaminated errors having parameters $\varepsilon = .25$, $\sigma_{1i} = 0.5$, $\sigma_{2i} = 10\sigma_{1i}$, so $\sigma_{2i} = 5$. Then the BLUE is the ordinary average, $\bar{X} = \hat{\theta}$, which is also a Bayes estimate if one were to assume that the error distribution is simple Normal and the prior probabilities equal. Adopt the NN approach (what else?) to identify the singing BIRD. Then we tabulate the

HIT PROBABILITIES

ALGORITHM	NUMBER OF WATCHERS, I						
	2	4	6	8	10	12	14
$\hat{\theta}$ = AVERAGE, NN	0.54	0.48	0.48	0.49	0.51	0.55	0.57
$\hat{\theta}$ = MEDIAN, NN	0.54	0.77	0.86	0.92	0.94	0.96	0.98

The effect mentioned is quite striking, with linear $\hat{\theta}$ hit probability awryyKov rff quickly, recovering slowly, and not approaching that of the median until a value of I much larger than any in our table is reached.

Table 1

Proportion of Correct Identifications
BIRDS Equally Likely

True Error Distribution

	Normal $\sigma = .5$				Student-t 1 df				Contaminated Normal $\varepsilon = 0.1$				Contaminated Normal $\varepsilon = 0.25$			
ber of WATCHERS:	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
ameter Estimate																
n	.88	.93	.96	.98	.60	.60	.61	.60	.77	.78	.79	.77	.64	.61	.59	.58
ian	.88	.89	.95	.95	.60	.75	.78	.83	.77	.85	.91	.92	.64	.78	.81	.86
r. T 1df	.88	.91	.95	.97	.61	.76	.80	.85	.77	.88	.92	.92	.64	.81	.85	.90
r. T 3df	.88	.93	.96	.98	.61	.76	.80	.84	.77	.88	.92	.95	.65	.80	.84	.89
er. CN ε=0.1	.88	.93	.96	.98	.60	.74	.78	.82	.77	.89	.93	.96	.64	.82	.86	.91
er. CN ε=0.25	.88	.93	.96	.98	.60	.75	.78	.83	.77	.89	.93	.96	.64	.82	.86	.91
yes Estimates																
rmal	.88	.93	.96	.98	.60	.59	.60	.60	.77	.78	.78	.77	.64	.60	.58	.57
1df	.86	.92	.95	.97	.68	.77	.82	.86	.81	.88	.93	.95	.75	.82	.87	.91
10df	.88	.93	.96	.98	.65	.74	.79	.83	.80	.88	.92	.94	.70	.78	.83	.87
ε=0.1	.88	.93	.96	.98	.67	.75	.79	.83	.82	.90	.94	.96	.75	.83	.88	.92
ε=0.25	.87	.95	.96	.98	.67	.75	.79	.83	.82	.89	.94	.96	.75	.83	.88	.92

Table 2

Proportion of Correct Identifications
BIRD with PARAMETER 1 SINGS

True Error Distribution

	Normal $\sigma=0.5$				Student-t 1 df				Contaminated Normal $\varepsilon=0.1$				Contaminated Normal $\varepsilon=0.25$			
Number of WATCHERS:	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
Parameter Estimate																
Mean	.92	.96	.98	.99	.75	.75	.75	.74	.85	.86	.86	.86	.77	.75	.74	.75
Median	.92	.93	.97	.97	.75	.85	.86	.89	.85	.91	.94	.95	.77	.86	.89	.92
Iter. T 1df	.92	.94	.97	.98	.74	.85	.86	.90	.85	.92	.95	.96	.77	.87	.88	.93
Iter. T 3df	.92	.95	.97	.98	.74	.84	.86	.89	.85	.92	.95	.96	.76	.87	.89	.93
Iter. CN $\varepsilon=0.1$	.92	.96	.98	.99	.75	.84	.86	.89	.85	.93	.96	.97	.77	.89	.92	.95
Iter. CN $\varepsilon=0.25$	.92	.95	.97	.98	.75	.84	.87	.90	.85	.93	.96	.97	.77	.89	.91	.95
Bayes Estimates																
Normal	.92	.96	.98	.99	.77	.79	.81	.82	.85	.87	.87	.87	.78	.76	.76	.78
T 1df	.91	.95	.97	.98	.78	.85	.88	.91	.87	.93	.96	.97	.82	.88	.92	.94
T 10df	.92	.96	.98	.99	.77	.84	.87	.89	.86	.92	.95	.96	.80	.86	.89	.92
CN $\varepsilon=0.1$	.92	.96	.98	.99	.76	.82	.87	.89	.87	.93	.96	.97	.82	.89	.92	.95
CN $\varepsilon=0.25$	.92	.95	.97	.98	.76	.82	.87	.89	.87	.93	.96	.97	.82	.89	.92	.95

Table 3

Proportion of Correct Identifications
BIRD with PARAMETER 3 SINGS

True Error Distribution

	Normal σ =0.5				Student-t 1 df				Contaminated ε =0.1				Normal ε =0.25			
er of WATCHERS:	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
meter Estimate																
l	.84	.91	.96	.97	.49	.50	.50	.50	.71	.72	.72	.72	.54	.51	.49	.48
ian	.84	.86	.93	.94	.49	.69	.72	.79	.71	.81	.88	.90	.54	.73	.77	.83
r. T 1df	.84	.89	.94	.96	.51	.70	.76	.81	.71	.84	.90	.93	.55	.76	.81	.87
r. T 3df	.84	.90	.95	.97	.51	.70	.74	.81	.71	.85	.90	.93	.56	.75	.80	.87
r. CN ε=0.1	.84	.91	.95	.97	.49	.68	.73	.78	.71	.86	.92	.95	.56	.77	.82	.89
r. CN ε=0.25	.84	.90	.95	.97	.49	.69	.73	.79	.71	.86	.91	.95	.54	.77	.82	.89
yes Estimates																
rmal	.84	.91	.96	.97	.49	.50	.50	.49	.71	.72	.72	.72	.54	.51	.48	.48
1df	.82	.89	.94	.96	.62	.71	.78	.82	.77	.85	.91	.94	.69	.78	.84	.89
10df	.84	.91	.95	.98	.56	.68	.73	.79	.75	.83	.89	.93	.63	.72	.79	.85
ε=0.1	.84	.91	.95	.97	.49	.68	.73	.78	.71	.86	.92	.95	.54	.77	.82	.89
ε=0.25	.84	.90	.95	.97	.49	.69	.73	.79	.71	.86	.91	.95	.54	.77	.82	.89

Table 4

LINE Pattern
Equally Likely BIRDS
Proportion of Correct Identifications

Assumed RHO=0.5

CORRELATION COEFFICIENT USED TO GENERATE THE ERROR

		RHO=0.5				RHO=0				RHO=¯0.5			
	Number of WATCHERS:	2	3	4	5	2	3	4	5	2	3	4	5
True Error Dist.	Estimate												
CN(ε=0)	MEDIAN	.92	.94	.98	.98	.96	.97	.99	.99	1	.99	1	1
	NORMMLE	.92	.96	.99	.98	.96	.99	.99	1	1	1	1	1
	TMLE 1df	.90	.95	.98	.99	.94	.98	.99	1	1	1	1	1
	TMLE 3df	.92	.95	.98	.99	.96	.98	.99	1	1	1	1	1
	TMLE 10df	.92	.95	.98	.99	.96	.99	.99	1	1	1	1	1
	CNMLE(ε=.1)	.92	.96	.99	.99	.96	.99	.99	1	1	1	1	1
	CNMLE(ε=.25)	.92	.96	.99	.99	.96	.99	.99	1	1	1	1	1
CN(ε=0.1)	MEDIAN	.82	.91	.95	.96	.85	.93	.96	.98	.87	.97	.97	.99
	NORMMLE	.82	.81	.83	.84	.85	.84	.83	.83	.87	.85	.84	.82
	TMLE 1df	.87	.94	.96	.97	.92	.97	.98	1	.97	.99	1	1
	TMLE 3df	.88	.94	.97	.98	.93	.97	.98	.99	.97	.99	1	1
	TMLE 10df	.88	.94	.97	.97	.92	.96	.98	.99	.95	.99	.99	1
	CNMLE(ε=0.1)	.89	.94	.97	.98	.93	.97	.99	1	.97	.99	1	1
	CNMLE(ε=.25)	.89	.94	.97	.98	.93	.97	.99	1	.97	.99	1	1
CN(ε=.25)	MEDIAN	.68	.85	.86	.92	.70	.87	.88	.93	.73	.88	.89	.94
	NORMMLE	.68	.67	.65	.68	.70	.67	.67	.65	.73	.68	.67	.65
	TMLE 1df	.82	.91	.93	.97	.86	.94	.97	.98	.93	.96	.98	.99
	TMLE 3df	.82	.91	.93	.96	.86	.94	.96	.98	.92	.95	.98	.99
	TMLE 10df	.80	.89	.92	.96	.85	.91	.94	.96	.89	.92	.96	.98
	CNMLE(ε=0.1)	.83	.92	.94	.97	.87	.94	.97	.98	.93	.97	.98	.99
	CNMLE(ε=.25)	.83	.92	.94	.97	.87	.95	.97	.98	.93	.97	.98	.99
T 1 d.f.	MEDIAN	.63	.79	.82	.87	.69	.83	.86	.91	.77	.88	.90	.93
	NORMMLE	.63	.63	.65	.64	.69	.68	.70	.69	.77	.77	.74	.76
	TMLE 1df	.78	.84	.89	.91	.80	.89	.94	.95	.88	.94	.97	.98
	TMLE 3df	.76	.84	.88	.91	.80	.89	.94	.94	.88	.94	.96	.98
	TMLE 10df	.75	.80	.85	.89	.79	.87	.92	.94	.87	.93	.95	.97
	CNMLE(ε=0.1)	.76	.81	.86	.89	.80	.87	.92	.94	.87	.92	.95	.97
	CNMLE(ε=.25)	.72	.82	.87	.89	.79	.87	.92	.94	.87	.92	.96	.97

Table 5

LINE Pattern
Equally Likely BIRDS
Proportion of Correct Identifications

Assumed ρ = ¯0.5

CORRELATION COEFFICIENT USED TO GENERATE THE ERROR

		RHO=0.5				RHO=0				RHO=¯0.5			
	Number of WATCHERS:	2	3	4	5	2	3	4	5	2	3	4	5
True Error Dist.	Estimate												
CN(ε=0)	MEDIAN	.92	.95	.97	.98	.96	.97	.99	.99	1	1	1	1
	NORMMLE	.92	.97	.98	.99	.96	.99	.99	1	1	1	1	1
	TMLE 1df	.89	.95	.97	.99	.95	.98	1	1	.99	1	1	1
	TMLE 3df	.91	.96	.98	.99	.96	.98	1	1	.99	1	1	1
	TMLE 10df	.91	.96	.98	.99	.96	.99	.99	1	1	1	1	1
	CNMLE(ε=0.1)	.89	.94	.97	.99	.95	.98	1	1	1	1	1	1
	CNMLE(ε=.25)	.88	.94	.96	.98	.94	.98	.99	1	.99	1	1	1
CN(ε=0.1)	MEDIAN	.82	.90	.95	.96	.86	.94	.96	.98	.87	.97	.98	.99
	NORMMLE	.82	.83	.82	.83	.86	.85	.83	.84	.87	.87	.84	.83
	TMLE 1df	.88	.92	.96	.98	.93	.97	.98	.99	.98	.99	1	1
	TMLE 3df	.89	.93	.96	.98	.93	.97	.99	1	.98	.99	1	1
	TMLE 10df	.89	.93	.96	.98	.93	.97	.98	.99	.97	.99	.99	1
	CNMLE(ε=0.1)	.87	.92	.96	.97	.93	.97	.99	.99	.98	1	1	1
	CNMLE(ε=.25)	.87	.91	.95	.97	.93	.97	.98	.99	.98	1	1	1
CN(ε=.25)	MEDIAN	.69	.82	.85	.92	.71	.86	.87	.94	.72	.89	.88	.95
	NORMMLE	.69	.68	.64	.66	.71	.68	.66	.66	.72	.67	.66	.66
	TMLE 1df	.82	.88	.92	.95	.87	.93	.97	.98	.93	.97	.99	.99
	TMLE 3df	.83	.88	.92	.95	.87	.93	.97	.98	.92	.97	.99	.99
	TMLE 10df	.83	.87	.92	.96	.87	.92	.96	.97	.90	.95	.98	1
	CNMLE(ε=0.1)	.82	.88	.92	.96	.87	.94	.97	.98	.93	.98	.99	1
	CNMLE(ε=.25)	.81	.88	.91	.95	.87	.94	.97	.98	.93	.98	.99	1
T 1 d.f.	MEDIAN	.63	.79	.82	.87	.68	.83	.85	.92	.76	.89	.91	.94
	NORMMLE	.63	.64	.64	.63	.68	.69	.70	.70	.76	.76	.78	.77
	TMLE 1df	.76	.81	.87	.91	.81	.89	.93	.96	.89	.94	.97	.98
	TMLE 3df	.76	.82	.88	.91	.81	.89	.93	.96	.89	.94	.97	.98
	TMLE 10df	.75	.81	.87	.90	.79	.87	.92	.95	.88	.93	.97	.98
	CNMLE(ε=0.1)	.74	.80	.85	.88	.79	.88	.91	.95	.88	.93	.97	.98
	CNMLE(ε=.25)	.74	.80	.85	.88	.79	.88	.91	.94	.88	.93	.97	.97

Table 6

LINE Pattern
BIRD (3,3) Always Sings
Proportion of Correct Identifications

Assumed RHO = 0.5

CORRELATION COEFFICIENT USED TO GENERATE THE ERROR

		RHO=0.5				RHO=0				RHO=¯0.5			
	Number of WATCHERS:	2	3	4	5	2	3	4	5	2	3	4	5
True Error Dist.	Estimate												
CN(ε=0)	MEDIAN	.90	.93	.97	.98	.95	.97	.99	.99	1	.99	1	1
	NORMMLE	.90	.96	.98	.99	.95	.99	.99	1	1	1	1	1
	TMLE 1df	.87	.94	.97	.99	.94	.98	.99	.99	.99	1	1	1
	TMLE 3df	.89	.95	.98	.99	.95	.99	.99	1	.99	1	1	1
	TMLE 10df	.89	.96	.98	.99	.95	.99	.99	1	.99	1	1	1
	CNMLE(ε=0.1)	.90	.96	.98	.99	.95	.99	.99	1	.99	1	1	1
	CNMLE(ε=.25)	.89	.96	.98	.99	.95	.98	.99	1	.99	1	1	1
CN(ε=0.1)	MEDIAN	.75	.88	.94	.95	.81	.93	.96	.97	.85	.96	.96	.99
	NORMMLE	.75	.78	.78	.77	.81	.80	.79	.80	.85	.81	.79	.80
	TMLE 1df	.84	.92	.96	.97	.90	.96	.98	.99	.97	.99	1	1
	TMLE 3df	.85	.92	.96	.98	.90	.97	.98	.99	.97	.98	1	1
	TMLE 10df	.84	.91	.96	.98	.89	.96	.98	.99	.95	.97	.99	1
	CNMLE(ε=0.1)	.85	.92	.96	.98	.91	.97	.98	.99	.97	.99	1	1
	CNMLE(ε=.25)	.85	.92	.96	.98	.91	.97	.98	.99	.97	.99	1	1
CN(ε=0.25)	MEDIAN	.59	.81	.83	.90	.62	.82	.85	.93	.66	.85	.87	.93
	NORMMLE	.59	.57	.58	.57	.62	.59	.57	.58	.66	.59	.59	.57
	TMLE 1df	.78	.86	.92	.95	.84	.91	.96	.98	.89	.95	.98	.99
	TMLE 3df	.78	.87	.93	.95	.83	.91	.95	.98	.89	.94	.97	.99
	TMLE 10df	.76	.85	.91	.94	.80	.89	.94	.96	.85	.92	.96	.97
	CNMLE(ε=0.1)	.79	.88	.93	.96	.84	.93	.96	.98	.90	.95	.98	.99
	CNMLE(ε=.25)	.79	.88	.93	.96	.84	.93	.97	.98	.90	.96	.98	.99
T 1 d.f.	MEDIAN	.53	.74	.76	.85	.61	.80	.81	.89	.69	.84	.87	.93
	NORMMLE	.53	.54	.54	.56	.61	.63	.61	.62	.69	.69	.71	.73
	TMLE 1df	.71	.81	.85	.90	.76	.86	.90	.94	.84	.91	.95	.98
	TMLE 3df	.70	.80	.85	.90	.75	.86	.90	.94	.84	.91	.95	.98
	TMLE 10df	.68	.78	.83	.89	.73	.85	.89	.93	.83	.89	.95	.98
	CNMLE(ε=0.1)	.70	.79	.83	.87	.75	.84	.88	.92	.83	.89	.93	.97
	CNMLE(ε=.25)	.70	.79	.83	.88	.75	.84	.89	.93	.83	.89	.93	.97

Table 7

BOX Pattern
Equally Likely BIRDS
Proportion of Correct Identifications

Assumed RHO = 0.5

CORRELATION COEFFICIENT USED TO GENERATE THE ERROR

		RHO=0.5				RHO=0				RHO=¯0.5			
	Number of WATCHERS:	2	3	4	5	2	3	4	5	2	3	4	5
True Error Dist.	Estimate												
CN(ε=0)	MEDIAN	.96	.97	.98	.99	.97	.97	.99	.99	.96	.96	.99	.95
	NORMMLE	.96	.98	.99	1	.96	.99	1	1	.94	.98	.99	1
	TMLE 1df	.96	.98	.99	.99	.96	.98	.99	1	.93	.97	.98	1
	TMLE 3df	.96	.98	.99	.99	.96	.98	.99	1	.94	.97	.99	1
	TMLE 10df	.96	.98	.99	1	.96	.99	1	1	.95	.98	.99	1
	CNMLE(ε=0.1)	.96	.98	.99	1	.96	.98	1	1	.93	.96	.98	.99
	CNMLE(ε=.25)	.96	.98	.99	1	.95	.98	.99	1	.92	.96	.98	.99
CN(ε=0.1)	MEDIAN	.83	.93	.96	.98	.84	.94	.96	.98	.84	.93	.96	.97
	NORMMLE	.83	.81	.82	.79	.84	.83	.81	.82	.83	.81	.82	.79
	TMLE 1df	.92	.96	.98	.99	.92	.97	.99	1	.90	.95	.97	.98
	TMLE 3df	.92	.97	.98	.99	.93	.97	.99	.99	.90	.95	.98	.98
	TMLE 10df	.90	.96	.98	.99	.91	.97	.98	.99	.90	.94	.98	.98
	CNMLE(ε=0.1)	.93	.97	.99	.99	.93	.97	.99	.99	.90	.94	.97	.98
	CNMLE(ε=.25)	.93	.97	.98	.99	.93	.97	.99	.99	.90	.94	.97	.98
CN(ε=0.25)	MEDIAN	.70	.87	.87	.92	.70	.85	.88	.93	.71	.87	.87	.93
	NORMMLE	.69	.63	.63	.61	.69	.64	.62	.62	.70	.64	.60	.61
	TMLE 1df	.87	.93	.96	.98	.87	.93	.96	.97	.84	.91	.94	.97
	TMLE 3df	.87	.93	.96	.97	.87	.93	.95	.97	.84	.91	.94	.97
	TMLE 10df	.85	.91	.95	.96	.85	.90	.94	.96	.83	.90	.93	.95
	CNMLE(ε=0.1)	.88	.94	.96	.98	.87	.93	.96	.98	.84	.91	.94	.97
	CNMLE(ε=.25)	.88	.94	.97	.98	.87	.93	.96	.98	.84	.90	.94	.97
T 1 d.f.	MEDIAN	.66	.83	.86	.91	.67	.83	.86	.91	.66	.83	.85	.91
	NORMMLE	.67	.71	.69	.69	.66	.67	.66	.66	.63	.63	.63	.63
	TMLE 1df	.81	.88	.92	.95	.80	.87	.92	.95	.77	.86	.90	.93
	TMLE 3df	.81	.88	.93	.95	.80	.87	.92	.95	.77	.85	.90	.92
	TMLE 10df	.79	.87	.92	.94	.77	.86	.90	.94	.75	.84	.89	.91
	CNMLE(ε=0.1)	.79	.87	.91	.94	.78	.85	.90	.93	.76	.85	.88	.91
	CNMLE(ε=.25)	.80	.87	.91	.94	.75	.85	.90	.93	.75	.84	.88	.91

DISTRIBUTION LIST

	NO. OF COPIES
Director of Research Administration (Code 012) Naval Postgraduate School Monterey, CA 93943-5000	1
Library (Code 0142) Naval Postgraduate School Monterey, CA 93943-5000	2
Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
Center for Naval Analyses 2000 Beauregard Street Alexandria, VA 22311	1
Operations Research Centers, Rm E40-164 Massachusetts Institute of Technology Attn: R.C. Larson and J. F. Shapiro Cambridge, MA 02139	1
Library (Code 55) Naval Postgraduate School Monterey, CA 93943-5000	1