APPROXIMATING THE DISTRIBUTION FUNCTION OF TIME-VARYING EQUIPMENT OUTPUTS BY RECURSIVE MOMENT ESTIMATION

By

JOHN CLEVELAND HOOKER

Bachelor of Science
University of Texas at Arlington
Arlington, Texas
1962

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1965

Submitted to the faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
July, 1968

APPROXIMATING THE DISTRIBUTION FUNCTION

OF TIME-VARYING EQUIPMENT OUTPUTS

BY RECURSIVE MOMENT ESTIMATION

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my adviser, Professor A.M. Breipohl for his encouragement, counsel and aid throughout my doctoral studies. I am particularly indebted to him for the original idea which made this thesis possible and for his patient guidance throughout my research. For the long hours of his time he so generously devoted on my behalf, I am truly grateful.

I also wish to thank the other members of my graduate committee; the chairman, Professor H.T. Fristoe, who always had an encouraging word, Professors B.L. Basore and J.L. Folks, who graciously agreed to serve on my committee during the final stages of my research.

To Professor E.K. McLachlan a special thanks is due for his guidance and advice during the early phases of my doctoral studies. I also wish to express my appreciation to two other faculty members, Professors C.M. Bacon and J.M. Walden, who took personal interest in my research and offered valuable comments.

The Sandia Corporation provided the financial support making this research possible which I gratefully acknowledge.

Finally, to my wife, Charliene, and son, Jay, I express my warmest appreciation for their encouragement and patience throughout my graduate studies. To my daughter, Jonna, my thanks for her timely arrival shortly after the draft of this thesis was completed. My wife's devotion, sacrifices, and typing have made this thesis a reality.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Statement of the Problem. The basic concept leading to this thesis is that of deciding on the basis of sample tests whether or not to repair a group of equipment. The engineer is often faced with such a decision as is exemplified in the following hypothetical situation.

A manufacturer of electronic equipment has produced a line of voltage generators which are now in operation. In addition to having produced the equipment, the manufacturer must maintain the generators. The manufacturer cannot continually monitor each generator but must rely on periodic checks of only a fixed number of the generators. From these checks the manufacturer must make either the decision to recall the generators and make the necessary repairs or the decision to leave the equipment in operation.

A very important ingredient in such a decision is the probability that a voltage output is outside specified limits at a specified time. Essentially this probability can be estimated by estimating the probability distribution function of the random variable which describes the voltage outputs at the specified time.

The probability distribution function of a random variable can be approximated as a series expansion of the moments of the random variable. Two series expansions which are common in the literature are the Gram-Charlier series and the Edgeworth series (2). See Appendix A for a discussion of these series.

In most situations the moments are unknown. Therefore, to approximate the probability distribution function using a series expansion the engineer needs estimates of the moments of the random variable which describes the voltage outputs at a specific time. It is this problem of estimation of moments with which this thesis is concerned.

1.2 General Approach to the Problem Solution. "When in doubt, compute the sample moments." In many situations, where observations of a random variable are possible, this advice offered by Deutsch (3) must suffice for the estimation of moments. However in some situations, such as that mentioned in Section 1.1, there may be other sources of information which should be put to use in the estimation of moments. Particularly the engineer may have a priori information about the moments which is derived from the design phase of the system development. In addition a system model which describes the behavior of the equipment outputs may be available.

This effort is directed toward the use of three sources of information in the estimation of moments. These are:

1) A System Model

2) A Priori Information

3) System Observations

Chapter II is devoted to the development of a system model and the subsequent derivation of an augmented moment model which provides the basis for the recursive estimation of moments. The system model used is of the form

$$X_n = C_n X_{n-1} + S_n \qquad (1.2.1)$$

where X_n is the random variable which represents the possible values of the equipment outputs at the time t_n and is given in terms of the previous random variable X_{n-1} and two system random variables C_n and S_n. This model is by no means unique and in many situations is not realistic, but the procedures used in Chapter II to determine the model require that the model have no more than two parameters, e.g., C_n and S_n of Equation 1.2.1. An example is given to illustrate this procedure.

From the system model an augmented moment model is derived. The augmented moment model is

$$\underline{\mu}_n = \underline{A}_n \underline{\mu}_{n-1} + \underline{\mu}_{S_n} \qquad (1.2.2)$$

where $\underline{\mu}_n$ and $\underline{\mu}_{n-1}$ are vectors of moments of X_n and X_{n-1}, respectively, and $\underline{A}_n$ and $\underline{\mu}_{S_n}$ are composed of moments of C_n and S_n. To derive this moment model it is assumed that the random variables C_n, X_{n-1}, and S_n of Equation 1.2.1 are independent.

In Chapter III the augmented moment model is used to develop a recursive moment estimation scheme. In this development the augmented moment model is considered to be a vector valued sample function from a stochastic process. In this framework the moments are random variables which at the time t_n take on specific values. The sample moments are computed from the observations of X_n. The sample moments are then used to determine unbiased data estimates, $\underline{\mu}_n^*$, which are formulated as uniform, minimum variance, minimum risk, unbiased estimators (5), (3), (2).

To develop the recursive moment estimation scheme the works of Papoulis (10) and Kalman (7) on recursive filtering are relied upon very heavily. The unbiased data estimates, $\underline{\mu}_n^*$ are assumed to be noisy observations of the random moments $\underline{\mu}_n$. The estimate $\hat{\underline{\mu}}_n$ is then derived as the linear estimate of $\underline{\mu}_n$ in terms of the observations, $\underline{\mu}_0^*, \underline{\mu}_1^*, \ldots, \underline{\mu}_n^*$, and the a priori estimate $\underline{\mu}_0'$, such that the mean squared error between $\hat{\underline{\mu}}_n$ and $\underline{\mu}_n$ is minimized. Several difficulties arise in the use of $\hat{\underline{\mu}}_n$ as derived in this manner. These difficulties are also presented in Chapter III and an alternative approach, the pseudo-minimum variance recursive moment estimation scheme, is introduced.

Another approach, the Bayesian recursive moment estimation scheme is presented in Appendix D. This approach is an attempt to make use of a reproducing a priori density function in Bayes' Rule to estimate $\underline{\mu}_n$.

To demonstrate the pseudo-minimum variance recursive moment estimation scheme and to investigate its estimating properties a simulating computer program was written. For comparison purposes the Bayesian recursive moment estimation scheme presented in Appendix D was included in this program. Chapter IV discusses the simulation and presents some typical results.

CHAPTER II

DEVELOPMENT OF THE SYSTEM MODELS

2.1 Introduction. This chapter is concerned with the development of a mathematical model of time variation of equipment output and the subsequent derivation of a moment model to be used in the recursive estimation of moments.

A statement of the physical problem is presented and a system model in the form of a first-order linear difference equation is developed. The development of this model is illustrated by an example. The model is then extended to a system model which describes the time variation of the random variables of the system. From the system model an augmented moment model is derived which is a first-order linear vector-matrix difference equation in terms of the moments of the random variables of the system.

A method is suggested by which estimates of the parameters of the augmented moment model can be determined.

2.2 Statement of the Physical Problem. A collection of K pieces of equipment, e.g., a set of 5,000 voltage generators, 1,000 similar radars, or 10,000 amplifiers of the same type, etc., is in operation. Periodically the outputs of k of the K pieces of equipment are observed. The outputs may be the voltage outputs of the generators, the signal-to-noise ratios of the radars, the gains of the amplifiers, etc. See Figure 1. It is desired to estimate at a time t_n the probability that a piece of the equipment is operating outside acceptable limits of operation, i.e., $P[X_n < a \text{ or } X_n > b]$, where X_n is a random variable which represents the possible values of the equipment outputs at the time t_n.

An estimate of $P[X_n < a \text{ or } X_n > b]$ can be made from the k observations at time t_n. This estimate could be determined by constructing the empirical distribution function of X_n from the k observations. Alternatively estimates of the moments of X_n could be formed from the k observations and the distribution function of X_n approximated using a truncated Gram-Charlier or Edgeworth series. See Appendix A. However, neither of these approaches makes use of the k observations at time t_{n-1} or the k observations at t_{n-2}, etc. In order to use the observations made at the n-1 previous sampling times the time varying changes in the equipment outputs must be modeled. The next section develops such a system model.

2.3 Development of the System Model. In order to get some understanding as to how equipment outputs change with time and environment, a lengthy test, a life test, is often performed upon a collection of typical pieces of equipment. Such a test can aid in determining a model of the time varying changes in the equipment outputs.

Consider the ith piece of equipment undergoing a life test. The life test environment is referred to as E_1, and $\underline{E}_{1,n}$ is a vector quantity representing the different constituents—temperature, rate of change of temperature, pressure, radiation, humidity, etc.—of

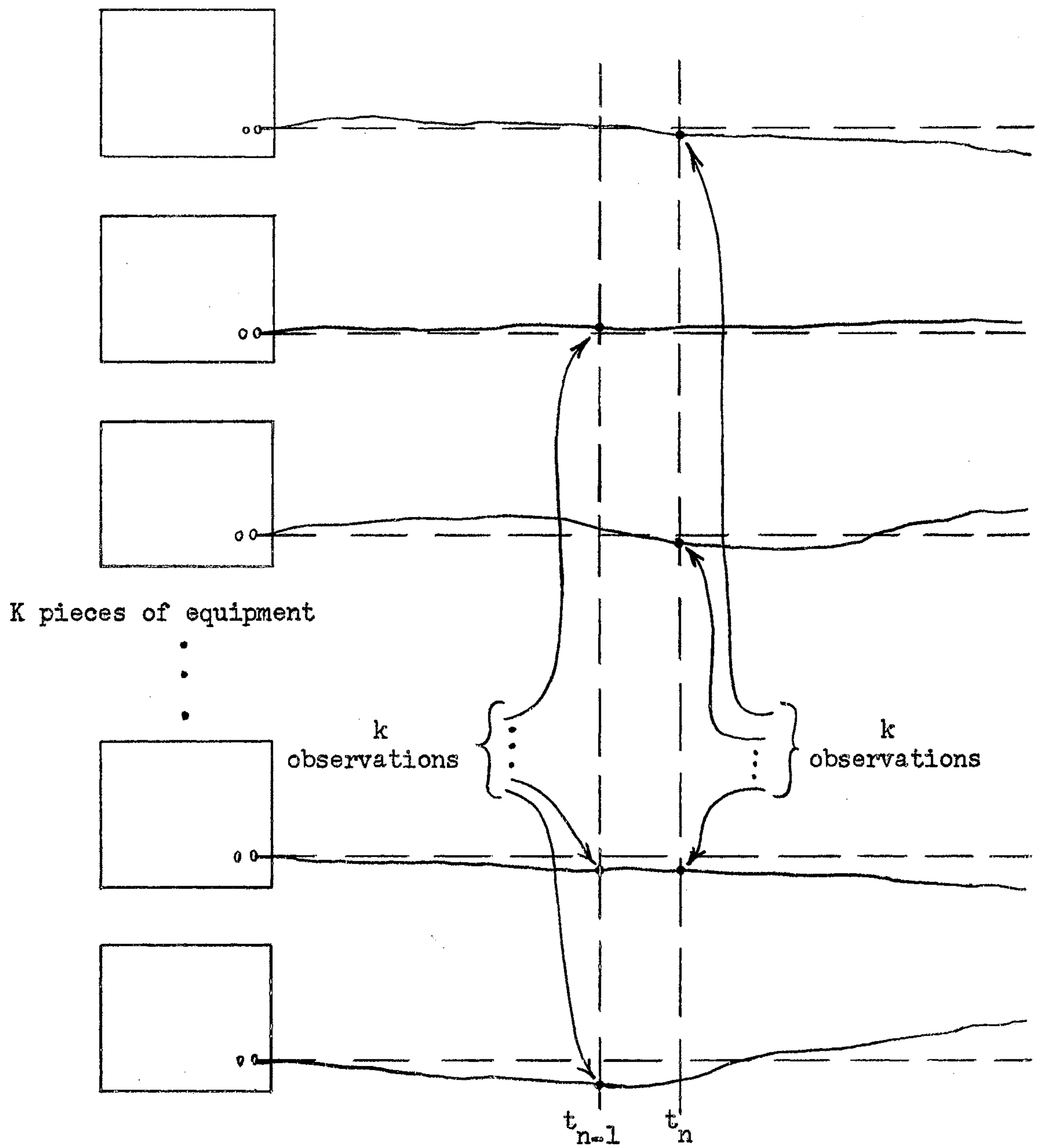

Figure 1. A System of Operating Equipment with Equipment Outputs

environment which cause change in the system outputs from t_{n-1} to t_n in environment E_1. The output of the ith piece of equipment at time t_n is a function of $\underline{E}_{1,n}$, the previous output at time t_{n-1}, and other possible variables, i.e.,

$$x^{(1)}_{i,n} = f(x_{i,n-1}, \ldots, \underline{E}_{1,n}), \qquad n \geq 1 \qquad (2.3.1)$$

where the subscripts i and n denote the ith piece of equipment and the time t_n, respectively, and the superscript (1) denotes the environment, E_1. See Figure 2.

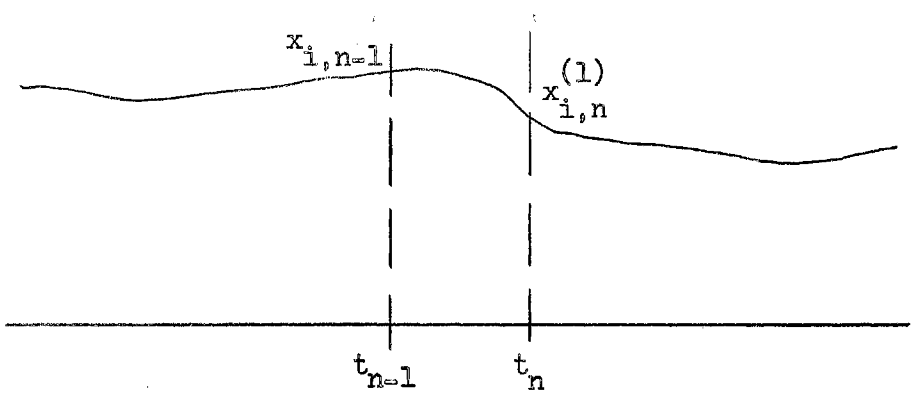

Figure 2. The ith Equipment Output in Life Test, E_1

The change that is observed in the ith equipment's output from t_{n-1} to t_n can be modeled in several ways. With the two observations of the values $x_{i,n-1}$ and $x^{(1)}_{i,n}$ the model of the change is restricted to be in terms of only one unknown. Arbitrarily the change is modeled here as a multiplier, i.e.,

$$x_{i,n}^{(1)} = c_{i,n}^{(1)} x_{i,n-1}, \qquad n \geq 1 \qquad (2.3.2)$$

The multiplier, $c_{i,n}^{(1)}$, represents the change in output which is observed when the equipment is operated in the life test environment, E_1; the change which is caused by the interreaction between the equipment and E_1 during the time from t_{n-1} to t_n. Note that $c_{i,n}^{(1)}$ is uniquely determined from Equation 2.3.2 by the observations of $x_{i,n-1}$ and $x_{i,n}^{(1)}$.

Now assume that the i<u>th</u> piece of equipment is operated in the system environment, E_2, from t_{n-1} to t_n. The output at time t_n is a function of $\underline{E}_{2,n}$, the previous output at time t_{n-1}, and other possible variables, i.e.,

$$x_{i,n}^{(2)} = f(x_{i,n-1},\ldots,\underline{E}_{2,n}), \qquad n \geq 1 \qquad (2.3.3)$$

See Figure 3.

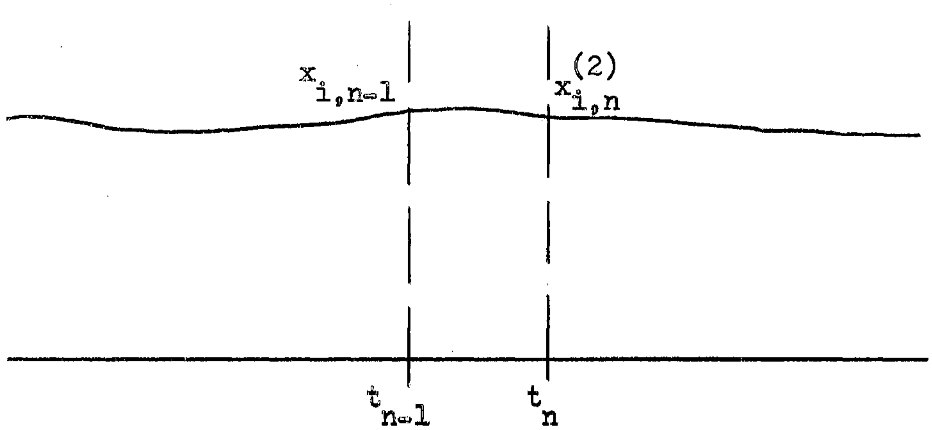

Figure 3. The i<u>th</u> Equipment Output in System Operation, E_2

Expanding $x^{(2)}_{i,n}$ in a Taylor series expansion about the vector $\underline{E}_{1,n}$,

$$x^{(2)}_{i,n} = f(x_{i,n-1},\ldots,\underline{E}_{1,n}) + \nabla f(x_{i,n-1},\ldots,\underline{E}_{1,n})\cdot(\underline{E}_{2,n}-\underline{E}_{1,n}) + \ldots, \quad n \geq 1 \qquad (2.3.4)$$

where "$\cdot$" denotes the dot or inner product and ∇f is the gradient of f with respect to the constituents of $\underline{E}_{\cdot,n}$ and in a sense is a measure of the sensitivity of the equipment output to a change of environment. A first order approximation to $x^{(2)}_{i,n}$ is

$$x^{(2)}_{i,n} \approx x^{(1)}_{i,n} + \nabla f(x_{i,n-1},\ldots,\underline{E}_{1,n})\cdot(\underline{E}_{2,n}-\underline{E}_{1,n}), \qquad n \geq 1$$

From Equation 2.3.2 and letting

$$s^{(2)}_{i,n} = \nabla f(x^{(2)}_{i,n-1},\ldots,\underline{E}_{1,n})\cdot(\underline{E}_{2,n}-\underline{E}_{1,n}),$$

$$x^{(2)}_{i,n} = c^{(1)}_{i,n}x_{i,n-1} + s^{(2)}_{i,n}, \qquad n \geq 1 \qquad (2.3.5)$$

The additive term, $s^{(2)}_{i,n}$, represents the change in the output which is caused by the difference between the system environment, E_2, and the life test environment, E_1.

Equation 2.3.5 is a first-order linear difference equation model of the change that takes place in the <u>ith</u> equipment output under the influence of the system environment, E_2. Although this model is not unique it is a satisfactory model in that it reflects the change that occurs and also the way information about the change is obtained. The multiplier, $c^{(1)}_{i,n}$, reflects the change which can be observed in a life test environment while the additive term, $s^{(2)}_{i,n}$, reflects the additional change which occurs when the equipment is placed in the system environment. From Equations 2.3.2 and 2.3.5 and observations of $x_{i,n-1}$ and $x_{i,n}^{(1)}$ during life test and observations of $x_{i,n-1}$ and $x_{i,n}^{(2)}$ during system operation there are two equations and two unknowns, $c_{i,n}^{(1)}$ and $s_{i,n}^{(2)}$. In which case, $c_{i,n}^{(1)}$ and $s_{i,n}^{(2)}$ can be uniquely determined.

Example. Consider a collection of voltage generators which are to be operated in an environment such that temperature is the only significant constituent. It is desired to model the voltage output, x, of a generator as a function of time under the influence of temperature. It is assumed that the time derivative of the voltage output, x, is proportional to the temperature, T, and the voltage output, i.e.,

$$\frac{dx}{dt} = KT(t)x$$

Solving this differential equation with the conditions $x = x_{n-1}$ at $t = t_{n-1}$ and $x = x_n$ at $t = t_n$

$$\int_{x_{n-1}}^{x_n} \frac{dx}{x} = \int_{t_{n-1}}^{t_n} KT(t)dt$$

$$\ln\left(\frac{x_n}{x_{n-1}}\right) = K \int_{t_{n-1}}^{t_n} T(t)dt$$

$$\frac{x_n}{x_{n-1}} = e^{K \int_{t_{n-1}}^{t_n} T(t)dt}$$

$$x_n = x_{n-1}\, e^{K \int_{t_{n-1}}^{t_n} T(t)dt} \qquad (2.3.6)$$

Thus for this example the change in voltage output from t_{n-1} to t_n is a function of the integral of the temperature, $T(t)$, from t_{n-1} to t_n. The effect is the same as that caused by a constant or average temperature, $T_{a,n}$, such that

$$T_{a,n}(t_n - t_{n-1}) = \int_{t_{n-1}}^{t_n} T(t)dt$$

Therefore Equation 2.3.6 can be expressed as

$$x_n = x_{n-1} \, e^{K \, T_{a,n}(t_n - t_{n-1})} \qquad (2.3.7)$$

If a generator is observed during a life test, environment E_1, and the voltage outputs at times t_{n-1} and t_n are observed to be $x_{i,n-1}$ and $x_{i,n}^{(1)}$, respectively, then the observed change in output from $x_{i,n-1}$ to $x_{i,n}^{(1)}$ is caused by the time integral of temperature,

$$T_{a,n}^{(1)}(t_n - t_{n-1}) = \int_{t_{n-1}}^{t_n} T^{(1)}(t)dt$$

and Equation 2.3.7 becomes

$$x_{i,n}^{(1)} = x_{i,n-1} \, e^{K_i T_{a,n}^{(1)}(t_n - t_{n-1})} \, , \quad n \geq 1 \qquad (2.3.8)$$

Thus the multiplier, $c_{i,n}^{(1)}$, of Equation 2.3.2 is

$$c_{i,n}^{(1)} = e^{K_i T_{a,n}^{(1)}(t_n - t_{n-1})} \qquad (2.3.9)$$

and from the observations of $x_{i,n-1}$ and $x_{i,n}^{(1)}$, $c_{i,n}^{(1)}$ is uniquely determined.

If the generators are placed in the operating system environment,

E_2, the voltage outputs at times t_{n-1} and t_n are $x_{i,n-1}$ and $x^{(2)}_{i,n}$, respectively. Again the change in voltage from $x_{i,n-1}$ to $x^{(2)}_{i,n}$ is caused by the time integral of temperature,

$$T^{(2)}_{a,n}(t_n - t_{n-1}) = \int_{t_{n-1}}^{t_n} T^{(2)}(t)dt$$

and Equation 2.3.7 becomes

$$x^{(2)}_{i,n} = x_{i,n-1}\, e^{K_i T^{(2)}_{a,n}(t_n - t_{n-1})}, \quad n \geq 1 \qquad (2.3.10)$$

Expanding $x^{(2)}_{i,n}$ in terms of the significant environmental effect, $T_{a,n}$, in a Taylor series expansion about the particular value $T^{(1)}_{a,n}$,

$$x^{(2)}_{i,n} = x_{i,n-1}\, e^{K_i T^{(2)}_{a,n}(t_n - t_{n-1})}\Big|_{T^{(2)}_{a,n}=T^{(1)}_{a,n}}$$

$$+ \frac{\partial}{\partial T^{(2)}_{a,n}}\left[x_{i,n-1}\, e^{K_i T^{(2)}_{a,n}(t_n - t_{n-1})}\right]\Big|_{T^{(2)}_{a,n}=T^{(1)}_{a,n}} (T^{(2)}_{a,n} - T^{(1)}_{a,n}) + \cdots$$

$$= x_{i,n-1}\, e^{K_i T^{(1)}_{a,n}(t_n - t_{n-1})}$$

$$+ x_{i,n-1}\, e^{K_i T^{(2)}_{a,n}(t_n - t_{n-1})} K_i(t_n - t_{n-1})\Big|_{T^{(2)}_{a,n}=T^{(1)}_{a,n}} T^{(2)}_{a,n} - T^{(1)}_{a,n} + \cdots$$

$$= x^{(1)}_{i,n} + K_i(t_n - t_{n-1})x_{i,n-1}\, e^{K_i T^{(1)}_{a,n}(t_n - t_{n-1})}(T^{(2)}_{a,n} - T^{(1)}_{a,n}) + \cdots$$

Taking the first order approximation of $x^{(2)}_{i,n}$ from the Taylor series expansion

$$x^{(2)}_{i,n} = c^{(1)}_{i,n}\, x_{i,n-1} + s^{(2)}_{i,n}\ , \quad n \geq 1 \qquad (2.3.11)$$

where $c^{(1)}_{i,n}$ is given by Equation 2.3.9 and

$$s^{(2)}_{i,n} = K_i(t_n - t_{n-1}) x_{i,n-1}\, e^{K_i T^{(1)}_{a,n}(t_n - t_{n-1})} (T^{(2)}_{a,n} - T^{(1)}_{a,n}) \qquad (2.3.12)$$

Equations 2.3.9 and 2.3.12 indicate for this example how the model parameters, $c^{(1)}_{i,n}$ and $s^{(2)}_{i,n}$, are related to the environmental constituents, $T^{(1)}_{a,n}$ and $T^{(2)}_{a,n}$, which cause change during life test and system operation. Equation 2.3.11 is the desired model for the voltage output of the ith piece of equipment.

As indicated prior to the above example if the values of $x_{i,n-1}$ and $x^{(1)}_{i,n}$ are observed during the life test and the corresponding values of $x_{i,n-1}$ and $x^{(2)}_{i,n}$ are observed during system operation, then using Equations 2.3.2 and 2.3.5 the two parameters, $c^{(1)}_{i,n}$ and $s^{(2)}_{i,n}$, can be uniquely determined. It is implicit here that the same unit is in system operation as in life test and that it is identifiable in both environments.

When a collection of equipment is in operation only k samples are taken of the total number of K units in operation, $k < K$. Usually M, the number of units observed during the life test, is considerably less than K, $M < K$, and not necessarily equal to k, $M \neq k$. Then $c^{(1)}_{i,n}$, $i = 1,2,\ldots,M$ can be determined. Also, unless the corresponding M units are observed during system operation, $s^{(2)}_{i,n}$ can not be determined. Instead from the M observations during life test and the k observations during system operation at each sampling time, t_n, only estimates of the population of equipments can be determined. A discussion of how estimates about $c_{i,n}^{(1)}$ and $s_{i,n}^{(2)}$ are obtained is presented in Section 2.5.

Since at best only estimates can be determined it is useful to extend the model, Equation 2.3.5, to a model relating the random variable X_n to the random variable X_{n-1} where X_n is the random variable representing the possible values of the K equipment outputs at the time t_n. The extended model is a first-order linear difference equation and is given by

$$X_n = C_n X_{n-1} + S_n , \quad n \geq 1 \qquad (2.3.13)$$

where C_n is a random variable which represents the possible values of $c_{i,n}$, $i = 1,\ldots,K$, and S_n is a random variable which represents the possible values of $s_{i,n}$, $i = 1,\ldots,K$. Note that the superscripts "(1)" and "(2)" representing the environments have been omitted in Equation 2.3.13. It is implicit in the remainder of the study that Equation 2.3.13 refers to the system operation environment.

2.4 Development of the Augmented Moment Model. To estimate the probability that a piece of equipment is operating outside acceptable limits of operation, $P[X_n < a \text{ or } X_n > b]$, a series expansion of the probability distribution function of X_n in terms of the moments of X_n can be used. See Appendix A. In this section an augmented moment model is derived from the system model developed in Section 2.3. This augmented moment model becomes the means whereby the estimates of the moments of X_i, $0 \leq i \leq n-1$, can be used in the estimation of the moments of X_n.

Although the augmented moment model and the techniques of moment estimation developed in this study can be extended to higher order moments, the estimation of only the first three moments of X_n is presented here. Of course with only k observations of X_n in many cases the estimates will become less accurate in a mean squared error sense as estimation of higher order moments is attempted.

The following notation will be used throughout this thesis. The first moment of a random variable is the mean or expectation of that random variable, i.e.,

$$\mu_{1,n} = E\{X_n\} \quad , \quad \mu_{1C_n} = E\{C_n\} \quad , \quad \text{and} \quad \mu_{1S_n} = E\{S_n\}$$

All other moments are central moments, i.e.,

$$\mu_{r,n} = E\{[X_n - \mu_{1,n}]^r\} \quad , \quad \mu_{rC_n} = E\{[C_n - \mu_{1C_n}]^r\} \quad ,$$

$$\mu_{rS_n} = E\{[S_n - \mu_{1S_n}]^r\} \quad , \quad r = 2, 3, \ldots$$

Assuming that in the system model, Equation 2.3.13, C_n, X_{n-1}, and S_n are independent random variables* the mean of X_n is

$$\mu_{1,n} = E\{X_n\} = E\{C_n X_{n-1} + S_n\} = E\{C_n\}E\{X_{n-1}\} + E\{S_n\}$$

$$= \mu_{1C_n} \mu_{1,n-1} + \mu_{1S_n} \quad , \qquad n \geq 1 \qquad (2.4.1)$$

---

*The assumption of independence may not always reflect the true circumstances. In the example presented in Section 2.3, Equation 2.3.12 indicates that X_{n-1} and S_n are very definitely dependent.

Similarly the second central moment, variance, of X_n is

$$\mu_{2,n} = E\{[X_n-\mu_{1,n}]^2\} = E\{[C_nX_{n-1}+S_n-\mu_{1C_n}\mu_{1,n-1}-\mu_{1S_n}]^2\}$$

$$= [\mu_{2C_n}+\mu_{1C_n}^2]\,\mu_{2,n-1} + \mu_{2C_n}\mu_{1,n-1}^2 + \mu_{2S_n}, \quad n \geq 1 \quad (2.4.2)$$

And the third central moment of X_n is

$$\mu_{3,n} = E\{[X_n-\mu_{1,n}]^3\} = E\{[C_nX_{n-1}+S_n-\mu_{1C_n}\mu_{1,n-1}-\mu_{1S_n}]^3\}$$

$$= [\mu_{3C_n}+3\mu_{2C_n}\mu_{1C_n}+\mu_{1C_n}^3]\,\mu_{3,n-1}$$

$$+ [3\mu_{3C_n}+6\mu_{2C_n}\mu_{1C_n}]\,\mu_{2,n-1}\mu_{1,n-1}+\mu_{3C_n}\mu_{1,n-1}^3+\mu_{3S_n},$$

$$n \geq 1 \quad (2.4.3)$$

The detailed developments of Equations 2.4.1, 2.4.2, and 2.4.3 are given in Appendix B.

Equations 2.4.1, 2.4.2, and 2.4.3 indicate a non-linear relationship between the moments of X_n and the moments of X_{n-1}. For example, Equation 2.4.2 gives $\mu_{2,n}$ as a function of $\mu_{2,n-1}$ and the square of $\mu_{1,n-1}$. By using $\mu_{1,n-1}^2$, $\mu_{1,n-1}^3$, and $\mu_{2,n-1}\mu_{1,n-1}$ as auxiliary variables a linear form can be construed. In this case Equation 2.4.2 gives $\mu_{2,n}$ as a linear function of $\mu_{2,n-1}$ and $\mu_{1,n-1}^2$. With these auxiliary variables an augmented moment vector $\underline{\mu}_n$, can be defined as

$$\underline{\mu}_n = \begin{bmatrix} \mu_{1,n} \\ \mu_{2,n} \\ \mu_{3,n} \\ \mu_{1,n}^2 \\ \mu_{1,n}^3 \\ \mu_{2,n}\mu_{1,n} \end{bmatrix}, \quad n \geq 1 \quad (2.4.4)$$

The variation of this augmented moment vector with n(time) can be written in the form of a first-order linear vector-matrix difference equation,

$$\underline{\mu}_n = \underline{A}_n \underline{\mu}_{n-1} + \underline{\mu}_{S_n} , \quad n \geq 1 \qquad (2.4.5)$$

where $\underline{A}_n$ is given in Figure 4 and

$$\underline{\mu}_{S_n} = \begin{bmatrix} \mu_{1S_n} \\ \mu_{2S_n} \\ \mu_{3S_n} \\ \mu^2_{1S_n} \\ \mu^3_{1S_n} \\ \mu_{2S_n}\mu_{1S_n} \end{bmatrix}$$

Equation 2.4.5 is the derived augmented moment model.

<u>2.5 Estimating the Moments of C_n and S_n.</u> In order to use the augmented moment model, Equation 2.4.5, the moments of C_n and S_n must be known. Unless the statistical properties of the changes due to the life test environmental stresses, C_n, and those due to the system environmental stresses, S_n, are known these moments will be unknown. In this section a method of determining estimates of the moments of C_n and S_n is suggested.

In Section 2.3 the change in output was observed in a controlled life test. This change was represented by the difference equation

$$x^{(1)}_{i,n} = c^{(1)}_{i,n} \, x_{i,n-1} , \qquad i = 1,\ldots,M, \quad n \geq 1 \qquad (2.3.2)$$

$$
\underline{A}_n = \begin{bmatrix}
\mu_{1C_n} & 0 & 0 & 0 & 0 & 0 \\
0 & [\mu_{2C_n}+\mu_{1C_n}^2] & 0 & \mu_{2C_n} & 0 & 0 \\
0 & 0 & [\mu_{3C_n}+3\mu_{2C_n}\mu_{1C_n}+\mu_{1C_n}^3] & 0 & \mu_{3C_n} & [3\mu_{3C_n}+6\mu_{2C_n}\mu_{1C_n}] \\
2\mu_{1C_n}\mu_{1S_n} & 0 & 0 & \mu_{1C_n}^2 & 0 & 0 \\
3\mu_{1C_n}\mu_{1S_n}^2 & 0 & 0 & 3\mu_{1C_n}^2\mu_{1S_n} & \mu_{1C_n}^3 & 0 \\
\mu_{1C_n}\mu_{2S_n} & [\mu_{2C_n}+\mu_{1C_n}^2]\mu_{1S_n} & 0 & \mu_{2C_n}\mu_{1S_n} & \mu_{2C_n}\mu_{1C_n} & [\mu_{2C_n}\mu_{1C_n}+\mu_{1C_n}^3]
\end{bmatrix}
$$

Figure 4. $\underline{A}_n$

It was also noted that $c^{(1)}_{i,n}$ is uniquely determined by observations of $x_{i,n-1}$ and $x^{(1)}_{i,n}$, i.e.,

$$c_{i,n} = \frac{x^{(1)}_{i,n}}{x_{i,n-1}} \quad , \qquad i = 1,\ldots,M, \quad n \geq 1 \qquad (2.5.1)$$

In a controlled life test where M pieces of equipment are observed periodically and $c^{(1)}_{i,n}$ for the $i\underline{th}$ piece of equipment is determined by Equation 2.5.1, an estimate of the mean of C_n is

$$\mu^*_{1C_n} = \frac{1}{M} \sum_{i=1}^{M} c^{(1)}_{i,n} \quad , \qquad n \geq 1 \qquad (2.5.2)$$

Similarly estimates of higher order central moments of C_n are

$$\mu^*_{rC_n} = \frac{1}{M} \sum_{i=1}^{M} (c^{(1)}_{i,n} - \mu^*_{1C_r})^r \quad , \qquad r = 2,3,\ldots, \quad n \geq 1 \quad (2.5.3)$$

or, using unbiased estimates as will be done in this study when possible, the unbiased estimates of the second and third central moments are

$$\mu^*_{2C_n} = \frac{1}{(M-1)} \sum_{i=1}^{M} (c^{(1)}_{i,n} - \mu^*_{1C_n})^2 \quad , \quad n \geq 1 \qquad (2.5.4)$$

$$\mu^*_{3C_n} = \frac{M}{(M-1)(M-2)} \sum_{i=1}^{M} (c^{(1)}_{i,n} - \mu^*_{1C_n})^3 \quad , \quad n \geq 1 \qquad (2.5.5)$$

See Appendix C for a development of unbiased estimates. Estimates of higher order moments of C_n can be determined in a similar fashion.

Thus, estimates of the moments of C_n can be determined from a life test of M pieces of equipment.

Estimates of the moments of S_n are more difficult to obtain. Since S_n models the change due to the difference in the environment of the life test and that of the system operation, accurate knowledge of the moments of S_n is difficult to obtain prior to actual observation of the system in operation. Instead the estimates of the moments of S_n must represent the a priori knowledge of how one believes the system environment affects the system.

The estimate of the mean of S_n will often be zero because a life test is often designed to simulate the actual system environment. The estimate of the second moment, variance, of S_n should reflect the uncertainty that one has in the effect of the system environment. If the uncertainty as to the difference between the life test environment and the system environment is great, μ_{2S_n} should be large. If one's confidence is high that the change due to the system environment is not very different from that observed in the life test, μ_{2S_n} should be small. Since little else can be said about the environmental changes, it is plausible to assume that S_n is normally distributed.

CHAPTER III

RECURSIVE MOMENT ESTIMATION

3.1 Restatement of the Problem and Introduction to Recursive Moment Estimation. In Section 2.2 the problem is given as one of estimating $P[X_n < a \text{ or } X_n > b]$ where X_n is a random variable representing the possible values of the equipment outputs at the time t_n. The estimate is to be formed by making estimates of moments of X_n and then approximating the distribution function of X_n using a truncated Gram-Charlier or Edgeworth series. The moments of X_n will be estimated using the augmented moment model, Equation 2.4.5. In this section the problem is restated and more definitively formalized in terms of the augmented moment model.

The augmented moment model of Equation 2.4.5 is a vector valued function which describes the way the moments, $\underline{\mu}_n$, Equation 2.4.4, of X_n, Equation 2.3.13, vary with time (n). In this respect Equation 2.4.5 can be thought of as a vector valued sample function from a stochastic process.

To develop the concept of the augmented moment model as a sample function consider the first moment, the mean, of X_n. From Equation 2.4.5 the mean of X_n is

$$\mu_{1,n} = \mu_{1C_n} \mu_{1,n-1} + \mu_{1S_n} , \qquad n \geq 1 \qquad (3.1.1)$$

Equation 3.1.1 is a function describing how the mean of X_n varies with time; at least how it changes from one sampling time to another. See Figure 5.

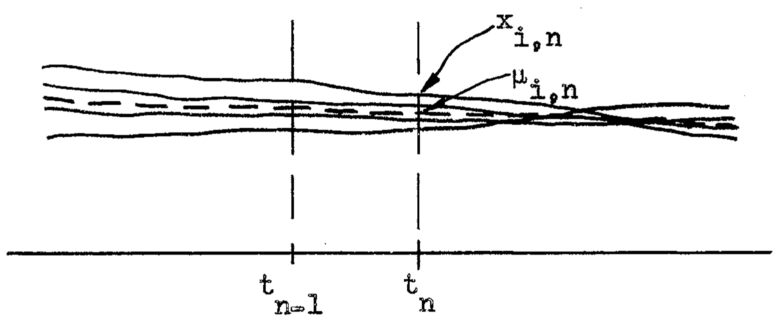

Figure 5. $\mu_{1,n}$ versus time (n)

Figure 5 depicts the function $\mu_{1,n}$ in relation to the equipment outputs, $x_{i,n}$, $i = 1,\ldots,K$. The system of K pieces of equipment generates the function, $\mu_{1,n}$, a sample function from a stochastic process. Other sample functions of means from this stochastic process are generated in the same manner. To elaborate, if there are other systems of equipment in operation, similar to the system of K pieces of equipment which generate $\mu_{1,n}$, then these systems also generate sample functions like Equation 3.1.1. If there are no other systems then a hypothetical stochastic process can be assumed from which the one sample function, $\mu_{1,n}$, is realized.

If time is fixed at t_n the value of the sample function, $\mu_{1,n}$, is fixed at a constant which is the mean of X_n. Similarly for all the sample functions of the means of the stochastic process; if time is fixed at t_n then each sample function takes on a constant value. If the stochastic process is considered as a whole and time is fixed at t_n there results a random variable, $\mu_{1,n}$, which represents all possible mean values at time t_n. One realization of this random variable is the value of $\mu_{1,n}$, the mean of X_n. Other realizations are the means of the other systems of equipment at t_n, either actual or hypothetical.

In a similar manner as that described above, each element of the augmented moment vector, $\underline{\mu}_n$, i.e., $\mu_{2,n}$, $\mu_{3,n}$, etc., can be considered as a sample function from a stochastic process. Thus the augmented moment model can be thought of as a vector valued sample function from a vector valued stochastic process.

In the context of the augmented moment model as a vector valued sample function from a stochastic process the problem of estimation of moments becomes one of estimating the value of the sample function $\underline{\mu}_n$ at each sampling time, t_n.

In the next section $\underline{\mu}_n^*$, the best estimate of $\underline{\mu}_n$ from the k observations of the random variable X_n, is developed. The criterion for "best" is taken to be minimum mean squared error. $\underline{\mu}_n^*$ is referred to as the unbiased data moment estimate. $\underline{\mu}_n^*$ is the best estimate of $\underline{\mu}_n$ given only the k observations of X_n.

In Section 3.3 a scheme of recursive moment estimation is developed. The procedure is a systematic method of determining $\hat{\underline{\mu}}_n$, the best estimate given the initial estimate, $\underline{\mu}_0'$, and the n + 1 data moment estimates, $\underline{\mu}_i^*$, i = 0,1,...,n. Section 3.4 discusses several difficulties in this recursive moment estimation approach and

Section 3.5 presents an alternative, pseudo-minimum variance recursive moment estimation procedure.

3.2 Unbiased Data Moment Estimates. When the stochastic process introduced in Section 3.1 is halted in time at t_n there results a random vector, $\underline{\mu}_n$, which represents all possible values of the moment vector. The vector valued sample function, the augmented moment model of Equation 2.4.5, takes on one possible value of this random vector, $\underline{\mu}_n$, the moment vector, Equation 2.4.4, of X_n.

It is desired in this section to determine the best estimate of the moment vector $\underline{\mu}_n$ given only the k observations of X_n.

From the k independent samples, $x_{1,n}, x_{2,n}, \ldots, x_{k,n}$, of X_n taken at time t_n, the sample moments of X_n are given by

$$m_{1,n} = \frac{1}{k} \sum_{i=1}^{k} x_{i,n} \quad , \quad m_{r,n} = \frac{1}{k} \sum_{i=1}^{k} (x_{i,n} - m_{1,n})^r \quad , \quad r = 2, 3, \ldots \tag{3.2.1}$$

These sample moments are estimates of the moments, $\mu_{1,n}, \mu_{r,n}$, $r = 2, 3, \ldots$, respectively. However all except $m_{1,n}$ are biased estimates. For example, $m_{2,n}$ is a biased estimate of $\mu_{2,n}$, for

$$E\{m_{2,n} | \underline{\mu}_{na}\} = E\{\frac{1}{k} \sum_{i=1}^{k} (x_{i,n} - m_{1,n})^2 | \underline{\mu}_{na}\} = \frac{k-1}{k} \mu_{2,n} \tag{3.2.2}$$

where $E\{\cdot | \underline{\mu}_{na}\}$ means the expectation given all the moments of X_n. The bias in the estimate of $\mu_{2,n}$ can be removed if, as can be deduced from Equation 3.2.2, the estimate $\mu^*_{2,n}$ given by

$$\mu^*_{2,n} = \frac{k}{k-1} m_{2,n} \tag{3.2.3}$$

is used instead of $m_{2,n}$. $\mu^*_{2,n}$ is obviously an unbiased estimate of

$\mu_{2,n}$ and is referred to as an unbiased data estimate.

Unbiased data estimates are used as estimates of the moments, $\underline{\mu}_n$. The vector $\underline{\mu}_n^*$ is defined as

$$\underline{\mu}_n^* = \begin{bmatrix} \mu_{1,n}^* \\ \mu_{2,n}^* \\ \mu_{3,n}^* \\ \mu_{1,n}^{2*} \\ \mu_{1,n}^{3*} \\ (\mu_{2,n}\mu_{1,n})^* \end{bmatrix}, \quad n \geq 1 \tag{3.2.4}$$

where each element of $\underline{\mu}_n^*$ is an unbiased estimate of the corresponding element of $\underline{\mu}_n$. In fact, the elements of $\underline{\mu}_n^*$ are the UMV-RUE's (uniform, minimum variance, minimum risk, unbiased estimators) of the elements of $\underline{\mu}_n$ and are given in terms of the sample moments, Equation 3.2.1, by Equations 3.2.5. See Appendix C for the derivation of UMV-RUE's.

$$\mu_1^* = m_1 \tag{3.2.5a}$$

$$\begin{bmatrix} \mu_2^* \\ \mu_1^{2*} \end{bmatrix} = \frac{1}{k-1} \begin{bmatrix} k & 0 \\ -1 & k-1 \end{bmatrix} \begin{bmatrix} m_2 \\ m_1^2 \end{bmatrix} \tag{3.2.5b}$$

$$\begin{bmatrix} \mu_3^* \\ (\mu_2\mu_1)^* \\ \mu_1^{3*} \end{bmatrix} = \frac{1}{(k-1)(k-2)} \begin{bmatrix} k^2 & 0 & 0 \\ -k & k(k-2) & 0 \\ 2 & -3(k-2) & (k-1)(k-2) \end{bmatrix} \begin{bmatrix} m_3 \\ m_2 m_1 \\ m_1^3 \end{bmatrix} \tag{3.2.5c}$$

Note that in Equations 3.2.5 the second subscript, n, has been omitted for simplicity of presentation. For the remainder of this section n will be omitted often. It is implicit that all random variables, data, moments, and estimates have the same time correspondence unless otherwise indicated.

A measure of the goodness of one of the elements of $\underline{\mu}_n^*$ as an estimate of the corresponding element of $\underline{\mu}_n$ is the mean squared error, e.g., $E\{(\mu^*_{2,n}-\mu_{2,n})^2\}$ is a measure of the goodness of $\mu^*_{2,n}$ as an estimate of $\mu_{2,n}$. It is necessary in the following chapters to have, in addition to the mean squared errors of the elements of $\underline{\mu}_n^*$, the error covariances of the elements of $\underline{\mu}_n^*$, e.g., $E\{(\mu^*_{2,n}-\mu_{2,n})(\mu^*_{3,n}-\mu_{3,n})\}$. It is convenient to place the mean squared errors and error covariances together in the error covariance matrix of $\underline{\mu}_n^*$ given by

$$\Psi_n^* = E\{(\underline{\mu}_n^*-\underline{\mu}_n)(\underline{\mu}_n^*-\underline{\mu}_n)^T\} \ .$$

However the error covariance matrix of the estimate $\underline{\mu}_n^*$ is unknown and in its place an estimate must be used.

To aid in determining an estimate of the error covariance matrix of $\underline{\mu}_n^*$, consider the covariance matrix, $\Psi_n^{*\prime}$, of $\underline{\mu}_n^*$ given the moments of X_n,

$$\Psi_n^{*\prime} = E\{(\underline{\mu}_n^*-\underline{\mu}_n)(\underline{\mu}_n^*-\underline{\mu}_n)^T|\underline{\mu}_{na}\}$$

Two typical elements of $\Psi_n^{*\prime}$ are expanded.

One element of $\Psi_n^{*\prime}$ is the variance $\sigma^2_{\mu_1^*}$ given by

$$\sigma^2_{\mu_1^*} = E\{(\mu_1^*-\mu_1)^2|\underline{\mu}_{na}\} = E\{\mu_1^{*2}|\underline{\mu}_{na}\}-\mu_1^2$$

From Equations 3.2.5

$$\sigma^2_{\mu_1^*} = E\{m_1^2 | \underline{\mu}_{na}\} - \mu_1^2$$

Using Equations C.4.3

$$\sigma^2_{\mu_1^*} = \frac{1}{k}\,\mu_2 + \mu_1^2 - \mu_1^2$$

$$= \frac{1}{k}\,\mu_2 \qquad (3.2.6)$$

Another element of $\Psi_n^{*\prime}$ is the covariance $\sigma_{\mu_1^*\mu_2^*}$ given by

$$\sigma_{\mu_1^*\mu_2^*} = E\{(\mu_1^*-\mu_1)(\mu_2^*-\mu_2) | \underline{\mu}_{na}\} = E\{\mu_1^*\mu_2^* | \underline{\mu}_{na}\} - \mu_1\mu_2$$

which from Equations 3.2.5 is

$$\sigma_{\mu_1^*\mu_2^*} = \frac{k}{k-1}\,E\{m_1 m_2 | \underline{\mu}_{na}\} - \mu_1\mu_2$$

and using Equations C.4.3

$$\sigma_{\mu_1^*\mu_2^*} = \frac{k}{k-1}\left[\frac{k-1}{k^2}\,\mu_3 + \frac{k-1}{k}\,\mu_2\mu_1\right] - \mu_1\mu_2$$

$$= \frac{1}{k}\,\mu_3 \qquad (3.2.7)$$

Similar developments of the remaining elements of $\Psi_n^{*\prime}$ result in Equations C.4.4. Since $\Psi_n^{*\prime}$ is a (6x6) symmetric matrix it has 21 distinct elements. These 21 elements are those given in Equations C.4.4.

Equations 3.2.6, 3.2.7, and C.4.4 all indicate the conclusion to be drawn here; that is, that the covariance matrix, $\Psi_n^{*\prime}$, is a function of the unknown moments of X_n, and therefore is unknown.

Since $\underline{\mu}_n^*$ is the UMV-RUE of $\underline{\mu}_n$ it does minimize the variances of $\underline{\mu}_n^*$ which are the diagonal elements of $\Psi_n^{*\prime}$ therefore the measures of goodness chosen are the diagonal terms of $\hat{\Psi}_n^*$, the UMV-RUE of $\Psi_n^{*\prime}$, which is determined strictly from the data observed. Each element of $\hat{\Psi}_n^*$ is the UMV-RUE of the corresponding element of $\Psi_n^{*\prime}$. For example, $(\sigma^2_{\mu_1^*})^*$ is the UMV-RUE of $\sigma^2_{\mu_1^*}$ and $(\sigma_{\mu_1^*\mu_2^*})^*$ is the UMV-RUE of $\sigma_{\mu_1^*\mu_2^*}$.

Equations C.4.4 indicate that $\Psi_n^{*\prime}$ is a linear function of moments and products of moments. In fact $\Psi_n^{*\prime} = \underline{B}\ \underline{Z}$, where $\underline{B}$ is a matrix of constants and $\underline{Z}$ is a matrix of moments and products of moments. $\underline{B}$ and $\underline{Z}$ are implicitely defined by Equations C.4.4. The following theorem concludes that $\hat{\Psi}_n^*$, the UMV-RUE of $\Psi_n^{*\prime}$, is the same linear function with the moments and products of moments replaced by their UMV-RUE's. This theorem is essentially proven as part of Theorem 2.7, p. 60, of Fraser (5).

<u>Theorem 3.2.1.</u> Given 1) a random variable X having the absolutely continuous distribution, $F_X(x;\underline{\theta})$ on R^1, the real line, 2) $t(\underline{x})$, a complete and sufficient statistic for $\{F_X(x;\underline{\theta}) \mid \underline{\theta} \in \Theta\}$, 3) $\Psi_n^{*\prime} = \underline{B}\ \underline{Z}$, where $\underline{B}$ is a matrix of constants and $\underline{Z}$ is a matrix of moments and products of moments of X, and 4) $\underline{Z}^*$, the matrix $\underline{Z}$ with each element replaced by its UMV-RUE, then $\hat{\Psi}_n^*$, the UMV-RUE of $\Psi_n^{*\prime}$, is

$$\hat{\Psi}_n^* = \underline{B}\ \underline{Z}^*$$

<u>Example.</u> Consider the variance $\sigma^2_{\mu_2^*}$. From Equation C.4.4

$$\sigma^2_{\mu_2^*} = \frac{1}{k}\mu_4 - \frac{(k-3)}{k(k-1)}\mu_2^2$$

Let

$$f_1'(\underline{x}) = x_1^4 - 4x_1^3x_2 + 6x_1^2x_2x_3 - 3x_1x_2x_3x_4$$

and

$$f_2'(\underline{x}) = x_1^2x_2^2 - 2x_1^2x_2x_3 + x_1x_2x_3x_4$$

then

$$E\{f_1'(\underline{X})\} = E\{X_1^4\} - 4E\{X_1^3\}\ E\{X_2\} + 6E\{X_1^2\}\ E\{X_2\}\ E\{X_3\}$$

$$-3E\{X_1\}\ E\{X_2\}\ E\{X_3\}\ E\{X_4\}$$

$$= \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4$$

$$= \mu_4$$

and

$$E\{f_2'(\underline{X})\} = E\{X_1^2\}E\{X_2^2\} - 2E\{X_1^2\}E\{X_2\}E\{X_3\} + E\{X_1\}E\{X_2\}E\{X_3\}E\{X_4\}$$

$$= \alpha_2^2 - 2\alpha_2\alpha_1^2 + \alpha_1^4$$

$$= \mu_2^2$$

where $\alpha_r = E\{X^r\}$, $r = 1, 2, \ldots$

Let

$$h(\underline{x}) = \frac{1}{k}\ f_1'(\underline{x}) - \frac{(k-3)}{k(k-1)}\ f_2'(\underline{x})$$

then

$$E\{h(\underline{X})\} = \frac{1}{k}\ E\{f_1'(\underline{X})\} - \frac{(k-3)}{k(k-1)}\ E\{f_1'(\underline{X})\}$$

$$= \frac{1}{k}\ \mu_4 - \frac{(k-3)}{k(k-1)}\ \mu_2^2$$

$(\sigma^2_{\mu_2 *})^*$, the UMV-RUE of $\sigma^2_{\mu_2 *}$, is the conditional expectation of $h(\underline{x})$ given the complete and sufficient statistic t.

$$(\sigma^2_{\mu_2^*})^* = E\{h(\underline{X})\,|\,t\} = E\{[\frac{1}{k} f_1'(\underline{X}) - \frac{(k-3)}{k(k-1)} f_2'(\underline{X})]\,|\,t\}$$

$$= \frac{1}{k} E\{f_1'(\underline{X})\,|\,t\} - \frac{(k-3)}{k(k-1)} E\{f_2'(\underline{X})\,|\,t\}$$

$$= \frac{1}{k} \mu_4^* - \frac{(k-3)}{k(k-1)} \mu_2^{2*}$$

where μ_4^* and μ_2^{2*} are the UMV-RUE's of μ_4 and μ_2^2, respectively. See Appendix C for a more comprehensive development of UMV-RUE's.

The UMV-RUE's needed in $\hat{\Psi}_n^*$ are given in Equations C.4.5 in terms of the sample moments, m_r, $r = 1,2,\ldots$. From the k observations of X, $x_1,x_2,\ldots,x_k$, the sample moments m_r, $r = 1,2,\ldots$, are calculated according to Equations 3.2.1. Then the estimate $\underline{\mu}_n^*$ is determined from Equations 3.2.5 and $\hat{\Psi}_n^*$, the estimate of the error covariance matrix of μ_n^*, is calculated using Equations C.4.4 by replacing the moments and products of moments in Equations C.4.4 by their corresponding UMV-RUE's of Equations C.4.5.

<u>3.3 Development of the Recursive Moment Estimates.</u> In the previous section $\underline{\mu}_n^*$, the best estimate of $\underline{\mu}_n$ given only the k observations of X_n, was developed. If the k observations of X_n were the only information available pertaining to $\underline{\mu}_n$ then $\underline{\mu}_n^*$ would have to suffice as the best estimate of $\underline{\mu}_n$. However at the n previous sampling times the estimates $\underline{\mu}_i^*$ of $\underline{\mu}_i$ have been made from observations of X_i, $0 \le i \le n-1$. In addition a priori knowledge may be available from which the estimate $\underline{\mu}_0'$ of $\underline{\mu}_0$ is derived. In this section a recursive estimation procedure is developed for which the estimate $\hat{\underline{\mu}}_n$ is the linear estimate of $\underline{\mu}_n$ in terms of $\hat{\underline{\mu}}_0,\underline{\mu}_1^*,\ldots,\underline{\mu}_n^*$, and I, the identity matrix, which minimizes tr $E\{[\underline{\mu}_n-\hat{\underline{\mu}}_n][\underline{\mu}_n-\hat{\underline{\mu}}_n]^T\}$.

$\hat{\underline{\mu}}_0$ is taken to be the linear estimate of $\underline{\mu}_0$ in terms of $\underline{\mu}_0'$ and $\underline{\mu}_0^*$ which minimizes tr $E\{[\underline{\mu}_0-\hat{\underline{\mu}}_0][\underline{\mu}_0-\hat{\underline{\mu}}_0]^T\}$

The vector $\hat{\underline{\mu}}_n$ is

$$\hat{\underline{\mu}}_n = \begin{bmatrix} \hat{\mu}_{1,n} \\ \hat{\mu}_{2,n} \\ \hat{\mu}_{3,n} \\ \hat{\mu}^2_{1,n} \\ \hat{\mu}^3_{1,n} \\ (\mu_{2,n}\hat{\mu}_{1,n}) \end{bmatrix}, \quad n \geq 1 \qquad (3.3.1)$$

where each element of $\hat{\underline{\mu}}_n$ is an estimate of the corresponding element of $\underline{\mu}_n$.

In terms of the stochastic process introduced in Section 3.1 this section is concerned with the development of an estimate of the value of $\underline{\mu}_n$ given $\hat{\underline{\mu}}_0, \underline{\mu}_1^*, \ldots, \underline{\mu}_n^*$. It is implicit from the context whether $\underline{\mu}_n$ is the random vector or a value of the random vector.

The following development of $\hat{\underline{\mu}}_n$ parallels the proof of the theorem on recursive filtering given by Papoulis (10). However the assumptions here are less restrictive than those of Papoulis. Whereas Papoulis deals only with the estimation of a random variable with zero mean, the estimate of a vector of random variables with non-zero means is developed here.

To the augmented moment model of Equation 2.4.5 an observation equation is attached. The augmented moment model and the observation equation are

$$\underline{\mu}_n = \underline{A}_n \underline{\mu}_{n-1} + \underline{\mu}_{S_n}$$
$$\underline{\mu}_n^* = \underline{\mu}_n + \underline{\gamma}_n \qquad (3.3.2)$$

where $\underline{A}_n$ is given in Figure 4, $\underline{\mu}_{S_n}$ in Equation 2.4.5, and $\underline{\mu}_n^*$ in Equations 3.2.4 and 3.2.5.

The following assumptions are made concerning Equations 3.3.2:

(A) $\underline{\mu}_{S_n}$ is a random vector with $E\{\underline{\mu}_{S_n}\} = \overline{\underline{\mu}}_{S_n}$ and

$\Psi_{S_n} = E\{[\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n}][\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n}]^T\}$

(B) $\underline{\mu}_{C_n} = [\mu_{1C_n}, \mu_{2C_n}, \mu_{3C_n}]^T$ is a random vector with $E\{\underline{\mu}_{C_n}\} = \overline{\underline{\mu}}_{C_n}$ and $\Psi_{C_n} = E\{[\underline{\mu}_{C_n} - \overline{\underline{\mu}}_{C_n}][\underline{\mu}_{C_n} - \overline{\underline{\mu}}_{C_n}]^T\}$ (See Figure 4 for the relation of $\underline{\mu}_{C_n}$ to $\underline{A}_n$).

(C) The random vectors $\underline{\mu}_{C_n}$, $\underline{\mu}_{S_n}$, and $\underline{\mu}_i$, $i < n$, are independent. Thus $E\{\underline{\mu}_{C_n} \underline{\mu}_i^T\} = \overline{\underline{\mu}}_{C_n} E\{\underline{\mu}_i^T\}$, $E\{\underline{\mu}_{S_n} \underline{\mu}_i^T\} = \overline{\underline{\mu}}_{S_n} E\{\underline{\mu}_i^T\}$, $E\{\underline{\mu}_{C_n} \underline{\mu}_{S_n}^T\} = \overline{\underline{\mu}}_{C_n} \overline{\underline{\mu}}_{S_n}^T$, $i < n$

(D) The random vectors $\underline{\gamma}_i$, $i = 0,1,2,\ldots$, are orthogonal. Thus

$$E\{\underline{\gamma}_i \underline{\gamma}_j^T\} = E\{[\underline{\mu}_i^* - \underline{\mu}_i][\underline{\mu}_j^* - \underline{\mu}_j]^T\} = \underline{0} \quad i \neq j$$
$$= \Psi_i^* \quad i = j$$

and $E\{\underline{\gamma}_i\} = \underline{0}$, $i = 0,1,2,\ldots$

(E) The random vectors $\underline{\mu}_i$, $\underline{\mu}_{S_i}$, $\underline{\mu}_{C_i}$, and thus $\underline{A}_i$ are orthogonal to $\underline{\gamma}_j$. $E\{\underline{\mu}_i \underline{\gamma}_j^T\} = E\{\underline{\mu}_{S_i} \underline{\gamma}_j^T\} = E\{\underline{\mu}_{C_i} \underline{\gamma}_j^T\} = E\{\underline{A}_i \underline{\gamma}_j\} = \underline{0}$, $i,j = 0,1,\ldots$

Under the above assumptions $\underline{B}_n$, $\underline{C}_n$ and $\underline{D}_n$ are determined such that for

$$\hat{\underline{\mu}}_n = \underline{B}_n \hat{\underline{\mu}}_{n-1} + \underline{C}_n \underline{\mu}_n^* + \underline{D}_n \qquad (3.3.3)$$

$\hat{\underline{\mu}}_n$ is the linear estimate of $\underline{\mu}_n$ in terms of $\hat{\underline{\mu}}_0, \underline{\mu}_1^*, \ldots, \underline{\mu}_n^*$ and I, the identity matrix, which minimizes tr $E\{[\underline{\mu}_n - \hat{\underline{\mu}}_n][\underline{\mu}_n - \hat{\underline{\mu}}_n]^T\}$.

From assumptions (C) and (E)

$$E\{\underline{A}_n \underline{\mu}_i^*\} = \overline{\underline{A}}_n E\{\underline{\mu}_i\} \quad , \quad E\{\underline{\mu}_{S_n} \underline{\mu}_i^{*T}\} = \overline{\underline{\mu}}_{S_n} E\{\underline{\mu}_i^T\} \quad , \quad i < n \qquad (3.3.4)$$

where $E\{\underline{A}_n\} = \overline{\underline{A}}_n$

Similarly from (C), (D), and (E)

$$E\{\underline{\mu}_i \underline{\mu}_j^{*T}\} = E\{\underline{\mu}_i \underline{\mu}_j^T\} \quad , \quad i, j = 0, 1, 2, \ldots \qquad (3.3.5)$$

and

$$E\{\underline{\mu}_i^* \underline{\mu}_j^{*T}\} = E\{\underline{\mu}_i \underline{\mu}_j^T\} + \Psi_i^* \quad , \quad i = j$$
$$= E\{\underline{\mu}_i \underline{\mu}_j^T\} \quad , \quad i \neq j \qquad (3.3.6)$$

Since $\hat{\underline{\mu}}_n$ is to be the linear function of $\hat{\underline{\mu}}_0, \underline{\mu}_1^*, \ldots, \underline{\mu}_n^*$, and I which minimizes tr $E\{[\underline{\mu}_n - \hat{\underline{\mu}}_n][\underline{\mu}_n - \hat{\underline{\mu}}_n]^T\}$, orthogonality must hold. Thus

$$E\{[\underline{\mu}_n - \hat{\underline{\mu}}_n]\, \underline{\mu}_n^{*T}\} = \underline{0} \qquad (3.3.7)$$

$$E\{[\underline{\mu}_n - \hat{\underline{\mu}}_n]\, \underline{\mu}_i^{*T}\} = \underline{0} \quad , \quad i = 0, 1, \ldots, n-1 \qquad (3.3.8)$$

$$E\{[\underline{\mu}_n - \hat{\underline{\mu}}_n]\} = \underline{0} \qquad (3.3.9)$$

where in Equation 3.3.8 $\underline{\mu}_0^*$ is taken to be $\hat{\underline{\mu}}_0$. See Kalman (7) and Papoulis (10) for discussions and developments of orthogonality.

From (C) and Equations 3.3.2

$$E\{\underline{\mu}_n\} = \overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\} + \overline{\underline{\mu}}_{S_n} \qquad (3.3.10)$$

Taking expected values in Equation 3.3.3, since $\underline{\mu}_i^*$ is the UMV-RUE of $\underline{\mu}_i$,

$$E\{\underline{\mu}_i^*\} = E\{E\{\underline{\mu}_i^* | \underline{\mu}_{ia}\}\} = E\{\underline{\mu}_i\} \quad , \quad i = 0,1,\ldots,$$

and since $E\{[\underline{\mu}_{n-1} - \hat{\underline{\mu}}_{n-1}]\} = \underline{0}$, solving Equation 3.3.9 for $\underline{D}_n$ yields

$$\underline{D}_n = (I - \underline{C}_n)E\{\underline{\mu}_n\} - \underline{B}_n E\{\underline{\mu}_{n-1}\} \qquad (3.3.11)$$

Furthermore using Equation 3.3.10

$$\underline{D}_n = [(I - \underline{C}_n)\bar{\underline{A}}_n - \underline{B}_n]E\{\underline{\mu}_{n-1}\} + (I - \underline{C}_n)\bar{\underline{\mu}}_{S_n} \qquad (3.3.12)$$

From Equation 3.3.2

$$E\{[\underline{\mu}_n - \underline{A}_n\underline{\mu}_{n-1} - \underline{\mu}_{S_n}]\,\underline{\mu}_i^{*T}\} \equiv \underline{0} \ , \quad i = 0,1,\ldots,n-1$$

Then

$$E\{\underline{\mu}_n\underline{\mu}_i^{*T}\} = \bar{\underline{A}}\ E\{\underline{\mu}_{n-1}\underline{\mu}_i^{*T}\} + \bar{\underline{\mu}}_{S_n} E\{\underline{\mu}_i^{*T}\} \ , \quad i = 0,1,\ldots,n-1 \qquad (3.3.13)$$

Solving Equation 3.3.10 for $\bar{\underline{\mu}}_{S_n}$, using the results in Equation 3.3.14 and solving for $\bar{\underline{A}}_n$ yields

$$\bar{\underline{A}}_n = [E\{\underline{\mu}_n\underline{\mu}_i^T\} - E\{\underline{\mu}_n\}E\{\underline{\mu}_i^T\}][E\{\underline{\mu}_{n-1}\underline{\mu}_i^T\} - E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_i^T\}]^{-1},$$
$$i = 0,1,\ldots,n-1 \qquad (3.3.14)$$

Expressing $\underline{\mu}_n - \hat{\underline{\mu}}_n$ as

$$\underline{\mu}_n - \hat{\underline{\mu}}_n = \underline{\mu}_n - \underline{B}_n\underline{\mu}_{n-1} - \underline{C}_n\underline{\mu}_n + \underline{B}_n(\underline{\mu}_{n-1} - \underline{\mu}_{n-1}) - \underline{C}_n\underline{\gamma}_n - \underline{D}_n \ ,$$

since $E\{[\underline{\mu}_{n-1} - \hat{\underline{\mu}}_{n-1}]\,\underline{\mu}_i^{*T}\} = \underline{0}$, $i = 0,1,\ldots,n-1$, and from (D) and (E) $E\{\underline{\gamma}_n\underline{\mu}_i^{*T}\} = \underline{0}$, $i = 0,1,\ldots,n-1$,

Equation 3.3.8 becomes

$$E\,[\underline{\mu}_n - \hat{\underline{\mu}}_n]\,\underline{\mu}_i^{*T}\} = E\{[(I - \underline{C}_n)\underline{\mu}_n - \underline{B}_n\underline{\mu}_{n-1} - \underline{D}_n]\,\underline{\mu}_i^{*T}\} = \underline{0}\,, \quad i = 0,1,\ldots,n-1 \tag{3.3.15}$$

Using Equation 3.3.11 in Equation 3.3.15 and solving for $(I - \underline{C}_n)^{-1}\underline{B}_n$ yields

$$(I - \underline{C}_n)^{-1}\underline{B}_n = [E\{\underline{\mu}_n\underline{\mu}_i^T\} - E\{\underline{\mu}_n\}E\{\underline{\mu}_i^T\}][E\{\underline{\mu}_{n-1}\underline{\mu}_i^T\} - E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_i^T\}]^{-1}, \quad i = 0,1,\ldots,n-1 \tag{3.3.16}$$

Therefore from Equations 3.3.14 and 3.3.16

$$(I - \underline{C}_n)^{-1}\underline{B}_n = \overline{\underline{A}}_n \qquad \text{or} \qquad \underline{B}_n = (I - \underline{C}_n)\overline{\underline{A}}_n \tag{3.3.17}$$

Using Equations 3.3.3 and 3.3.11 in Equation 3.3.7

$$E\{[\underline{\mu}_n - \underline{B}_n\hat{\underline{\mu}}_{n-1} - \underline{C}_n\underline{\mu}_n^* - (I - \underline{C}_n)E\{\underline{\mu}_n\} + \underline{B}_nE\{\underline{\mu}_{n-1}\}]\,\underline{\mu}_n^{*T}\} = \underline{0}$$

which with Equations 3.3.5 and 3.3.6 becomes

$$E\{\underline{\mu}_n\underline{\mu}_n^T\} - E\{\underline{\mu}_n\}E\{\underline{\mu}_n^T\} = \underline{B}_n[E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_n^T\} - E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_n^T\}] + \underline{C}_n[E\{\underline{\mu}_n\underline{\mu}_n^T\} - E\{\underline{\mu}_n\}E\{\underline{\mu}_n^T\} + \Psi_n^*] \tag{3.3.18}$$

The error covariance matrix for $\hat{\underline{\mu}}_n$ is given by

$$\hat{\Psi}_n = E\{[\underline{\mu}_n - \hat{\underline{\mu}}_n][\underline{\mu}_n - \hat{\underline{\mu}}_n]^T\} = E\{[\underline{\mu}_n - \hat{\underline{\mu}}_n]\,\underline{\mu}_n^T\}$$

which, making use of Equations 3.3.3 and 3.3.11, becomes

$$\hat{\Psi}_n = E\{\underline{\mu}_n\underline{\mu}_n^T\} - E\{\underline{\mu}_n\}E\{\underline{\mu}_n^T\} - \underline{B}_n[E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_n^T\} - E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_n^T\}] - \underline{C}_n[E\{\underline{\mu}_n\underline{\mu}_n^T\} - E\{\underline{\mu}_n\}E\{\underline{\mu}_n^T\}] \tag{3.3.19}$$

Using Equation 3.3.18 in Equation 3.3.19 yields

$$\hat{\Psi}_n = \underline{C}_n \Psi_n^* \quad \text{or} \quad \underline{C}_n = \hat{\Psi}_n \Psi_n^{*-1} \tag{3.3.20}$$

Then with Equations 3.3.12, 3.3.17, and 3.3.20 $\hat{\underline{\mu}}_n$ of Equations 3.3.3 becomes

$$\hat{\underline{\mu}}_n = [I - \hat{\Psi}_n \Psi_n^{*-1}][\overline{\underline{A}}_n \hat{\underline{\mu}}_{n-1} + \overline{\underline{\mu}}_{S_n}] + \hat{\Psi}_n \Psi_n^{*-1} \underline{\mu}_n^* \tag{3.3.21}$$

Therefore to complete the development of $\hat{\underline{\mu}}_n$ an expression for $\hat{\Psi}_n$ in terms of $\hat{\Psi}_{n-1}$ and Ψ_n^* is necessary.

From Equations 3.3.2

$$E\{\hat{\underline{\mu}}_{n-1}[\underline{\mu}_n - \underline{A}_n \underline{\mu}_{n-1} - \underline{\mu}_{S_n}]^T\} \equiv \underline{0}$$

from which

$$E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_n^T\} = E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T + E\{\underline{\mu}_{n-1}\}\overline{\underline{\mu}}_{S_n}^T \tag{3.3.22}$$

The previous error covariance matrix is

$$\hat{\Psi}_{n-1} = E\{[\underline{\mu}_{n-1} - \hat{\underline{\mu}}_{n-1}][\underline{\mu}_{n-1} - \hat{\underline{\mu}}_{n-1}]^T\} = E\{[\underline{\mu}_{n-1} - \hat{\underline{\mu}}_{n-1}]\,\underline{\mu}_{n-1}^T\}$$

$$= E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\} - E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_{n-1}^T\}$$

so that Equation 3.3.22 can be written as

$$E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_n^T\} = [E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\} - \hat{\Psi}_{n-1}]\,\overline{\underline{A}}_n^T + E\{\underline{\mu}_{n-1}\}\overline{\underline{\mu}}_{S_n}^T$$

Solving Equation 3.3.10 for $\hat{\underline{\mu}}_{S_n}^T$, $E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_n^T\}$ becomes

$$E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_n^T\} = [E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\} - E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_{n-1}^T\} - \hat{\Psi}_{n-1}]\overline{\underline{A}}_n^T + E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_n^T\} \tag{3.3.23}$$

Using Equation 3.3.23 Equation 3.3.18 becomes

$$(I - \underline{C}_n)\Psi_{n,n} = \underline{B}_n[\Psi_{n-1,n-1} - \hat{\Psi}_{n-1}]\,\bar{\underline{A}}_n^T + \underline{C}_n\Psi_n^* \qquad (3.3.24)$$

where $$\Psi_{n,n} = E\{\underline{\mu}_n\underline{\mu}_n^T\} - E\{\underline{\mu}_n\}E\{\underline{\mu}_n^T\}$$

and $$\Psi_{n-1,n-1} = E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\} - E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_{n-1}^T\}$$

Substituting Equations 3.3.17 and 3.3.20 into Equation 3.3.24

$$(I - \hat{\Psi}_n\Psi_n^{*-1})\Psi_{n,n} = (I - \hat{\Psi}_n\Psi_n^{*-1})\bar{\underline{A}}_n[\Psi_{n-1,n-1} - \hat{\Psi}_{n-1}]\,\bar{\underline{A}}_n^T + \hat{\Psi}_n \qquad (3.3.25)$$

Solving Equation 3.3.25 for $\hat{\Psi}_n$ yields

$$\hat{\Psi}_n = \{\Psi_{n,n} + \bar{\underline{A}}_n[\hat{\Psi}_{n-1} - \Psi_{n-1,n-1}]\,\bar{\underline{A}}_n^T\}\{\Psi_{n,n} + \bar{\underline{A}}_n[\hat{\Psi}_{n-1} - \Psi_{n-1,n-1}]\,\bar{\underline{A}}_n^T + \Psi_n^*\}^{-1}\Psi_n^* \qquad (3.3.26)$$

Let

$$\underline{\mu}_n' = \bar{\underline{A}}_n\hat{\underline{\mu}}_{n-1} + \bar{\underline{\mu}}_{S_n} \qquad (3.3.27)$$

and

$$\Psi_n' = E\{[\underline{\mu}_n - \underline{\mu}_n'][\underline{\mu}_n - \underline{\mu}_n']^T\} \qquad (3.3.28)$$

then Equation 3.3.21 becomes

$$\hat{\underline{\mu}}_n = [I - \hat{\Psi}_n\Psi_n^{*-1}]\,\underline{\mu}_n' + \hat{\Psi}_n\Psi_n^{*-1}\underline{\mu}_n^* \qquad (3.3.29)$$

Now to show that

$$\Psi_{n,n} + \bar{\underline{A}}_n[\hat{\Psi}_{n-1} - \Psi_{n-1,n-1}]\,\bar{\underline{A}}_n^T = \Psi_n'$$

and that Ψ_n' can be expressed in terms of $\hat{\Psi}_{n-1}$.

First consider $\Psi_{n,n}$

$$\Psi_{n,n} = E\{[\underline{\mu}_n - E\{\underline{\mu}_n\}][\underline{\mu}_n - E\{\underline{\mu}_n\}]^T\}$$

$$= E\{[\underline{A}_n\underline{\mu}_{n-1} + \underline{\mu}_{S_n} - \overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\} - \overline{\underline{\mu}}_{S_n}][\underline{A}_n\underline{\mu}_{n-1} + \underline{\mu}_{S_n} - \overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\} - \overline{\underline{\mu}}_{S_n}]^T\}$$

$$= E\{\underline{A}_n\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\underline{A}_n^T\} + \overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T + \Psi_{S_n}$$

$$-\overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T + E\{\underline{A}_n\underline{\mu}_{n-1}(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})^T\}$$

$$-\overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\}E\{\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T - \overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\}E\{(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})^T\}$$

$$+E\{(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})\underline{\mu}_{n-1}^T\underline{A}_n^T\} - E\{\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n}\}E\{\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T$$

Adding and subtracting the term $\overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T$ to $\Psi_{n,n}$

$$\Psi_{n,n} = E\{\underline{A}_n\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\underline{A}_n^T\} - \overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T + \overline{\underline{A}}_n\Psi_{n-1,n-1}\overline{\underline{A}}_n^T + \Psi_{S_n}$$

$$+E\{\underline{A}_n\underline{\mu}_{n-1}(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})^T\} + E\{(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})\underline{\mu}_{n-1}^T\underline{A}_n^T\} \qquad (3.3.30)$$

Then using Equation 3.3.30 in $\Psi_{n,n} + \overline{\underline{A}}_n[\hat{\Psi}_{n-1} - \Psi_{n-1,n-1}]\overline{\underline{A}}_n^T$

$$\Psi_{n,n} + \overline{\underline{A}}_n[\hat{\Psi}_{n-1} - \Psi_{n-1,n-1}]\overline{\underline{A}}_n^T$$

$$= E\{\underline{A}_n\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\underline{A}_n^T\} - \overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T + \overline{\underline{A}}_n\hat{\Psi}_{n-1}\overline{\underline{A}}_n^T + \Psi_{S_n}$$

$$+ E\{\underline{A}_n\underline{\mu}_{n-1}(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})^T\} + E\{(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})\underline{\mu}_{n-1}^T\underline{A}_n^T\} \qquad (3.3.31)$$

Now consider Ψ_n' of Equation 3.3.28

$$\Psi_n' = E\{[\underline{A}_n\underline{\mu}_{n-1} + \underline{\mu}_{S_n} - \overline{\underline{A}}_n\hat{\underline{\mu}}_{n-1} - \overline{\underline{\mu}}_{S_n}]$$

$$\cdot[\underline{A}_n\underline{\mu}_{n-1} + \underline{\mu}_{S_n} - \overline{\underline{A}}_n\hat{\underline{\mu}}_{n-1} - \overline{\underline{\mu}}_{S_n}]^T\}$$

$$\Psi_n' = E\{\underline{A}_n\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\underline{A}_n^T\} + \overline{\underline{A}}_n E\{\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}_{n-1}^T\}\overline{\underline{A}}_n^T + \Psi_{S_n}$$

$$-\overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\hat{\underline{\mu}}_{n-1}^T\}\overline{\underline{A}}_n^T + E\{\underline{A}_n\underline{\mu}_{n-1}(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})^T\}$$

$$-\overline{\underline{A}}_n E\{\hat{\underline{\mu}}_{n-1}\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T - \overline{\underline{A}}_n E\{\hat{\underline{\mu}}_{n-1}\} E\{(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})^T\}$$

$$+E\{(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})\underline{\mu}_{n-1}^T\underline{A}_n^T\} - E\{\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n}\} E\{\hat{\underline{\mu}}_{n-1}^T\}\overline{\underline{A}}_n^T$$

Adding and subtracting the term $\overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T$ to Ψ_n'

$$\Psi_n' = E\{\underline{A}_n\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\underline{A}_n^T\} - \overline{\underline{A}}_n E\{\underline{\mu}_{n-1}\underline{\mu}_{n-1}^T\}\overline{\underline{A}}_n^T + \overline{\underline{A}}_n\hat{\Psi}_{n-1}\overline{\underline{A}}_n^T + \Psi_{S_n}$$

$$+ E\{\underline{A}_n\underline{\mu}_{n-1}(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})^T\} + E\{(\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})\underline{\mu}_{n-1}^T\underline{A}_n^T\} \qquad (3.3.32)$$

The right side of Equation 3.3.31 is the same as the right side of Equation 3.3.32 so that

$$\Psi_{n,n} + \overline{\underline{A}}_n[\hat{\Psi}_{n-1} - \Psi_{n-1,n-1}]\overline{\underline{A}}_n^T = \Psi_n' \qquad (3.3.33)$$

Therefore Equation 3.3.26 becomes

$$\hat{\Psi}_n = \Psi_n'[\Psi_n' + \Psi_n^*]^{-1}\Psi_n^* \qquad (3.3.34)$$

and $\hat{\underline{\mu}}_n$ of Equation 3.3.29 becomes

$$\hat{\underline{\mu}}_n = \Psi_n^*[\Psi_n' + \Psi_n^*]^{-1}\underline{\mu}_n' + \Psi_n'[\Psi_n' + \Psi_n^*]^{-1}\underline{\mu}_n^* \qquad (3.3.35)$$

Equation 3.3.32 is an expression of Ψ_n' in terms of $\hat{\Psi}_{n-1}$ but consider a slightly different development of Ψ_n'

$$\Psi_n' = E\{[\underline{A}_n(\underline{\mu}_{n-1} - \hat{\underline{\mu}}_{n-1}) + (\underline{A}_n - \overline{\underline{A}}_n)\hat{\underline{\mu}}_{n-1} + (\underline{\mu}_{S_n} - \overline{\underline{\mu}}_{S_n})]$$

$$\cdot [\underline{A}_n(\underline{\mu}_{n-1} - \hat{\underline{\mu}}_{n-1}) + (\underline{A}_n - \overline{\underline{A}}_n)\hat{\underline{\mu}}_{n-1} + (\underline{\mu}_{S_n} - \underline{\mu}_{S_n})]^T\}$$

$$\Psi_n' = E\{\underline{A}_n(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})^T\underline{A}_n^T\} + E\{(\underline{A}_n-\bar{\underline{A}}_n)\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n-\bar{\underline{A}}_n)^T\} + \Psi_{S_n}$$
$$+ E\{\underline{A}_n(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n-\bar{\underline{A}}_n)^T\} + E\{\underline{A}_n(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})^T\}$$
$$+ E\{(\underline{A}_n-\bar{\underline{A}}_n)\hat{\underline{\mu}}_{n-1}(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})^T\underline{A}_n^T\} + E\{(\underline{A}_n-\bar{\underline{A}}_n)\hat{\underline{\mu}}_{n-1}(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})^T\}$$
$$+ E\{(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})^T\underline{A}_n^T\} + E\{(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n-\bar{\underline{A}}_n)^T\} \quad (3.3.36)$$

Since $\underline{\mu}_{C_n}$, $\underline{\mu}_{S_n}$, and $\underline{\mu}_i$ are independent the expectations with respect to $\underline{\mu}_{C_n}$, $\underline{\mu}_{S_n}$, and $\underline{\mu}_i$ can be taken separately and then using the orthogonality relations for $\hat{\underline{\mu}}_{n-1}$

$$E\{\underline{A}_n(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n-\bar{\underline{A}}_n)^T\} = E\{\underline{A}_n E\{(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})\hat{\underline{\mu}}_{n-1}^T\}(\underline{A}_n-\bar{\underline{A}}_n)^T\} = \underline{0}$$

and

$$E\{\underline{A}_n(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})^T\} = E\{\underline{A}_n E\{\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1}\}(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})^T\} = \underline{0}$$

Similarly

$$E\{(\underline{A}_n-\bar{\underline{A}}_n)\hat{\underline{\mu}}_{n-1}(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})^T\underline{A}_n^T\} = \underline{0}$$

and

$$E\{(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})(\underline{\mu}_{n-1}-\hat{\underline{\mu}}_{n-1})^T\underline{A}_n^T\} = \underline{0}$$

Then Ψ_n' becomes

$$\Psi_n' = E\{\underline{A}_n\hat{\Psi}_{n-1}\underline{A}_n^T\} + E\{(\underline{A}_n-\bar{\underline{A}}_n)\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n-\bar{\underline{A}}_n)^T\} + \Psi_{S_n}$$
$$+ E\{(\underline{A}_n-\bar{\underline{A}}_n)\hat{\underline{\mu}}_{n-1}(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})^T\} + E\{(\underline{\mu}_{S_n}-\bar{\underline{\mu}}_{S_n})\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n-\bar{\underline{A}}_n)^T\}$$
$$(3.3.37)$$

The results of this development are now summarized. If the estimate $\hat{\underline{\mu}}_{n-1}$ and its error covariance matrix $\hat{\Psi}_{n-1}$ are available, $\underline{\mu}_n'$ and its error covariance matrix Ψ_n' are determined from Equations 3.3.27 and 3.3.37, respectively, which are

$$\underline{\mu}_n' = \underline{\bar{A}}_n \hat{\underline{\mu}}_{n-1} + \underline{\bar{\mu}}_{S_n} \tag{3.3.38}$$

$$\begin{aligned} \Psi_n' = & E\{\underline{A}_n \hat{\Psi}_{n-1} \underline{A}_n^T\} + E\{(\underline{A}_n - \underline{\bar{A}}_n)\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n - \underline{\bar{A}}_n)^T\} + \Psi_{S_n} \\ & + E\{(\underline{A}_n - \underline{\bar{A}}_n)\hat{\underline{\mu}}_{n-1}(\underline{\mu}_{S_n} - \underline{\bar{\mu}}_{S_n})^T\} + E\{(\underline{\mu}_{S_n} - \underline{\bar{\mu}}_{S_n})\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n - \underline{\bar{A}}_n)^T\} \end{aligned} \tag{3.3.39}$$

Then $\hat{\underline{\mu}}_n$ and its covariance matrix $\hat{\Psi}_n$ are determined from Equations 3.3.35 and 3.3.34, respectively, which are

$$\hat{\underline{\mu}}_n = \Psi_n^*[\Psi_n' + \Psi_n^*]^{-1} \underline{\mu}_n' + \Psi_n'[\Psi_n' + \Psi_n^*]^{-1} \underline{\mu}_n^* \tag{3.3.40}$$

$$\hat{\Psi}_n = \Psi_n'[\Psi_n' + \Psi_n^*]^{-1} \Psi_n^* \tag{3.3.41}$$

Some comments on the difficulties which arise when Equations 3.3.38 through 3.3.41 are implemented are offered in the next section.

There are several interesting cases of this development which are worth enumerating here. They are:

<u>Case I.</u> $\underline{A}_n$ unknown, $\underline{\mu}_{S_n}$ unknown; $\underline{A}_n$ and $\underline{\mu}_{S_n}$ dependent.

This is the most general case; the one for which Equations 3.3.38 through 3.3.41 were developed.

<u>Case II.</u> $\underline{A}_n$ unknown, $\underline{\mu}_{S_n}$ unknown; $\underline{A}_n$ and $\underline{\mu}_{S_n}$ independent.

This case is not possible with the augmented moment model since $\underline{A}_n$ is a function of $\underline{\mu}_{S_n}$ but it does hold interest for those situations in which $\underline{A}_n$ and $\underline{\mu}_{S_n}$ are not related. In this case Equations 3.3.38 through 3.3.41 become

$$\underline{\mu}_n' = \underline{\bar{A}}_n \hat{\underline{\mu}}_{n-1} + \underline{\bar{\mu}}_{S_n} \tag{3.3.42a}$$

$$\Psi_n' = E\{\underline{A}_n \hat{\Psi}_{n-1} \underline{A}_n\} + E\{(\underline{A}_n - \underline{\bar{A}}_n)\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n - \underline{\bar{A}}_n)^T\} + \Psi_{S_n} \tag{3.3.42b}$$

$$\hat{\underline{\mu}}_n = \Psi_n^*[\Psi_n' + \Psi_n^*]^{-1} \underline{\mu}_n' + \Psi_n'[\Psi_n' + \Psi_n^*]^{-1} \underline{\mu}_n^* \tag{3.3.42c}$$

$$\hat{\Psi}_n = \Psi_n'[\Psi_n' + \Psi_n^*]^{-1} \Psi_n^* \qquad (3.3.42d)$$

Case III. $\underline{A}_n$ known, $\underline{\mu}_{S_n}$ unknown.

This case also is not possible with the augmented moment model but it has special importance which makes it worthy of presentation. In this case the recursive moment estimation equations are

$$\underline{\mu}_n' = \underline{A}_n\hat{\underline{\mu}}_{n-1} + \bar{\underline{\mu}}_{S_n} \qquad (3.3.43a)$$

$$\Psi_n' = \underline{A}_n\Psi_{n-1}\underline{A}_n^T + \Psi_{S_n} \qquad (3.3.43b)$$

$$\hat{\underline{\mu}}_n = \Psi_n^*[\Psi_n' + \Psi_n^*]^{-1} \underline{\mu}_n' + \Psi_n'[\Psi_n' + \Psi_n^*]^{-1} \underline{\mu}_n^* \qquad (3.3.43c)$$

$$\hat{\Psi}_n = \Psi_n'[\Psi_n' + \Psi_n^*]^{-1} \Psi_n^* \qquad (3.3.43d)$$

These results (with $\bar{\underline{\mu}}_{S_n} = \underline{0}$) are the same as those obtained by Kalman (7) and are the vector form of those obtained by Papoulis (10).

Case IV. $\underline{A}_n$ unknown, $\underline{\mu}_{S_n}$ known.

This case would apply to the augmented moment model if $\underline{\mu}_{C_n}$ were unknown while $\underline{\mu}_{S_n}$ was known. The moment estimation equations are

$$\underline{\mu}_n' = \bar{\underline{A}}_n\hat{\underline{\mu}}_{n-1} + \underline{\mu}_{S_n} \qquad (3.3.44a)$$

$$\Psi_n' = E\{\underline{A}_n\hat{\Psi}_{n-1}\underline{A}_n^T\} + E\{(\underline{A}_n-\bar{\underline{A}}_n)\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}_{n-1}^T(\underline{A}_n-\bar{\underline{A}}_n)^T\} \qquad (3.3.44b)$$

$$\hat{\underline{\mu}}_n = \Psi_n^*[\Psi_n' + \Psi_n^*]^{-1} \underline{\mu}_n' + \Psi_n'[\Psi_n' + \Psi_n^*]^{-1} \underline{\mu}_n^* \qquad (3.3.44c)$$

$$\hat{\Psi}_n = \Psi_n'[\Psi_n' + \Psi_n^*]^{-1} \Psi_n^* \qquad (3.3.44d)$$

Case V. $\underline{A}_n$ known, $\underline{\mu}_{S_n}$ known.

This is the simplest case and will be used extensively in the next chapter. It occurs when $\underline{\mu}_{C_n}$ and $\underline{\mu}_{S_n}$ are known. The recursive moment estimation equations are

$$\underline{\mu}_n' = \underline{A}_n\hat{\underline{\mu}}_{n-1} + \underline{\mu}_{S_n} \qquad (3.3.45a)$$

$$\Psi_n' = \underline{A}_n\hat{\Psi}_{n-1}\underline{A}_n^T \qquad (3.3.45b)$$

$$\hat{\underline{\mu}}_n = \Psi_n^*[\Psi_n' + \Psi_n^*]^{-1}\,\underline{\mu}_n' + \Psi_n'[\Psi_n' + \Psi_n^*]^{-1}\,\underline{\mu}_n^* \qquad (3.3.45c)$$

$$\hat{\Psi}_n = \Psi_n'[\Psi_n' + \Psi_n^*]^{-1}\,\Psi_n^* \qquad (3.3.45d)$$

3.4 Inherent and Practical Difficulties. One of the inherent difficulties in recursive moment estimation was introduced previously. It is that Ψ_n^* is unknown. Section 3.2 presents a detailed discussion of Ψ_n^* and proposes the use of $\hat{\Psi}_n^*$ in its place. Section 3.2 should be referred to for the development of $\hat{\Psi}_n^*$. In that development $\hat{\Psi}_n^*$ is the UMV-RUE of $\Psi_n^{*'} = E\{[\underline{\mu}_n^*-\underline{\mu}_n][\underline{\mu}_n^*-\underline{\mu}_n]^T|\mu_{na}\}$, which is the conditional covariance matrix of $\underline{\mu}_n^*$ given the moments of X_n. Although the theoretical development of Section 3.3 requires that Ψ_n^* be used in the weighting of $\underline{\mu}_n^*$ and $\underline{\mu}_n'$ (See Equation 3.3.41), in practice when $\underline{\mu}(n)$ is the value of $\underline{\mu}_n$ that occurs $\Psi_n^{*'}$ would be a reasonable weighting parameter to use, but still $\Psi_n^{*'}$ is unknown and in its place $\hat{\Psi}_n^*$ is used.

In some of the more general cases of recursive moment estimation enumerated in Section 3.3 there are yet other unknowns. Due to the independence of $\underline{\mu}_{C_n}$, $\underline{\mu}_{S_n}$ and $\hat{\underline{\mu}}_{n-1}$ Equation 3.3.40 can be written

$$\Psi_n' = E\{\underline{A}_n\hat{\Psi}_{n-1}\underline{A}_n^T\} + E\{(\underline{A}_n-\overline{\underline{A}}_n)E\{\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}_{n-1}^T\}(\underline{A}_n-\overline{\underline{A}}_n)^T\} + \Psi_{S_n}$$
$$+ E\{(\underline{A}_n-\overline{\underline{A}}_n)E\{\hat{\underline{\mu}}_{n-1}\}(\underline{\mu}_{S_n}-\overline{\underline{\mu}}_{S_n})^T\} + E\{(\underline{\mu}_{S_n}-\overline{\underline{\mu}}_{S_n})E\{\hat{\underline{\mu}}_{n-1}^T\}(\underline{A}_n-\overline{\underline{A}}_n)^T\} \qquad (3.4.1)$$

Two obvious unknowns required by this equation are the mean value of $\hat{\underline{\mu}}_{n-1}$, $E\{\hat{\underline{\mu}}_{n-1}\}$ (which is also $E\{\underline{\mu}_{n-1}\}$), and the matrix $E\{\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}^T_{n-1}\}$.

An alternative to the use of $E\{\hat{\underline{\mu}}_{n-1}\}$ and $E\{\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}^T_{n-1}\}$ is offered. It is suggested that $\hat{\underline{\mu}}(n-1)$ the value of $\hat{\underline{\mu}}_{n-1}$ which has been determined be used for $E\{\hat{\underline{\mu}}_{n-1}\}$, since $E\{\hat{\underline{\mu}}_{n-1}\} = E\{\underline{\mu}_{n-1}\}$ is unknown and $\hat{\underline{\mu}}(n-1)$ is the estimate of $\underline{\mu}_{n-1}$. To determine $E\{\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}^T_{n-1}\}$ the covariance matrix

$$E\{[\hat{\underline{\mu}}_{n-1}-E\{\hat{\underline{\mu}}_{n-1}\}][\hat{\underline{\mu}}_{n-1}-E\{\hat{\underline{\mu}}_{n-1}\}]^T\} = E\{\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}^T_{n-1}\} - E\{\hat{\underline{\mu}}_{n-1}\}E\{\hat{\underline{\mu}}^T_{n-1}\}$$

is necessary, but it is unknown. Since $\hat{\underline{\mu}}(n-1)$ is used for $E\{\hat{\underline{\mu}}_{n-1}\}$ and the nearest thing to an estimate of this covariance matrix is $\hat{\Psi}_{n-1}$, it is suggested that $E\{\hat{\underline{\mu}}_{n-1}\hat{\underline{\mu}}^T_{n-1}\}$ be approximated by

$$\hat{\Psi}_{n-1} + \hat{\underline{\mu}}(n-1)\hat{\underline{\mu}}^T(n-1)$$

When $\underline{\mu}_{C_n}$ and/or $\underline{\mu}_{S_n}$ are unknown, Cases I, II, III, and IV of Section 3.3, $\bar{\underline{\mu}}_{S_n}$, $\bar{\underline{\mu}}_{C_n}$, and their covariance matrices must be known. In Section 2.5 a discussion of estimation of $\underline{\mu}_{S_n}$ and $\underline{\mu}_{C_n}$ is presented. The estimates presented there can be used for $\bar{\underline{\mu}}_{S_n}$ and $\bar{\underline{\mu}}_{C_n}$. Estimation of Ψ_{C_n} could follow the same procedure for determining $\hat{\Psi}^*_n$ presented in Section 3.2. But estimation of Ψ_{S_n} must undoubtedly be based on engineering judgement just as is the estimation of $\underline{\mu}_{S_n}$ presented in Section 2.5.

In addition to these inherent difficulties which essentially arise from lack of information concerning the statistical properties of the augmented moment model, there are some practical problems which stem from the use of Ψ^*_n in the combination of $\underline{\mu}'_n$ with $\underline{\mu}^*_n$.

Particularly, two of these problems are that; 1) $\hat{\Psi}^*_n$ may not always be a positive definite matrix as an error covariance matrix should be and, 2) $\Psi_n' + \hat{\Psi}_n^*$ is ill-conditioned for the matrix inversion which is required in both $\hat{\underline{\mu}}_n$ and $\hat{\Psi}_n$. To illustrate these two conditions consider the following example.

From a normal distribution with mean 10 and variance 1, N(10,1), a sample of size 50 was drawn. $\hat{\Psi}_n^*$ was then constructed according to the procedure outlined at the conclusion of Section 3.2. The resulting $\hat{\Psi}_n^*$ is given here with the lower part of the matrix omitted since $\hat{\Psi}_n^*$ is symmetric.

$$\hat{\Psi}_n^* = \begin{bmatrix} .01351 & -.00334 & .001040 & .272 & 4.09 & -.0248 \\ & .01891 & -.01525 & .0675 & -1.023 & .1882 \\ & & .0354 & .0213 & .326 & -.1532 \\ & & & 5.46 & 82.3 & -.502 \\ & & & & 1240. & -7.63 \\ & & & & & 1.880 \end{bmatrix} \qquad (3.4.2)$$

Consider the (2 x 2) principal minor

$$\begin{vmatrix} .01351 & 4.09 \\ 4.09 & 1240. \end{vmatrix} = 16.75 - 16.75 = 0$$

Here only slide rule accuracy has been used. The actual $\hat{\Psi}_n^*$ was generated in a computer simulation which will be discussed in the next chapter. In the computer simulation the same principal minor as that given above is

$$\begin{vmatrix} .013514378 & 4.0937072 \\ 4.0937072 & 1239.7335 \end{vmatrix} = 16.7542544 - 16.7584386$$

$$= -.0041842 < 0 \qquad (3.4.3)$$

Since this principal minor is negative $\hat{\Psi}_n^*$ is not positive definite.

Since $\hat{\Psi}_n^*$ is not necessarily a positive definite matrix it can not be assumed that $|\hat{\Psi}_n^*| \neq 0$ or that $|\hat{\Psi}_n^* + \Psi_n'| \neq 0$. Thus $\hat{\Psi}_n^* + \Psi_n'$ may not have in inverse. Even if $\hat{\Psi}_n^* + \Psi_n'$ does have an inverse there can be difficulty in determining it. When a matrix contains such large numbers as 1240 and such small numbers as .001040 as $\hat{\Psi}_n^*$ does, the matrix is not easily and accurately inverted. If Ψ_n' is a comparable matrix to $\hat{\Psi}_n^*$ this condition will remain and $\Psi_n' + \hat{\Psi}_n^*$ will be difficult to invert. Of course there are very sophisticated computer routines which will do a fairly accurate inversion on such an ill-conditioned matrix, but they are generally very time consuming.

3.5 Pseudo-Minimum Variance Recursive Moment Estimation: An Alternative. To eliminate some of the difficulties encountered in the previous section an alternate approach is proposed which modifies the method of determining $\hat{\underline{\mu}}_n$ from $\underline{\mu}_n'$ and $\underline{\mu}_n^*$, Equation 3.3.41, and $\hat{\Psi}_n$ from Ψ_n' and $\hat{\Psi}_n^*$, Equation 3.3.42, but which does not affect the various ways of determining $\underline{\mu}'_n$ from $\hat{\underline{\mu}}_{n-1}$ and Ψ_n' from $\hat{\Psi}_{n-1}$ enumerated in Section 3.3. Essentially this alternative combines $\underline{\mu}_n'$ and $\underline{\mu}_n^*$ element by element so as to minimize the mean squared error of the resulting elements of $\hat{\underline{\mu}}_n$.

To facilitate the presentation of this alternative $\underline{\mu}_n$, $\underline{\mu}_n'$, $\underline{\mu}_n^*$, and $\hat{\underline{\mu}}_n$ are redefined with a slight modification in notation. Let

$\underline{\mu}_n = \{\mu_{i,n}\}$, $i = 1,\ldots,6$, so that equating this to $\underline{\mu}_n$ of Equation 2.4.4,

$$\{\mu_{i,n}\} = \begin{bmatrix} \mu_{1,n} \\ \mu_{2,n} \\ \mu_{3,n} \\ \mu_{4,n} \\ \mu_{5,n} \\ \mu_{6,n} \end{bmatrix} = \begin{bmatrix} \mu_{1,n} \\ \mu_{2,n} \\ \mu_{3,n} \\ \mu_{1,n}^2 \\ \mu_{1,n}^3 \\ \mu_{2,n}\mu_{1,n} \end{bmatrix} = \underline{\mu}_n \qquad (3.5.1)$$

Similarly

$$\underline{\mu}'_n = \{\mu'_{i,n}\}, \quad \underline{\mu}_n = \{\mu^*_{i,n}\}, \quad \hat{\underline{\mu}}_n = \{\hat{\mu}_{i,n}\}, \quad i = 1,\ldots,6 \qquad (3.5.2)$$

In accordance with this notation Ψ'_n, $\hat{\Psi}^*_n$, and $\hat{\Psi}_n$ are given by

$$\Psi'_n = \{\sigma'_{ij,n}\}, \quad \hat{\Psi}^*_n = \{\hat{\sigma}^*_{ij,n}\}, \quad \hat{\Psi}_n = \{\hat{\sigma}_{ij,n}\}, \quad i,j = 1,\ldots,6 \qquad (3.5.3)$$

where for example $\sigma'_{ij,n} = E\{[\mu_{i,n}-\mu'_{i,n}][\mu_{j,n}-\mu'_{j,n}]\}$ is the error covariance between the i<u>th</u> and j<u>th</u> elements of $\underline{\mu}'_n$ of Equations 3.5.2.

If the augmented moment model

$$\underline{\mu}_n = \underline{A}_n\underline{\mu}_{n-1} + \underline{\mu}_{S_n}$$

could be reduced to

$$\mu_{i,n} = a_{ii,n}\mu_{i,n-1} + \mu_{iS_n}, \quad i = 1,\ldots,6 \qquad (3.5.4)$$

then a one dimensional development analogous to the development of Section 3.3 produces the following recursive moment estimation equations

$$\mu'_{i,n} = \bar{a}_{ii,n}\hat{\mu}_{i,n-1} + \bar{\mu}_{iS_n} \tag{3.5.5}$$

$$\sigma'_{ii,n} = E\{a^2_{ii,n}\}\hat{\sigma}_{ii,n} + E\{(a_{ii,n}-\bar{a}_{ii,n})^2\}E\{\hat{\mu}^2_{i,n-1}\} + E\{(\mu_{iS_n}-\bar{\mu}_{iS_n})^2\}$$
$$+ 2E\{(a_{ii,n}-\bar{a}_{ii,n})(\mu_{iS_n}-\bar{\mu}_{iS_n})\}E\{\hat{\mu}_{i,n-1}\} \tag{3.5.6}$$

$$\hat{\mu}_{i,n} = \frac{\sigma^*_{ii,n}}{\sigma'_{ii,n}+\sigma^*_{ii,n}}\mu'_{i,n} + \frac{\sigma'_{ii,n}}{\sigma'_{ii,n}+\sigma^*_{ii,n}}\mu^*_{i,n} \tag{3.5.7}$$

$$\hat{\sigma}_{ii,n} = \frac{\sigma'_{ii,n}\sigma^*_{ii,n}}{\sigma'_{ii,n}+\sigma^*_{ii,n}} \tag{3.5.8}$$

In addition the error covariance $\hat{\sigma}_{ij,n} = E\{[\mu_{i,n}-\hat{\mu}_{i,n}][\mu_{j,n}-\hat{\mu}_{j,n}]\}$ is

$$\hat{\sigma}_{ij,n} = \frac{1}{(\sigma'_{ii,n}+\sigma^*_{ii,n})(\sigma'_{jj,n}+\sigma^*_{jj,n})}[\sigma^*_{ii,n}\sigma^*_{jj,n}\sigma'_{ij,n}+\sigma'_{ii,n}\sigma'_{jj,n}\sigma^*_{ij,n}] \tag{3.5.9}$$

Equations 3.3.41 and 3.3.42 would be the same as Equations 3.5.7 and 3.5.8 if the covariances between the elements of $\underline{\mu}'_n$ and the covariances between the elements of $\underline{\mu}^*_n$ were zero, i.e., if the off diagonal terms of Ψ'_n and Ψ^*_n were zero.

In the case where $a_{ii,n}$ is known Equations 3.5.5 and 3.5.6 become

$$\mu'_{i,n} = a_{ii,n}\mu_{i,n-1} + \bar{\mu}_{iS_n} \tag{3.5.10}$$

$$\sigma'_{ii,n} = a^2_{ii,n}\hat{\sigma}_{ii,n} + E\{(\mu_{iS_n}-\bar{\mu}_{iS_n})^2\} \tag{3.5.11}$$

If in addition $\bar{\mu}_{iS_n} = \underline{0}$ this one dimensional development is carried out by Papoulis (10), although he does not reduce his results to the same form as Equations 3.5.10, 3.5.11, 3.5.7. and 3.5.8.

This simplification of Equations 3.5.5 and 3.5.6 is analogous to Case III of Section 3.3. There likewise are one dimensional analogies for each of the other four cases presented.

Since, as has been discussed previously, $\Psi_n^* = \{\sigma_{ij,n}^*\}$ is unknown, Equations 3.5.7, 3.5.8, and 3.5.9 must use estimates of $\sigma_{ii,n}^*$. These estimates are taken to be the corresponding elements of $\hat{\Psi}_n^* = \{\hat{\sigma}_{ij,n}^*\}$, i.e., $\hat{\sigma}_{ii,n}^*$. Then Equations 3.5.7, 3.5.8 and 3.5.9 become

$$\hat{\mu}_{i,n} = \frac{\hat{\sigma}_{ii,n}^*}{\sigma_{ii,n}'+\sigma_{ii,n}^*}\mu_{i,n}' + \frac{\sigma_{ii,n}'}{\sigma_{ii,n}'+\hat{\sigma}_{ii,n}^*}\mu_{i,n}^* \qquad (3.5.12)$$

$$\hat{\sigma}_{ii,n} = \frac{\sigma_{ii,n}'\hat{\sigma}_{ii,n}^*}{\sigma_{ii,n}'+\hat{\sigma}_{ii,n}^*} \qquad (3.5.13)$$

$$\hat{\sigma}_{ij,n} = \frac{1}{(\sigma_{ii,n}'+\hat{\sigma}_{ii,n}^*)(\sigma_{jj,n}'+\hat{\sigma}_{jj,n}^*)}\left[\hat{\sigma}_{ii,n}^*\hat{\sigma}_{jj,n}^*\sigma_{ij,n}'+\sigma_{ii,n}'\sigma_{jj,n}'\hat{\sigma}_{ij,n}^*\right] \qquad (3.5.14)$$

Therefore, to alleviate some of the difficulties encountered in the recursive moment estimation scheme of Section 3.3., namely, that $\hat{\Psi}_n^*$ may not be positive definite and that $\Psi_n' + \hat{\Psi}_n^*$ may not be invertable, either theoretically or practically, it is proposed that Equations 3.5.12 through 3.5.14 be used in place of Equations 3.3.41 and 3.3.42.

Equations 3.5.12 and 3.5.13 are very similar in form to the results of combining two unbiased estimators so as to minimize the variance of the resulting unbiased estimator. See Fraser (6), Problem 6, p. 244. This similarity leads to the phrase "pseudo-minimum variance recursive moment estimation" to identify this alternative recursive moment estimation procedure.

Thus far no comments have been offered as to how this recursive scheme begins. Assuming that $\underline{\mu}_0'$ and Ψ_0' are the a priori estimate of

$\underline{\mu}_0$ and its error covariance matrix, Equations 3.5.12 through 3.5.14 are used to combine $\underline{\mu}_0'$ and Ψ_0' with $\underline{\mu}_0^*$ and $\hat{\Psi}_0^*$ to determine $\hat{\underline{\mu}}_0$ and $\hat{\Psi}_0^*$. From Equation 3.5.12 it is observed that, if the confidence in the a priori estimate $\mu_{i,0}'$ is great, as would be reflected by a small mean squared error, $\sigma_{ii,0}'$, $\hat{\mu}_{i,0}$ would be influenced mostly by $\mu_{i,0}'$, while if there is little confidence, a large $\sigma_{ii,0}'$, $\hat{\mu}_{i,0}$ would be influenced mostly by $\mu_{i,0}^*$.

In order to summarize the results of this section and this chapter the following pseudo-minimum variance recursive moment estimation algorithm is presented. Case V, $\underline{A}_n$ known, $\underline{\mu}_{S_n}$ known, is used. However the same procedure holds by changing only the equations involving $\underline{\mu}_n'$ and $\hat{\underline{\mu}}_{n-1}$ and Ψ_n' and $\hat{\Psi}_{n-1}$.

The Pseudo-Minimum Variance Recursive Moment Estimation Algorithm:

(1) Determine the prediction estimate, $\underline{\mu}_n'$, from $\hat{\underline{\mu}}_{n-1}$:

$$\mu_n' = \underline{A}_n\hat{\underline{\mu}}_{n-1} + \underline{\mu}_{S_n} \tag{3.5.15}$$

and the error covariance matrix of $\underline{\mu}_n'$, Ψ_n', from $\hat{\Psi}_{n-1}$:

$$\Psi_n' = \underline{A}_n\hat{\Psi}_{n-1}\underline{A}_n^T \tag{3.5.16}$$

(2) From the observations of X_n, $\underline{\mu}_n^*$, the data estimate or observation of $\underline{\mu}_n$, is computed according to Equations 3.2.4 and 3.2.5 and the estimated error covariance matrix of $\underline{\mu}_n^*$, $\hat{\Psi}_n^*$, is determined using the UMV-RUE's of moments and products of moments, Equations C.4.5, in Equations C.4.4.

(3) The pseudo-minimum variance estimate, $\hat{\underline{\mu}}_n$, is determined from

$$\hat{\mu}_{i,n} = \frac{1}{\sigma'_{ii,n}+\hat{\sigma}^*_{ii,n}}\left[\hat{\sigma}^*_{ii,n}\mu'_{i,n} + \sigma'_{ii,n}\mu^*_{i,n}\right], \quad i = 1,\ldots,6 \qquad (3.5.17)$$

and its error covariance matrix is calculated from

$$\hat{\sigma}_{ii,n} = \frac{\sigma'_{ii,n}\hat{\sigma}^*_{ii,n}}{\sigma'_{ii,n}+\hat{\sigma}^*_{ii,n}}, \quad i = 1,\ldots,6 \qquad (3.5.18)$$

and

$$\hat{\sigma}_{ij,n} = \frac{1}{(\sigma'_{ii,n}+\hat{\sigma}^*_{ii,n})(\sigma'_{jj,n}+\hat{\sigma}^*_{jj,n})}\left[\hat{\sigma}^*_{ii,n}\hat{\sigma}^*_{jj,n}\sigma'_{ij,n}+\sigma'_{ii,n}\sigma'_{jj,n}\hat{\sigma}^*_{ij,n}\right], \quad i = 1,\ldots,6 \qquad (3.5.19)$$

CHAPTER IV

SIMULATION AND DISCUSSION OF RESULTS

4.1 Introduction. This chapter is concerned with a discussion of a computer aided simulation of the pseudo-minimum variance recursive moment estimation algorithm and some results of several simulations. Only the important points of the simulation are presented with major emphasis placed on the results.

4.2 Simulating Program. In order to demonstrate the pseudo-minimum variance recursive moment estimation algorithm and to investigate its moment learning ability a Fortran IV computer program was written and implemented on an IBM 7040 computer. For comparison purposes the Bayesian recursive moment estimation algorithm, developed in Appendix D, was included in the program.

The program simulated the system model, $X_n = C_n X_{n-1} + S_n$, by recursively constructing a number of its sample functions. The sample functions were then recursively sampled without replacement, i.e., no one sample function was used more than once at one sampling time, and from these samples $\underline{\mu}_n^*$ and $\hat{\Psi}_n^*$ were determined. The number of sample functions and the number of samples taken at each sampling were specified initially. Using the initial assumptions for $\underline{\mu}_0$ and Ψ_0, i.e., $\underline{\mu}_0'$ and Ψ_0', and $\underline{\mu}_n^*$ and $\hat{\Psi}_n^*$, determined at each sampling, the pseudo-minimum variance and Bayesian recursive moment estimation algorithms, Equations 3.5.15 through 3.5.19 and Equations D.6.1 through D.6.3, respectively, were programed to determine the estimates $\underline{\mu}_n'$ and $\hat{\underline{\mu}}_n$, and the corresponding error covariance matrices, Ψ_n' and $\hat{\Psi}_n$. To support these estimates and to aid in evaluating their accuracies the augmented moment model, $\underline{\mu}_n = \underline{A}_n \underline{\mu}_{n-1} + \underline{\mu}_{S_n}$, was also recursively computed. Then for each of the estimates, $\underline{\mu}_n'$, $\underline{\mu}_n^*$ and $\hat{\underline{\mu}}_n$, the ratios of each of their elements to the corresponding elements of $\underline{\mu}_n$ was determined.

In order to construct the sample functions the initial random variable X_0 was assumed to be normally distributed with mean $\mu_{1,n}$ and variance $\mu_{2,n}$, i.e., $N(\mu_{1,n}, \mu_{2,n})$. By sampling this distribution (by means of a random number generator) the initial values of the sample functions were determined. Then for each $n \geq 1$, C_n and S_n were assumed to be $N(\mu_{1C_n}, \mu_{2C_n})$ and $N(\mu_{1S_n}, \mu_{2S_n})$, respectively. By recursively sampling these two distributions and using the resulting values with the initial values of the sample functions the sample functions were constructed. The program was written so that $(\mu_{1C_i}, \mu_{2C_i}) = (\mu_{1C_j}, \mu_{2C_j})$ and $(\mu_{1S_i}, \mu_{2S_i}) = (\mu_{1S_j}, \mu_{2S_j})$, $i, j = 1, 2, \ldots$.

It should be noted that, although X_0, C_n and S_n were normally distributed random variables, for $n \geq 1$, X_n was not a normally distributed random variable. Also by setting the values of (μ_{1C_n}, μ_{2C_n}) and (μ_{1S_n}, μ_{2S_n}) for all n the simulation was restricted to Case V of Section 3.3 where $\underline{A}_n$ and $\underline{\mu}_{S_n}$ are known. In particular $\underline{A}_i = \underline{A}_j$ and $\underline{\mu}_{S_i} = \underline{\mu}_{S_j}$, $i, j = 1, 2, \ldots$.

<u>4.3 Simulation Results and Discussion.</u> Figures 6 through 14

present some typical results from simulations performed using the computer program described in the previous section. In each case 1,000 sample functions were generated. From these 1,000 sample functions 50 samples were taken at each sampling time, $n = 0,1,2,\ldots$. In each of these simulations $\mu_{1,0}$, (μ_{1C_n}, μ_{2C_n}), and (μ_{1S_n}, μ_{2S_n}) remained constant with $\mu_{1,0} = 10$, $(\mu_{1C_n}, \mu_{2C_n}) = (1.0, 0.01)$, and $(\mu_{1S_n}, \mu_{2S_n}) = (0.0, 0.01)$. Also the a priori estimate of $\underline{\mu}_0$ was set at $\underline{\mu}_0' = \underline{0}$.

In Figures 6 through 14 the subscript v on an estimate refers to an estimate determined by the pseudo-minimum variance recursive moment estimation algorithm. The subscript B refers to one determined by the Bayesian recursive moment estimation algorithm. In each figure the ratio of the estimate to the corresponding element of $\underline{\mu}_n$ is presented.

In the first set of simulations, Figures 6 through 12, Ψ_0' was fixed at $\Psi_0' = 0.1 \times 10^{10} I$. This has the effect of removing the a priori estimate, $\underline{\mu}_0' = \underline{0}$, from the pseudo-minimum variance estimate. Three simulations were performed with values for $\mu_{2,0}$ of 0.01, 1, and 4, respectively. Figures 6 through 8 depict the results of the simulation with $\mu_{2,0} = 0.01$. A complete set of curves is presented showing the estimates of $\mu_{1,n}$, $\mu_{2,n}$ and $\mu_{3,n}$. In Figures 9 and 10 only estimates of $\mu_{2,n}$ and $\mu_{3,n}$ are presented when $\mu_{2,0} = 1$ since the estimation of $\mu_{1,n}$ in this case was, pictorially, essentially the same as that in Figure 6. Likewise, for the same reason when $\mu_{2,0} = 4$ only estimates of $\mu_{2,n}$ and $\mu_{3,n}$ are presented. See Figures 11 and 12.

The last set of figures, Figures 13 and 14, present the results of a simulation in which $\mu_{2,0} = 1.0$ and $\Psi_0' = I$. This value of Ψ_0'

causes the pseudo-minimum variance recursive moment estimation algorithm to weight the a priori estimate of $\underline{\mu}_0$ so highly that the data estimates of $\mu_{2,n}$ and $\mu_{3,n}$ are considered only slightly. Again pictorial presentation of the estimates of $\mu_{1,n}$ was so much like Figure 6 that it was omitted.

In Figure 6 where $\mu_{2,0} = 0.01$ and $\Psi_0' = 0.1 \times 10^{10} I$, the fact that the Bayesian estimation algorithm is an averaging of the projections of the a priori estimate, $\underline{\mu}_0'$, and the data estimates, $\underline{\mu}_i^*$, $i = 0,1,\ldots,n$, is clearly demonstrated. From Equation D.6.3 with $w_0' = 1$

$$\hat{\mu}_{1,0} = \frac{1}{2}\,\mu_{1,0}' + \frac{1}{2}\,\mu_{1,0}^*$$

From the simulation $\mu_{1,0}' = 0.0$ and $\mu_{1,0}^* = 9.998$ so that

$$\hat{\mu}_{1,0} = \frac{1}{2}(0.0) + \frac{1}{2}(9.998) = 4.999$$

and

$$\hat{\mu}_{1,0B}/\mu_{1,0} = 4.999/10 = 0.4999$$

which is verified in Figure 6. From the augmented moment model

$$\mu_{1,1}' = \mu_{1C_1}\,\hat{\mu}_{1,0} + \mu_{1S_1}$$

From the simulation $\mu_{1C_1} = 1.0$ and $\mu_{1S_1} = 0.0$, so that

$$\mu_{1,1}' = 1.0(4.999) + 0.0 = 4.999$$

and

$$\mu_{1,1B}'/\mu_{1,1} = 0.4999$$

Then since $\mu_{1,1}^* = 10.06$

$$\hat{\mu}_{1,1} = \frac{2}{3}\mu'_{1,1} + \frac{1}{3}\mu^*_{1,1} = \frac{2}{3}(4.999) + \frac{1}{3}(10.06)$$

$$= \frac{1}{3}(0.0) + \frac{1}{3}(9.998) + \frac{1}{3}(10.06)$$

$$= 6.686$$

so that

$$\hat{\mu}_{1,1B}/\mu_{1,1} = 0.6686$$

etc.

Likewise the fact that the larger a priori error covariance matrix Ψ'_0 causes the a priori estimate $\mu'_{1,0}$ to be of little effect in the pseudo-minimum variance estimate is obvious since the data estimate $\mu^*_{1,0}$ and the pseudo-minimum variance estimate $\hat{\mu}_{1,0}$ are the same value of 9.998.

In the pseudo-minimum variance estimate, since the sample mean, $\mu^*_{1,0}$, is such a good estimate of $\mu_{1,0}$ (the estimated error covariance between $\mu^*_{1,0}$ and $\mu_{1,0}$ is 0.002), its value of 9.998 is essentially projected through the augmented moment model and used as the estimate of $\mu_{1,n}$ at each value of n. Note in Figure 6 that there is a slight change in $\mu'_{1,nv}/\mu_{1,n}$ at n = 3 and n = 20. At these values of n the estimated error covariance of $\mu^*_{1,n}$ is small enough in comparison to the estimated error covariance of $\mu'_{1,n}$ that $\mu^*_{1,n}$ slightly modifies the estimate of $\mu_{1,n}$. Otherwise $\mu^*_{1,n}$ produces no noticeable change in $\mu'_{1,n}$ and $\mu'_{1,n}$ becomes $\hat{\mu}_{1,n}$.

In each of the other two estimations, Figures 7 and 9 in this simulation, the averaging property of the Bayesian algorithm is not quite so obvious, since the projection of the estimates through the augmented moment model also involves the estimates of the augmenting moment terms $\mu^2_{1,n}, \mu^3_{1,n}$ and $\mu_{2,n}\mu_{1,n}$. It is noticeable that the data estimates, $\mu^*_{2,n}$ and $\mu^*_{3,n}$, vary much more from $\mu_{2,n}$ and $\mu_{3,n}$ than $\mu^*_{1,n}$ does from $\mu_{1,n}$ in Figure 6, causing more variation in $\hat{\mu}_{2,n}$ and $\hat{\mu}_{3,n}$. This is to be expected since $\mu^*_{2,n}$ and $\mu^*_{3,n}$ have larger variances than $\mu^*_{1,n}$.

Figure 7 does show that in the pseudo-minimum variance estimation the large Ψ'_0 causes $\hat{\mu}_{2,0}$ to be $\mu^*_{2,0}$. It also indicates that at n = 1 the estimate $\mu'_{2,1}$ is such a good estimate of $\mu_{2,1}$ that $\hat{\mu}_{2,1}$ is essentially $\mu'_{2,1}$ and that for n = 2,3,... the estimate of $\mu_{2,n}$ is essentially $\hat{\mu}_{2,1}$ projected through the augmented moment model.

In Figure 8 as in each figure depicting the estimates of $\mu_{3,n}$, the estimates at n = 0 are not accurately presented. The figures indicate that the estimates of $\mu_{3,0}$ are all zero. This occurs since the presentation is that of the estimate divided by $\mu_{3,0}$. Since it was assumed that X_0 was normally distributed $\mu_{3,0} = 0$. In the simulation the computer tried to divide by zero but instead of giving an answer of infinite or stopping the simulation the ratio was evaluated as zero. Thus $\mu_{3,0V}/\mu_{3,0}$, $\mu_{3,0B}/\mu_{3,0}$ and $\mu^*_{3,0}/\mu_{3,0}$ all appear to be zero. However for larger values of n the presentation is accurate.

Figure 8 indicates that at n = 2 the pseudo-minimum variance estimate $\mu'_{3,2}$ is such a good estimate of $\mu_{3,2}$ that the following estimates of $\mu_{3,n}$ are essentially the projections of $\mu'_{3,2}$. Notice that for some larger values of n, e.g., n = 23 and 43, the value of $\mu^*_{3,n}$ exerts a slight influence on the value of $\hat{\mu}_{3,n}$ in the combination of $\mu'_{3,n}$ and $\mu^*_{3,n}$.

Figures 9 through 12 present results from simulations in which

$\mu_{2,0} = 1$ and $\mu_{2,0} = 4$. These are similar to Figures 7 and 8. However, it is obvious that the increased variance on X_0 affects the pseudo-minimum variance estimates of $\mu_{2,n}$ and $\mu_{3,n}$. These estimates do not approach $\mu_{2,n}$ and $\mu_{3,n}$ as quickly as they did in the first simulation. In fact the pseudo-minimum variance estimates of $\mu_{3,n}$ in Figures 10 and 12 appear to be following the trend of $\mu^*_{3,n}$ below and away from $\mu_{3,n}$ for large values of n. In Figures 9 and 11 both the Bayesian and the pseudo-minimum variance estimates of $\mu_{2,n}$ appear to reach a fairly steady percentage error for large values of n. In Figure 11 the Bayesian estimate of $\mu_{2,n}$ has a smaller error than the pseudo-minimum variance estimate for values of n above n = 15. This is also true for the estimates of $\mu_{3,n}$ in Figures 10 and 12.

In Figures 13 and 14 the results with $\Psi_0' = I$ indicate that the pseudo-minimum variance estimates of $\mu_{2,n}$ and $\mu_{3,n}$ approach values which are approximately 35% and 5% of $\mu_{2,n}$ and $\mu_{3,n}$, respectively. This algorithm has taken the a priori estimate $\underline{\mu}_0' = \underline{0}$ as a good estimate of $\underline{\mu}_0$ and essentially projected this value through the augmented moment model to determine $\hat{\mu}_{2,n}$ and $\hat{\mu}_{3,n}$. The slopes of the curves depicting the early estimates of $\mu_{2,n}$ and $\mu_{3,n}$ appear to be negative. This is not actually the case. For example $\hat{\mu}_{2,1}$ increases to $\mu'_{2,1}$ but it does not increase as much as $\mu_{2,1}$ to $\mu_{2,2}$. So that $\hat{\mu}_{2,1v}/\mu_{2,1}$ is actually larger than $\mu'_{2,2v}/\mu_{2,2}$.

It is interesting to note in Figure 14, that even though $\mu'_{3,0} = 0$ and $\mu_{3,0} = 0$ the estimate of $\mu_{3,n}$ for large n is not at all near the value of $\mu_{3,n}$. This is due to the fact that $\mu^3_{1,0}$ and $\mu_{2,0}\mu_{1,0}$ are not zero, but the a priori estimate of each is zero and with $\Psi_0' = I$ a false confidence is maintained in these estimates. These estimates essentially become the estimates $\hat{\mu}_{1,0}^3$ and $(\mu_{2,0}\hat{\mu}_{1,0})$ which are used in the augmented moment model to project $\hat{\mu}_{3,0}$ to $\mu_{3,1}'$. The error covariance of $\mu_{3,0}'$ is determined mainly from Ψ_0' and thus $\mu_{3,1}'$ is considered a good estimate of $\mu_{3,1}$. This error remains and is compounded as n increases.

Other simulations were performed for various parameter values. One simulation with $\mu_{2,0} = 1$ and $\Psi_0' = 1,000\ I$ and all other parameters the same as in Figures 6 through 14, produced results in which the pseudo-minimum variance estimates of $\mu_{1,n}$ were within 4% of $\mu_{1,n}$, the estimates of $\mu_{2,n}$ within 10% of $\mu_{2,n}$, but the estimates of $\mu_{3,n}$ for n greater than 20 were only about 50% of $\mu_{3,n}$. In another set of simulations with $\mu_{2,0} = 1$ and $\Psi_0' = 0.1 \times 10^{10} I$ all other parameters remained the same except for μ_{2S_n}. Simulations were performed for values of μ_{2S_n} of 1, 10, 100. Since μ_{2S_n} was a known value the results of these simulations were very similar to Figures 6, 9, and 10, except that the curves for $\mu_{1,n}^*/\mu_{1,n}$, $\mu_{2,n}^*/\mu_{1,n}$, and $\mu_{3,n}^*/\mu_{1,n}$ exhibited much more variation than they do in Figures 6, 9, and 10.

Thus for Case V of Section 3.3 where $\underline{A}_n$ and $\underline{\mu}_{S_n}$ are known, as simulated, $\mu_{2,n}$ the variance of the initial random variable X_0 and Ψ_0' the estimated error covariance matrix between $\underline{\mu}_0$ and $\underline{\mu}_0'$ appear to be critical parameters to the pseudo-minimum variance recursive moment estimation algorithm. Conversely, since the moments of C_n and S_n are known in this case, they are apparently not very critical.

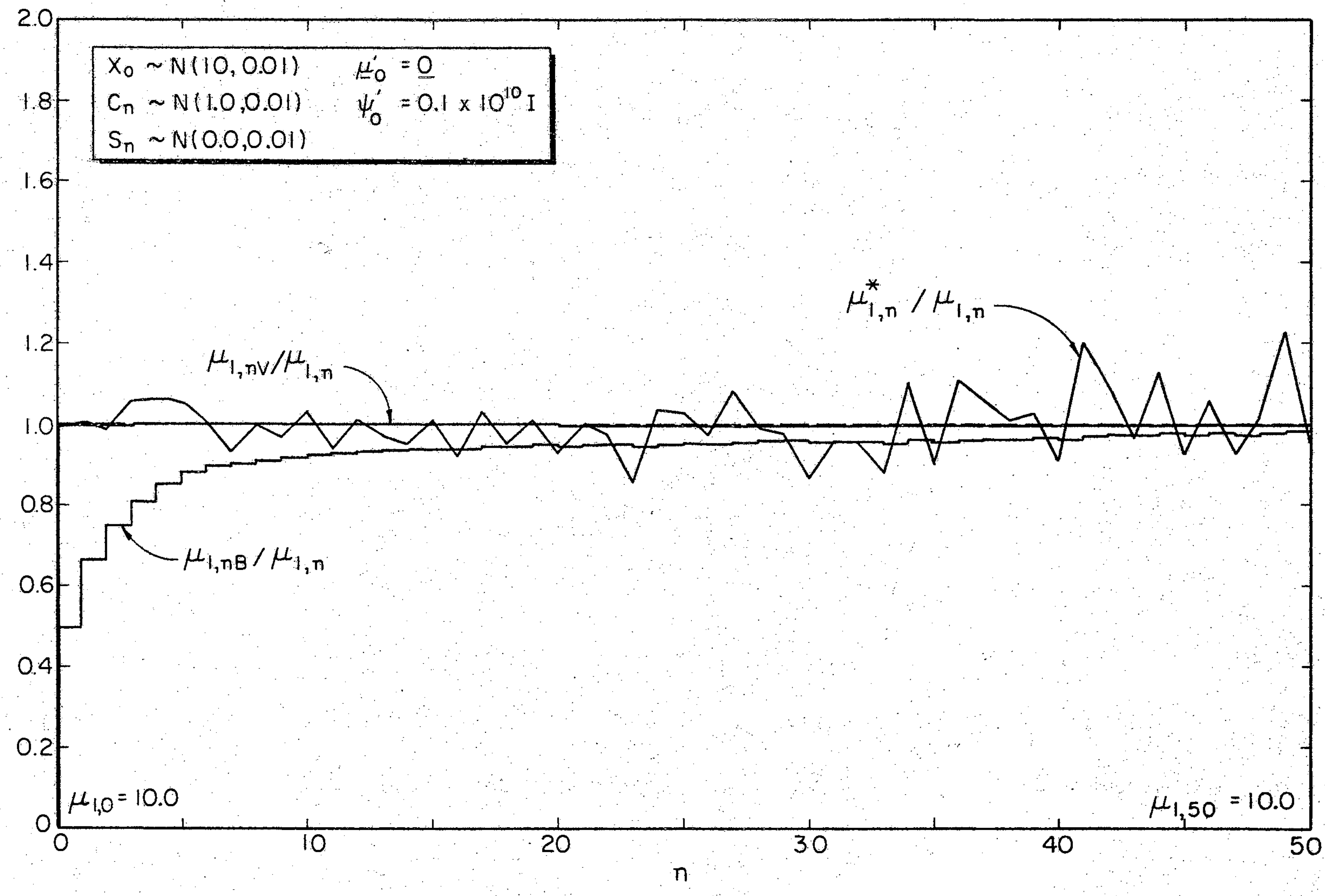

Figure 6. Estimates of $\mu_{1,n}$

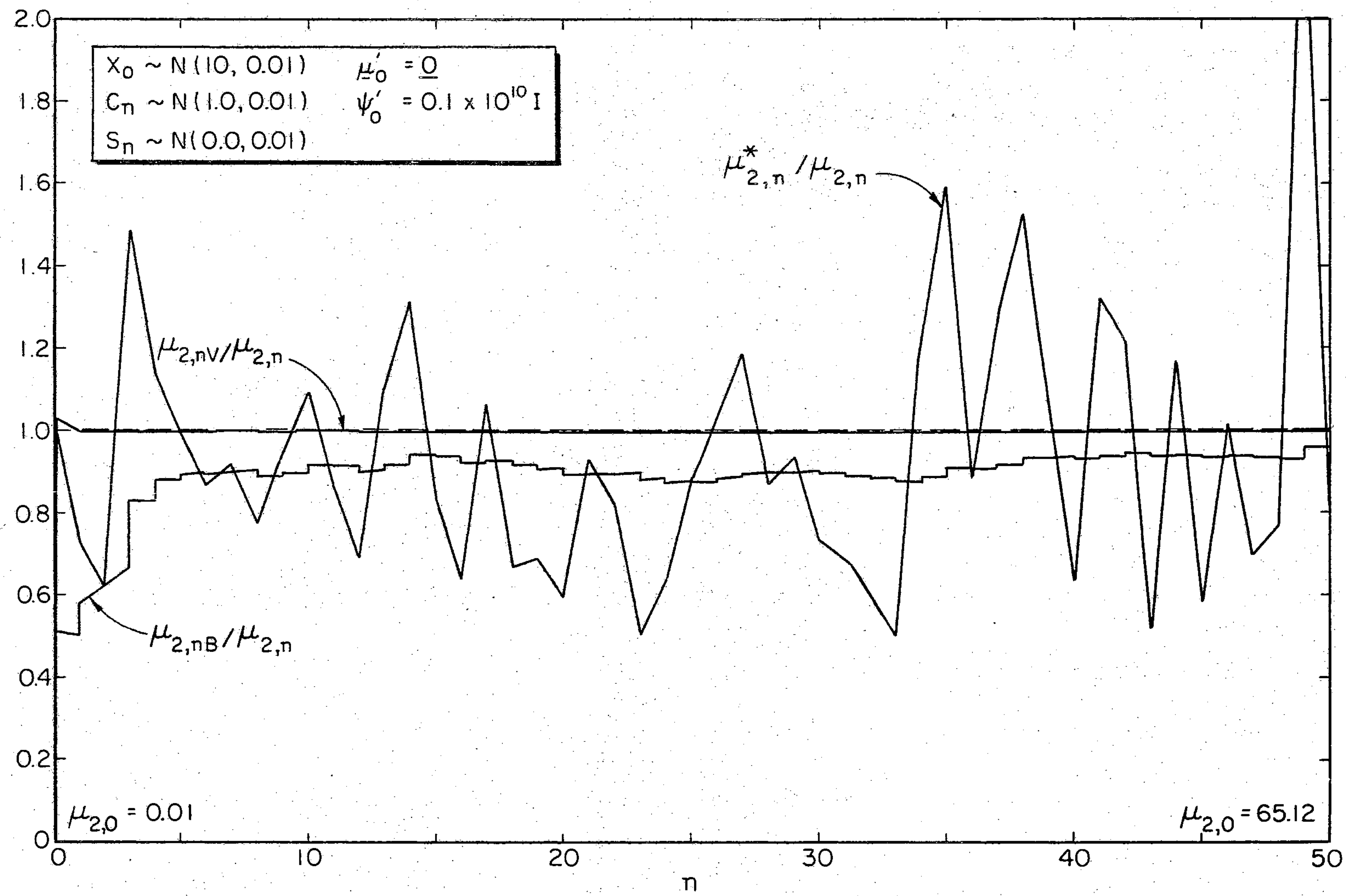

Figure 7. Estimates of $\mu_{2,n}$

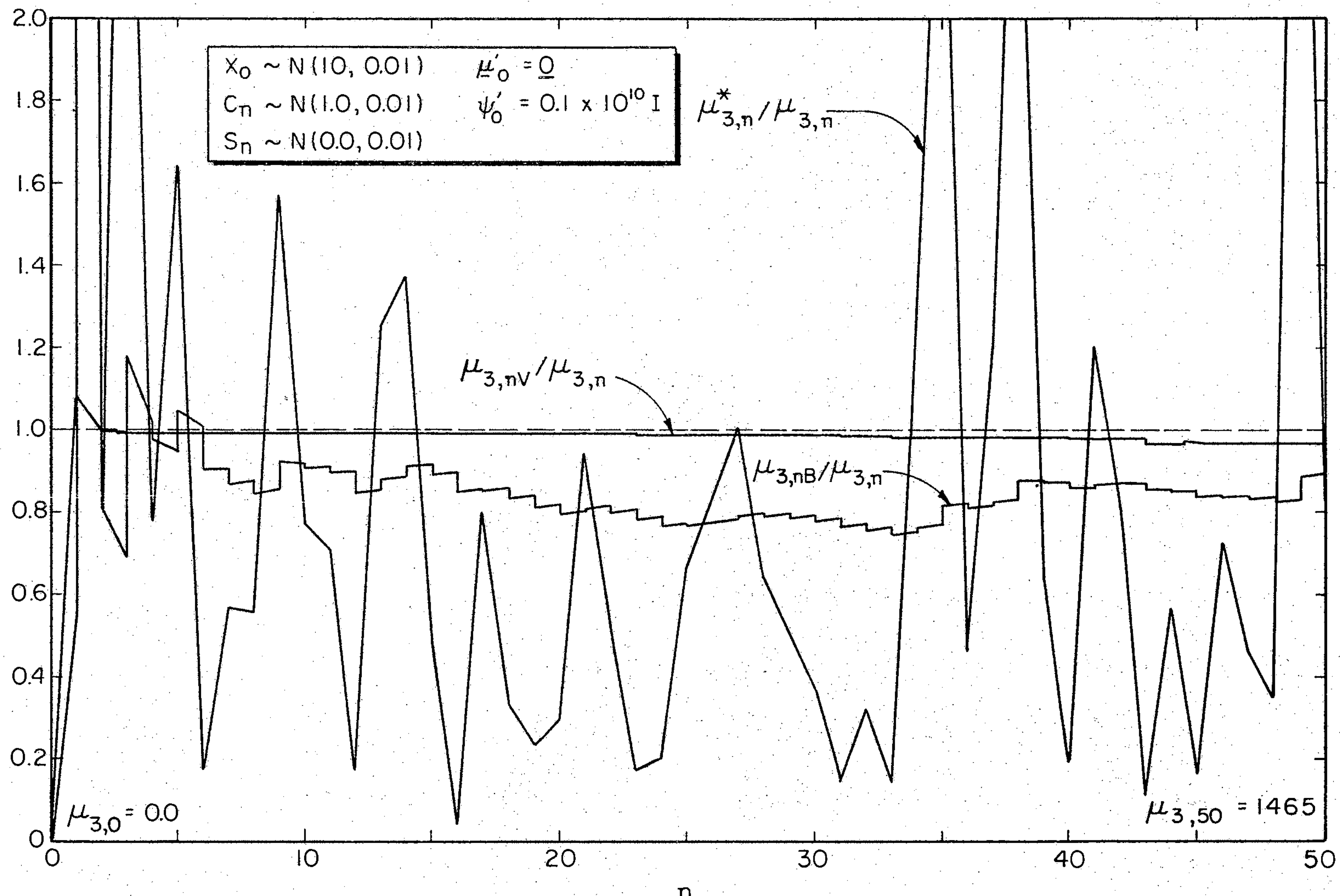

Figure 8. Estimates of $\mu_{3,n}$

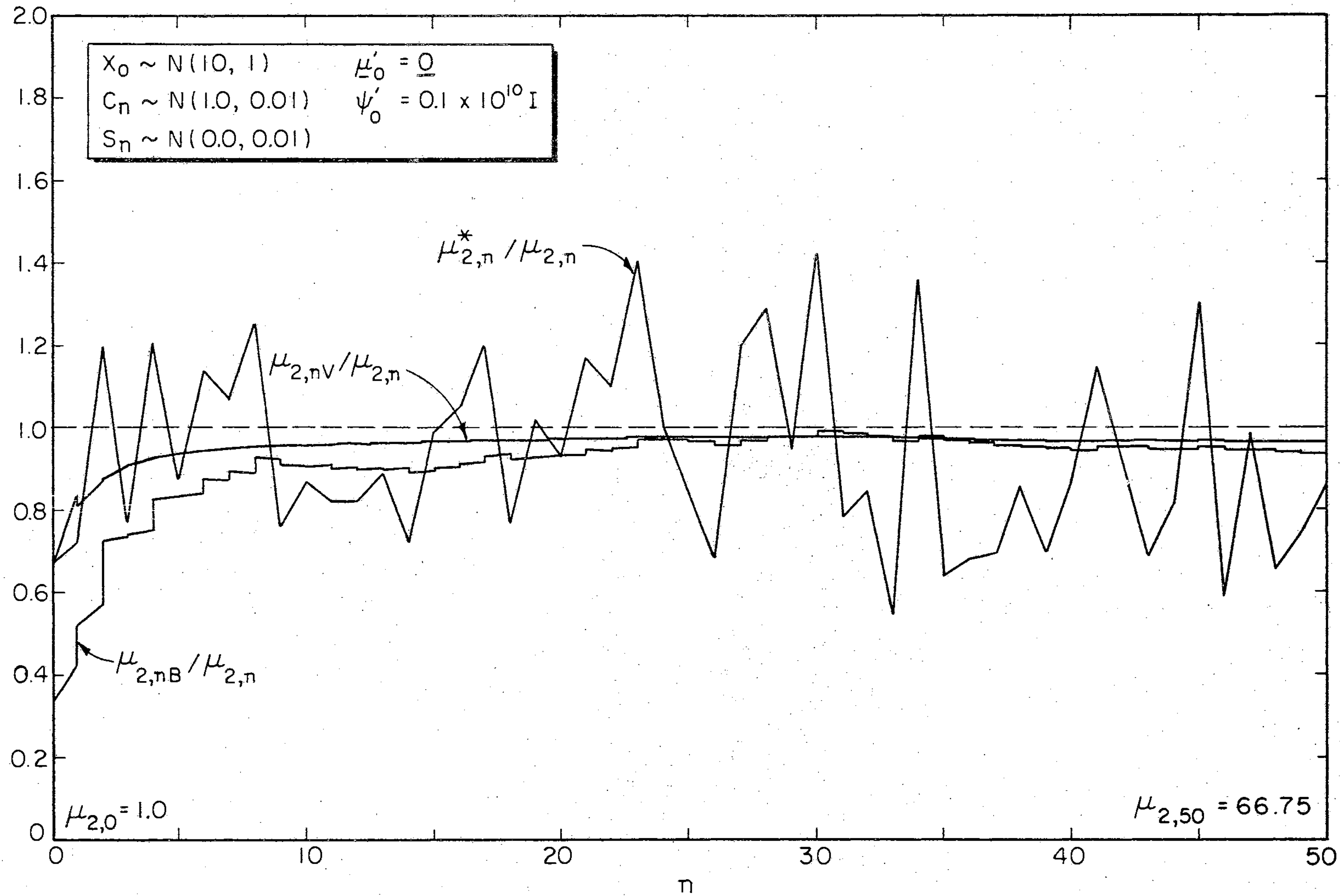

Figure 9. Estimates of $\mu_{2,n}$

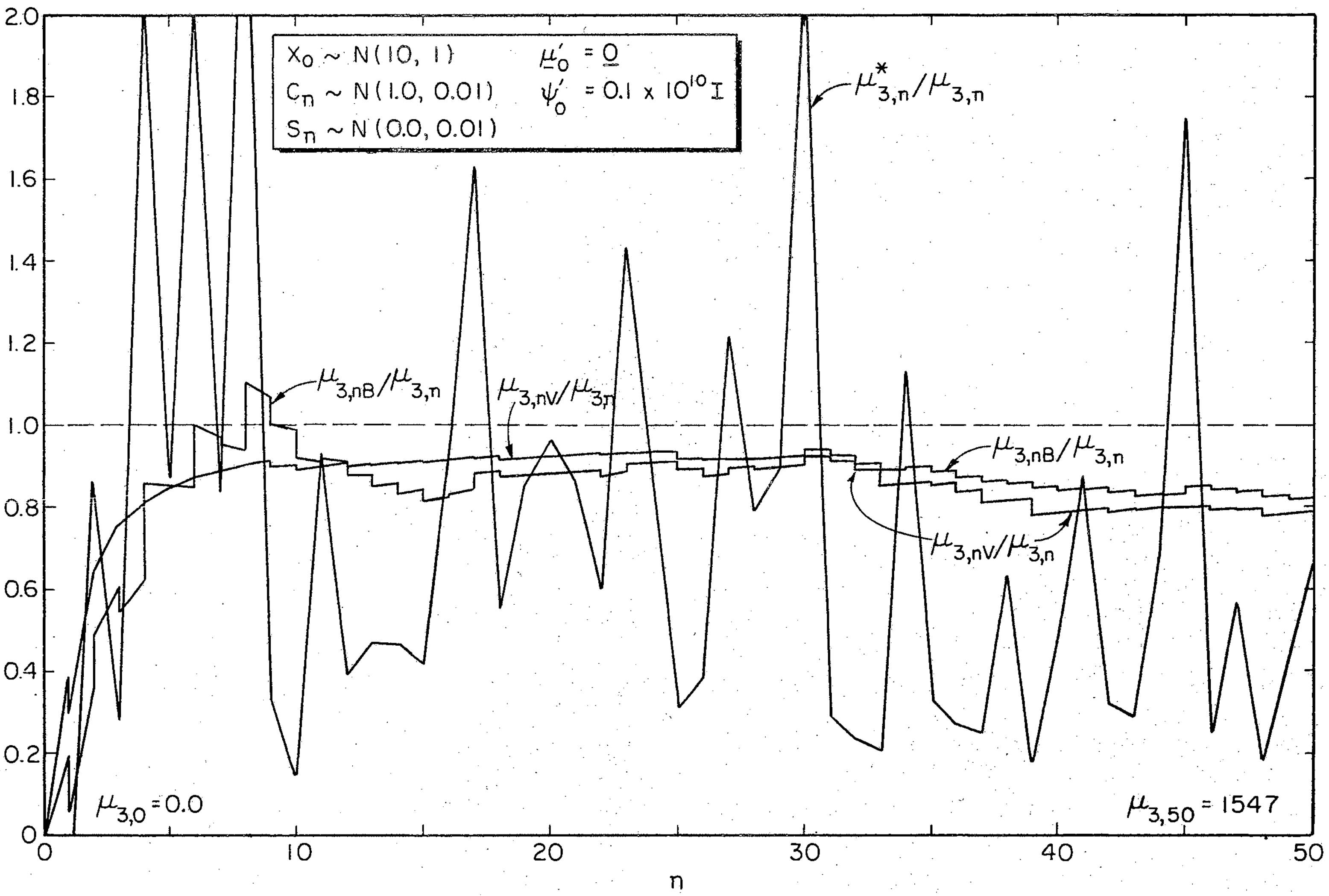

Figure 10. Estimates of $\mu_{3,n}$

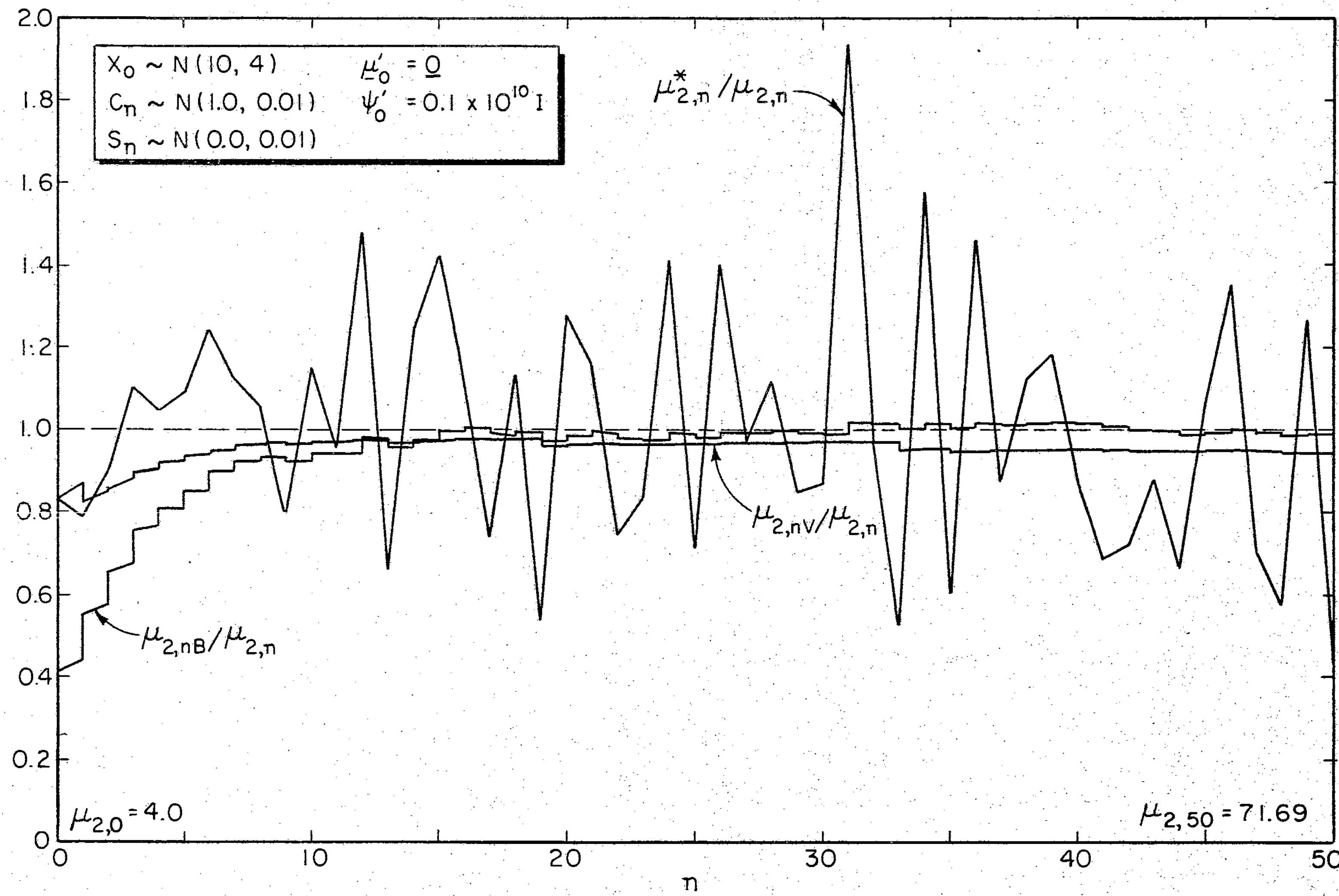

Figure 11. Estimates of $\mu_{2,n}$

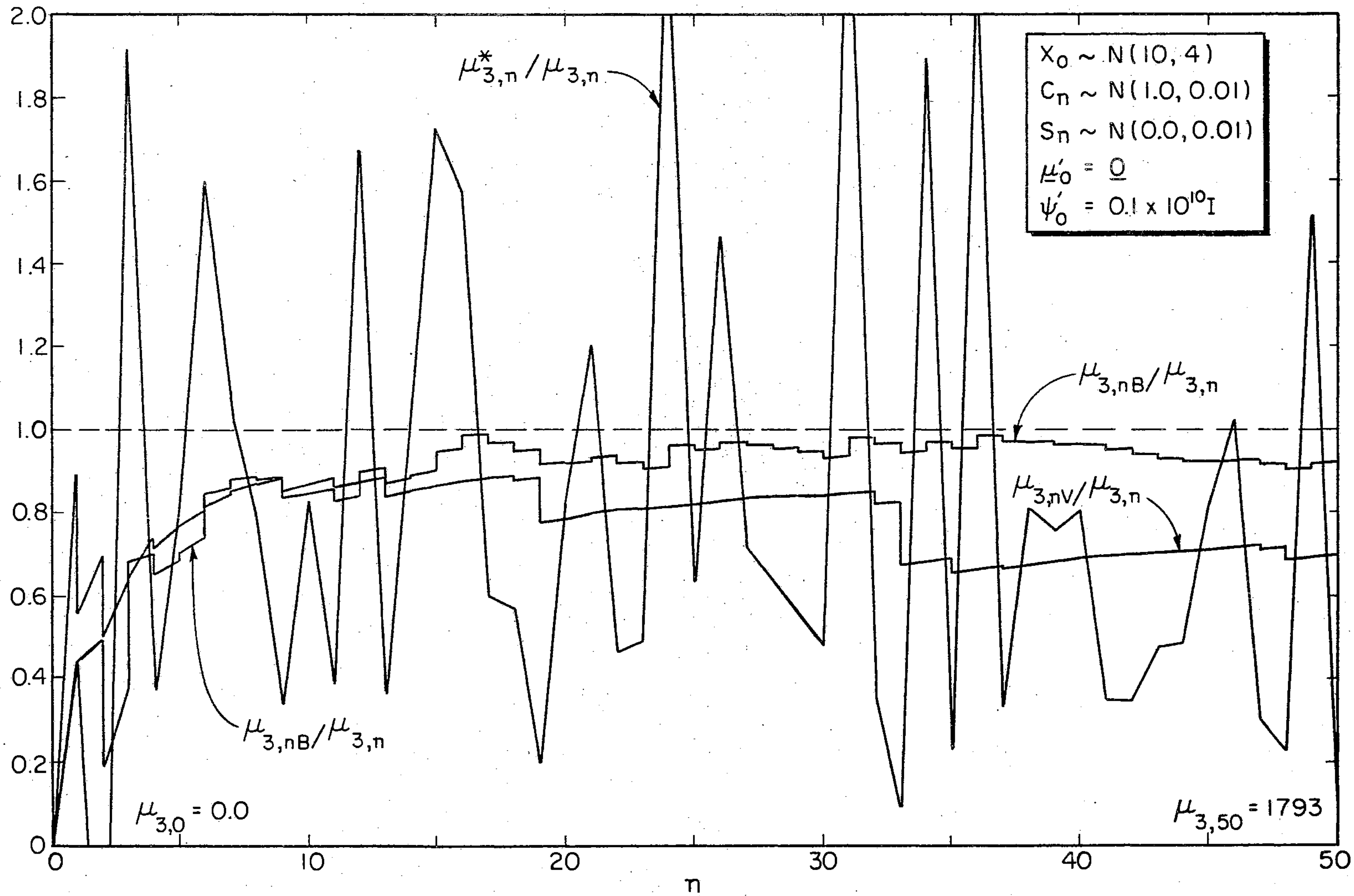

Figure 12. Estimates of $\mu_{3,n}$

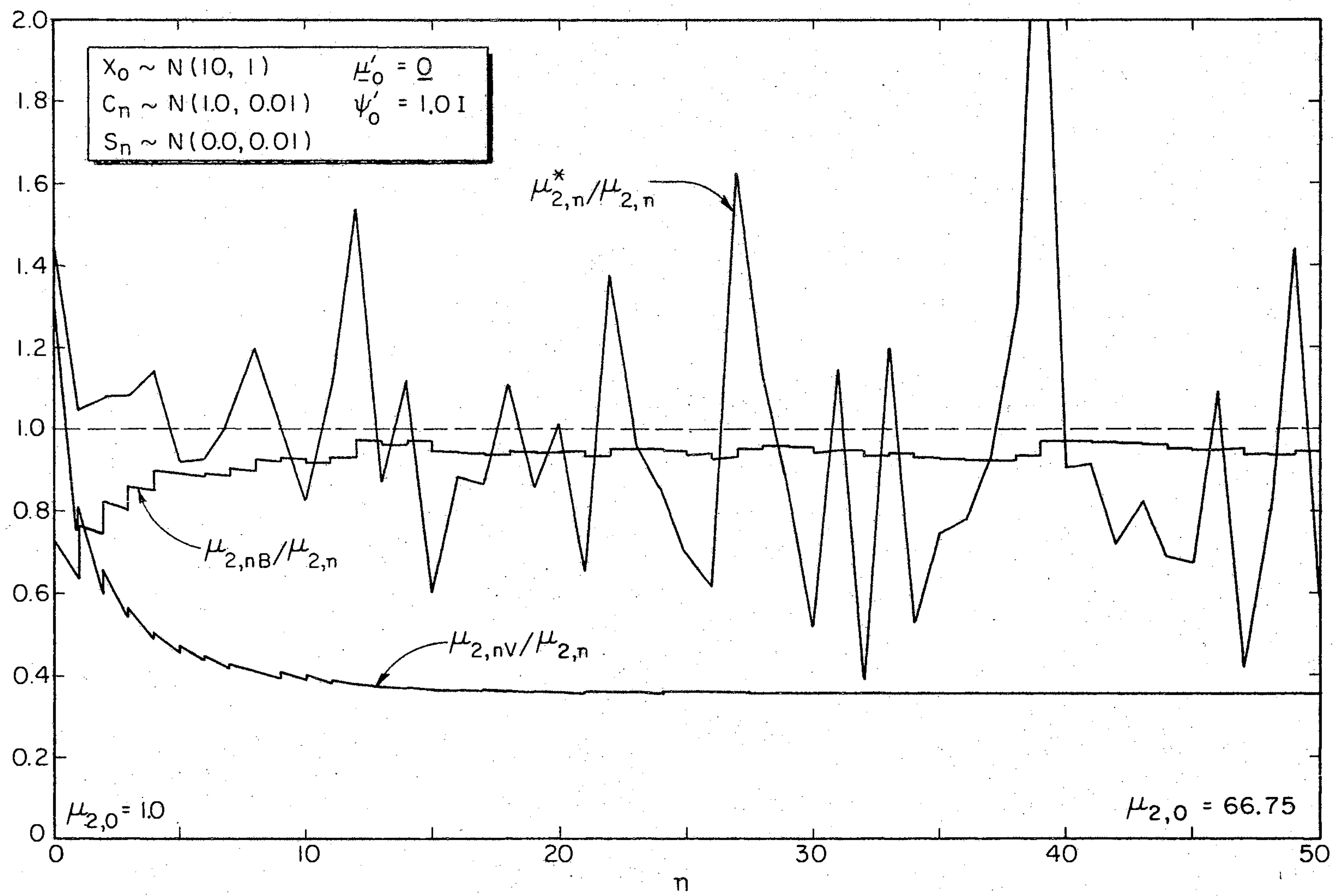

Figure 13. Estimates of $\mu_{2,n}$

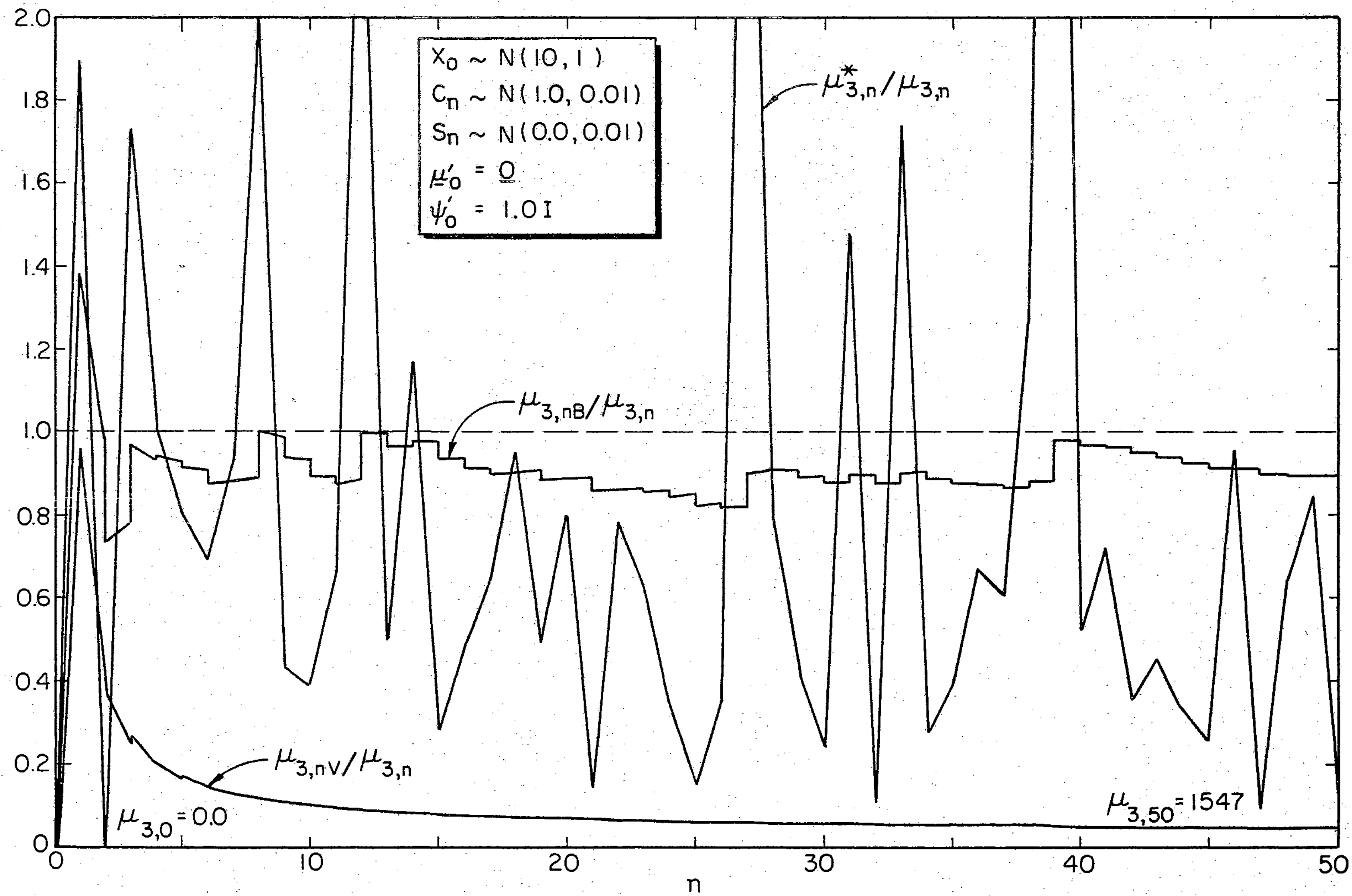

Figure 14. Estimates of $\mu_{3,n}$

CHAPTER V

SUMMARY AND CONCLUSIONS

5.1 Summary. The objective of this study was to develop a procedure for the estimation of the distribution function of a random variable representing time-varying equipment outputs. The Gram-Charlier or Edgeworth series expansions of the distribution function in terms of the moments of the random variable are often used to approximate the distribution function. For this type of approximation the problem was reduced to one of estimation of the moments of the time-varying random variable.

Two methods for the estimation of moments were developed. These make use of not only unbiased sample moments determined from system observations, but also a system model and a priori information.

Chapter II presents the development of a system model of time-varying equipment outputs and the subsequent derivation of a moment model. The system model used was a first-order linear difference equation and the resulting moment model was a first-order vector-matrix difference equation.

Chapter III presents the theoretical development of the recursive moment estimation scheme. This scheme makes use of the sample moments and the moment model to determine the best linear mean squared error estimate of the moments in terms of the a priori estimates and all the unbiased data estimates computed through the estimation time. Several difficulties which appeared in this development are discussed and an alternative approach, the pseudo-minimum variance recursive moment estimation scheme, is presented.

The other method for the estimation of moments, the Bayesian recursive moment estimation scheme, is presented in Appendix D. The Bayesian approach was an attempt to make use of a reproducing a priori density function in Bayes' Rule to estimate the moments.

Chapter IV discusses a simulating computer program and presents some typical results of simulations of the two methods of recursive moment estimation.

5.2 Conclusions. The procedure presented in Chapter II for the development of a system model is an appraoch which is useful in the modeling of time-varying equipment outputs. The form of the model is not unique, but with information on system behavior available only from life tests and system tests, the procedure is restricted to the development of a model with only two parameters.

For the derivation of the moment model from the system model it was assumed that the random variables of the system model were independent. This assumption may not always hold. In fact, in the example used in Chapter II to demonstrate the system model development, it is obvious that X_{n-1} and S_n are not independent.

Pseudo-minimum variance recursive moment estimation provides a means to make use of a system model, a priori information, and system observations to estimate moments. It makes use of at least estimates of the error covariance matrices between the estimates and the moments to be estimated in the weights necessary to combine estimates. In

Bayesian recursive moment estimation the weights are predetermined constants. The pseudo-minimum variance recursive moment estimates are modified minimum mean squared error estimates, while the Bayesian recursive estimates are averages of projected estimates. As a result of this the pseudo-minimum variance estimates tend to approach the moments faster than the Bayesian estimates.

The derivation of the pseudo-minimum variance recursive moment estimation algorithm is not unique to the system model chosen or the resulting moment model. However the form of the algorithm no doubt will change with a change in models.

The pseudo-minimum variance recursive moment estimation algorithm is easily implemented on a digital computer, both for simulation and actual use with a system in operation.

The pseudo-minimum variance estimates are better estimates than the sample moments in a modified mean squared error sense. Thus the pseudo-minimum variance moment estimates will, in this same sense, yield better results in Gram-Charlier or Edgeworth series approximations to the distribution function.

When the pseudo-minimum variance moment estimation is used some thought should be given to the choice of the system model. It may well be that a system model different from that used in this study is more realistic and may even produce a simpler algorithm. See Section 5.3.

The pseudo-minimum variance recursive moment estimation algorithm does have some limitations which should be noted. The algorithm is no better than the system model. The model can reflect only the variations which are observed in life test and system test. Changes in the system during operation which depart from these, such as catastrophic failures, can not be modeled. When it becomes apparent that something of this nature has occurred other tests are required to determine the necessary model changes before continuing. Just as with sample moments the moment estimates are more accurate when determined from more data. This is reflected both in the moment model development and the computation of sample moments. In Chapter II it is indicated that higher order moments of the system random variables are difficult to obtain. Using a fixed amount of data in many cases the higher order sample moments will be less accurate estimates than the lower order sample moments. This inaccuracy of higher order estimates is clearly indicated in the simulation results presented in Chapter IV.

The sample moments were used to develop unbiased data estimates. Several unbiased estimates of higher order moments and products of moments were derived in Appendix C. These estimates, which as far as the author could find are not available in the literature, may be of some use in other areas of endeavor.

5.3 Recommendations for Further Study. As indicated in earlier remarks of this chapter some consideration should be given to the development of other system models and the resulting moment models. For example if the system model developed was of the form

$$X_n = X_{n-1} + C_n + S_n$$

and the random variables X_{n-1}, C_n, and S_n were independent the resulting moment model (for three moments) would be

$$\underline{\mu}_n = \underline{\mu}_{n-1} + \underline{\mu}_{C_n} + \underline{\mu}_{S_n}$$

where

$$\underline{\mu}_n = \begin{bmatrix} \mu_{1,n} \\ \mu_{2,n} \\ \mu_{3,n} \end{bmatrix}, \quad \underline{\mu}_{C_n} = \begin{bmatrix} \mu_{1C_n} \\ \mu_{2C_n} \\ \mu_{3C_n} \end{bmatrix}, \text{ and } \underline{\mu}_{S_n} = \begin{bmatrix} \mu_{1S_n} \\ \mu_{2S_n} \\ \mu_{3S_n} \end{bmatrix}$$

In this moment model the moment vectors are not augmented. The error covariance matrices, Ψ_n', Ψ_n^*, and $\hat{\Psi}_n$ are all (3 x 3). Thus the estimate $\hat{\Psi}_n^*$ would be easier to obtain; requiring fewer unbiased estimates of higher order moments and products of moments.

Even though it is felt that when using a series expansion to approximate the distribution function the pseudo-minimum variance estimation scheme produces the best moment estimates some consideration should be given to other approaches. Approximation of the distribution function might be accomplished by constructing an empirical distribution function. Either this empirical distribution function could be constructed from all system observations at all sampling times through the present by projection of the observations through the system model, or empirical distribution functions could be constructed at each sampling time and then through some form of the system model these distribution functions projected and combined.

BIBLIOGRAPHY

1. Abramson, N. and D. Braverman. "Learning to Recognize Patterns in a Random Environment." IRE Transactions on Information Theory. Vol. IT-8. (September, 1962).

2. Cramér, Harald. Mathematical Methods of Statistics. Princeton: Princeton University Press, 1946.

3. Deutsch, Ralph. Estimation Theory. Englewood Cliffs: Prentice-Hall, 1965.

4. Hogg, R.V. and A.T. Craig. Introduction to Mathematical Statistics. New York: Macmillan, 1965.

5. Fraser, D.A.S. Nonparametric Methods in Statistics. New York: John Wiley and Sons, 1957.

6. Fraser, D.A.S. Statistics - An Introduction. New York: John Wiley and Sons, 1958.

7. Kalman, R.E. "A New Approach to Linear Filtering and Prediction Problems." Transactions of the ASME: Journal of Basic Engineering. Vol. 82. (March, 1960).

8. Keehn, D.G. "Learning the Mean Vector and Covariance Matrix of Gaussian Signals in Pattern Recognition." Stanford Electronics Laboratories Technical Report 2003-6. Stanford: Stanford University, 1963.

9. Nilsson, Nils J. Learning Machines. New York: McGraw-Hill, 1965.

10. Papoulis, Athanasois. Probability, Random Variables, and Stochastic Processes. New York: McGraw-Hill, 1965.

11. Spragins, John. "A Note on the Iterative Application of Bayes' Rule". IEEE Transactions on Information Theory. Vol. IT-11. (October, 1965).

12. Tchouproff, A.A. "On the Mathematical Expectation of the Moments of Frequency Distributions." Biometrika. Vol. 12. (1918-1919), and Biometrika. Vol. 13. (1921).

APPENDIX A

THE GRAM-CHARLIER SERIES AND THE EDGEWORTH SERIES EXPANSIONS OF A DISTRIBUTION FUNCTION

Consider the standardized random variable $y = \frac{X - \mu_1}{\sqrt{\mu_2}}$ where $\mu_1 = E\{X\}$ and $\mu_2 = E\{[X - \mu_1]^2\}$. The density function, $f(y)$, of Y is given by Cramér (2) expanded in a Gram-Charlier series as

$$f(y) = C_0\phi(y) + \frac{C_1}{1!}\phi^{(1)}(y) + \frac{C_2}{2!}\phi^{(2)}(y) + \frac{C_3}{3!}\phi^{(3)}(y) + \cdots \tag{A.1}$$

where C_r, $r = 1,2,\ldots$, are constant coefficients,

$$\phi(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}, \text{ the normal density function, } N(0,1),$$

and

$$\phi^{(r)}(y) = \frac{d^r}{dy^r}\phi(y), \quad r = 1,2,\ldots$$

The derivatives of the normal density function are given by

$$\phi^{(r)}(y) = (-1)^r H_r(y)\,\phi(y)$$

where $H_r(y)$, $r = 1,2,\ldots$, are the Hermite polynomials. The Hermite polynomials are defined by

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}, \quad n = 0,1,2,\ldots$$

the first few of which are given by

$$H_0 = 1$$

$$H_1 = y$$

$$H_2 = y^2 - 1$$

$$H_3 = y^3 - 3y$$

$$H_4 = y^4 - 6y^2 + 3$$

The constant coefficients, c_r are given by

$$c_r = (-1)^r \int_{-\infty}^{+\infty} H_r(y) f(y) dy$$

the first few of which are

$$C_0 = 1$$

$$C_1 = C_2 = 0$$

$$C_3 = -\frac{\mu_3}{(\mu_2)^{3/2}}$$

$$C_4 = \frac{\mu_4}{\mu_2^2} - 3$$

Since $C_0 = 1$ and $C_1 = C_2 = 0$, the Gram-Charlier series expansion of f(y) becomes

$$f(y) = \phi(y) + \frac{C_3}{3!} \phi^{(3)}(y) + \frac{C_4}{4!} \phi^{(4)}(y) + \cdots \qquad (A.2)$$

It can be shown that under certain conditions, Equation A.2 will converge to the true density function of Y (2). However Carmér (2) shows that generally the Gram-Charlier series is not an asymptotic expansion, i.e., addition of another term to an approximation using a finite number of terms in the Gram-Charlier series does not necessarily reduce the error between the approximation and the true density function.

The Edgeworth series expansion of f(y) is given by

$$f(y) = \phi(x) - \frac{1}{3!}\frac{\mu_3}{(\mu_2)^{3/2}}\phi^{(3)}(x) + \frac{1}{4!}\left(\frac{\mu_4}{\mu_2^2} - 3\right)\phi^{(4)}(x) + \frac{10}{6!}\left(\frac{\mu_3}{(\mu_2)^{3/2}}\right)^2\phi^{(6)}(x) + \cdots \tag{A.3}$$

The development and additional terms of the Edgeworth series may be found in Cramér (2). The Edgeworth series, unlike the Gram-Charlier series, is, under fairly general conditions, an asymptotic expansion.

Since in this study only the first three moments are used, only the first two terms (through the third order terms) of either the Gram-Charlier or the Edgeworth series can be used. Under this restriction the approximations of f(y) by both the Gram-Charlier and the Edgeworth series are identical. Therefore f(y) is approximately given by

$$f(y) \doteq \phi(y) - \frac{1}{3!}\frac{\mu_3}{(\mu_2)^{3/2}}\phi^{(3)}(y)$$

$$\doteq \phi(y) + \frac{1}{3!}\left(\frac{\mu_3}{(\mu_2)^{3/2}}\right)(y^3 - 3y)\,\phi(y) \tag{A.4}$$

and the distribution function of y is

$$F(y) \doteq \Phi(y) - \frac{1}{3!}\left(\frac{\mu_3}{(\mu_2)^{3/2}}\right)(y^2 - 1)\,\phi(y) \tag{A.5}$$

where

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-z^2/2} dz$$

Before such an approximation, Equation A.4 or A.5, can be used in the context of this study a modification must be made. This modification is necessary; for since the mean and the variance of X, the random variable under consideration, are unknown a standardized random variable can not be used. This modification is accomplished by performing a change of variables. Since $Y = \frac{X - \mu_1}{\sqrt{2\pi}}$,

$$X = \sqrt{\mu_2}\, Y - \mu_1$$

$$f_X(x) = \frac{f_Y(y)}{\left|\frac{dx}{dy}\right|} = \frac{1}{\sqrt{\mu_2}} f_Y(y = \frac{x - \mu_1}{\sqrt{\mu_2}})$$

$$\doteq \phi(x;\ \mu_1,\ \mu_2)$$

$$+ \frac{1}{3!}\left(\frac{\mu_3}{\mu_2^3}\right)[x^3 - 3\mu_1 x^2 + 3(\mu_1^2 - \mu_2)x - \mu_1^3 + 3\mu_1\mu_2]\phi(x;\mu_1,\mu_2) \qquad \text{(A.6)}$$

and

$$F_X(x) = F_Y\left(\frac{x - \mu_1}{\sqrt{\mu_2}}\right)$$

$$\doteq \Phi(x;\ \mu_1,\mu_2)$$

$$- \frac{1}{3!}\left(\frac{\mu_3}{(\mu_2)^{3/2}}\right)[x^2 - 2x\mu_1 + \mu_1^2 - \mu_2]\ \phi(x;\ \mu_1,\mu_2) \qquad \text{(A.7)}$$

where

$$\phi(x;\ \mu_1,\mu_2) = \frac{1}{\sqrt{2\pi\,\mu_2}}\, e^{-\frac{(x-\mu_1)^2}{2\mu_2}}$$

and

$$\Phi(x; \mu_1, \mu_2) = \frac{1}{\sqrt{2\pi\, \mu_2}} \int_{-\infty}^{x} e^{-\frac{(z-\mu_1)^2}{2\mu_2}} \, dz$$

Equation A.7 can be used to approximate F(x) by using the estimates of μ_1, μ_2, and μ_3 developed in the body of this thesis.

APPENDIX B

MOMENTS THROUGH THE SYSTEM MODEL

B.1 Introduction. The System model

$$X_n = C_n X_{n-1} + S_n \tag{B.1.1}$$

can be considered as a model of the transition of the random variable X_{n-1} to the random variable X_n. In this appendix, assuming that C_n, X_{n-1}, and S_n are independent random variables, the relationships between the first (mean), second (variance), and third central moments of X_n and X_{n-1} are established.

B.2 The Mean, $\mu_{1,n}$. Since C_n, X_{n-1}, and S_n are independent random variables

$$\mu_{1,n} = E\{X_n\} = E\{C_n X_{n-1} + S_n\} = E\{C_n\}E\{X_{n-1}\} + E\{S_n\}$$

$$= \mu_{1C_n}\mu_{1,n-1} + \mu_{1S_n} \tag{B.2.1}$$

where μ_{1C_n} is the mean of C_n, $\mu_{1,n-1}$ is the mean of X_{n-1}, etc.

B.3 The Variance, $\mu_{2,n}$. The variance, the second central moment, of X_n is

$$\mu_{2,n} = E\{[X_n - \mu_{1,n}]^2\} = E\{[C_n X_{n-1} + S_n - \mu_{1C_n}\mu_{1,n-1} - \mu_{1S_n}]^2\}$$

$$= E\{[(C_n X_{n-1} - \mu_{1C_n}\mu_{1,n-1}) + (S_n - \mu_{1S_n})]^2\} \tag{B.3.1}$$

which due to the independence of C_n, X_{n-1}, and S_n is the variance of C_nX_{n-1} plus the variance of S_n

$$\mu_{2,n} = E\{[C_nX_{n-1} - \mu_{1C_n}\mu_{1,n-1}]^2\} + E\{[S_n - \mu_{1S_n}]^2\} \qquad (B.3.2)$$

Recalling the relations between the variance of the product of two independent random variables and the moments of each random variable, $\mu_{2,n}$ becomes

$$\mu_{2,n} = \mu_{2C_n}\mu_{2,n-1} + \mu_{1C_n}^2\mu_{2,n-1} + \mu_{2C_n}\mu_{1,n-1}^2 + \mu_{2S_n}$$

$$= [\mu_{2C_n} + \mu_{1C_n}^2]\,\mu_{2,n-1} + \mu_{2C_n}\mu_{1,n-1}^2 + \mu_{2S_n} \qquad (B.3.3)$$

where μ_{2C_n} is the variance of C_n, $\mu_{2,n-1}$ is the variance of X_{n-1}, etc.

<u>B.4 The Third Central Moment, $\mu_{3,n}$.</u> The third central moment of X_n is

$$\mu_{3,n} = E\{[X_n - \mu_{1,n}]^3\} = E\{[C_nX_{n-1} + S_n - \mu_{1C_n}\mu_{1,n-1} - \mu_{1S_n}]^3\}$$

$$= E\{[(C_nX_{n-1} - \mu_{1C_n}\mu_{1,n-1}) + (S_n - \mu_{1S_n})]^3\} \qquad (B.4.1)$$

which due to the independence of C_n, X_{n-1}, and S_n is the third central moment of C_nX_{n-1} plus the third central moment of S_n

$$\mu_{3,n} = E\{[C_nX_{n-1} - \mu_{1C_n}\mu_{1,n-1}]^3\} + E\{[S_n - \mu_{1S_n}]^3\} \qquad (B.4.2)$$

Expanding $E\{[C_nX_{n-1} - \mu_{1C_n}\mu_{1,n-1}]^3\}$

$$E\{[C_nX_{n-1} - \mu_{1C_n}\mu_{1,n-1}]^3\} = E\{C_n^3X_{n-1}^3 - 3C_n^2X_{n-1}^2\mu_{1C_n}\mu_{1,n-1}$$

$$+ 3C_nX_{n-1}\mu_{1C_n}^2\mu_{1,n-1}^2 - \mu_{1C_n}^3\mu_{1,n-1}^3\}$$

$$= E\{C_n^3\}E\{X_{n-1}^3\} - 3E\{C_n^2\}E\{X_{n-1}^2\}\mu_{1C_n}\mu_{1,n-1} + 2\mu_{1C_n}^3\mu_{1,n-1}^3 \qquad (B.4.3)$$

Recalling that

$$\mu_{3,n-1} = E\{X_{n-1}^3\} - 3E\{X_{n-1}^2\}\,\mu_{1,n-1} + 2\mu_{1,n-1}^3$$

and

$$\mu_{2,n-1} = E\{X_{n-1}^2\} - \mu_{1,n-1}^2$$

and using the same relations for μ_{3C_n} and μ_{2C_n}

$$E\{[C_n X_{n-1} - \mu_{1C_n}\mu_{1,n-1}]^3\}$$

$$= E\{C_n^3\}\mu_{3,n-1} + 3[E\{C_n^3\} - E\{C_n^2\}\,\mu_{1C_n}]\,\mu_{2,n-1}\mu_{1,n-1}$$

$$+[E\{C_n^3\} - 3E\{C_n^2\}\mu_{1C_n} + 2\mu_{1C_n}^3]\,\mu_{1,n-1}^3$$

$$= \mu_{3C_n}\mu_{3,n-1} + 3\mu_{2C_n}\mu_{1C_n}\mu_{3,n-1} + \mu_{1C_n}^3\mu_{3,n-1}$$

$$+3\mu_{3C_n}\mu_{2,n-1}\mu_{1,n-1} + 6\mu_{2C_n}\mu_{1C_n}\mu_{2,n-1}\mu_{1,n-1} + \mu_{3C_n}\mu_{1,n-1}^3$$

$$= [\mu_{3C_n} + 3\mu_{2C_n}\mu_{1C_n} + \mu_{1C_n}^3]\,\mu_{3,n-1}$$

$$+ [3\mu_{3C_n} + 6\mu_{2C_n}\mu_{1C_n}]\,\mu_{2,n-1}\mu_{1,n-1} + \mu_{3C_n}\mu_{1,n-1}^3 \qquad (B.4.4)$$

Therefore the third central moment of X_n is

$$\mu_{3,n} = [\mu_{3C_n} + 3\mu_{2C_n}\mu_{1C_n} + \mu_{1C_n}^3]\,\mu_{3,n-1}$$

$$+ [3\mu_{3C_n} + 6\mu_{2C_n}\,\mu_{1C_n}]\,\mu_{2,n-1}\mu_{1,n-1} + \mu_{3C_n}\mu_{1,n-1}^3 + \mu_{3S_n}$$

$$(B.4.5)$$

APPENDIX C

UNIFORM, MINIMUM VARIANCE, MINIMUM RISK, UNBIASED ESTIMATORS

C.1 Introduction. In this appendix some useful theorems are presented which lead to the development of UMV-RUE's (uniform, minimum variance, minimum risk, unbiased estimators). Some discussion of the interpretation of these theorems and their application to the determination of UMV-RUE's is made. The procedure for the construction of UMV-RUE's is then presented in the form of examples and, finally, some useful relationships for the development of UMV-RUE's are presented.

The theorems and procedures of this appendix are essentially taken from Fraser (5) with modifications so that they agree with the content and notation of this thesis. The reader is referred to Fraser (5), Chapter 1 and 2, for a more comprehensive and theoretical presentation.

C.2 The Rao-Blackwell and Lehmann-Scheffé Theorems. Very fundamental to the development of UMV-RUE's are the Rao-Blackwell and Lehman-Scheffé Theorems. These two theorems are presented here in forms suitable to the purpose of this thesis.

Rao-Blackwell Theorem. If $t(\underline{x})$ is a sufficient statistic for the family of distribution functions indexed by a parameter vector, θ, $\{F_X(x;\underline{\theta}|\theta \in \Theta\}$, and $f(\underline{x})$ is an unbiased estimator of $g(\underline{\theta})$, then $h(t) = E\{f(X) | t(\underline{x})\}$ is an unbiased estimator based on $t(\underline{x})$. The variance of $h(t)$ is less than the variance of $f(\underline{x})$, $\sigma_f^2(\underline{\theta}) > \sigma_h^2(\underline{\theta})$, unless $f(\underline{x}) = h(t(\underline{x}))$ almost everywhere $(F_X(x;\underline{\theta}))$. With a strictly convex loss function, $R(\underline{\theta})$, the inequality $R_f(\underline{\theta}) > R_h(\underline{\theta})$ holds unless $f(\underline{x}) = h(t(\underline{x}))$ almost everywhere $(F_X(x;\underline{\theta}))$, in which case $R_f(\underline{\theta}) = R_h(\underline{\theta})$.

Lehmann-Scheffé Theorem. If there is a complete and sufficient statistic $t(\underline{x})$ for $\{F_X(x;\underline{\theta}) | \underline{\theta} \in \Theta\}$, then every estimable real parameter $g(\underline{\theta})$ has a unique unbiased estimator with minimum variance and minimum risk (strictly convex loss); the estimator is the only unbiased estimator which is a function of $t(\underline{x})$.

The Rao-Blackwell Theorem indicates that if there exists a sufficient statistic for the class of probability distribution functions, one of which is under consideration, and if an unbiased estimator of a parameter is known, then the conditional expectation of that estimator given the sufficient statistic is also an unbiased estimator of the parameter. Furthermore the conditional estimator has smaller variance and risk than the unconditional estimator. The Lehmann-Scheffé Theorem further indicates that, if the sufficient statistic is also complete, the conditional estimator is a unique unbiased estimator and has smaller variance and risk than any other unbiased estimator of the parameter.

C.3 UMV-RUE's of the Parameters of an Absolutely Continuous Distribution. Consider the k independent samples $\underline{x} = (x_1, x_2, \ldots, x_k)$ of a random variable X having the absolutely continuous distribution,

$F_X(x;\underline{\theta})$ on R^1, the real line. In Chapter II, Problem 14 of Fraser (5) it is shown that the order statistic, $t(x) = (x_{(1)},\ldots,x_{(k)})$, is a complete sufficient statistic for the class of absolutely continuous distributions on R^1. In the following examples UMV-RUE's of some of the parameters of $F_X(x;\underline{\theta})$ will be determined.

<u>Example 3.1 The UMV-RUE of μ_1, the Mean of X.</u> This example can be found in Fraser (5), pp. 58-59. Let $f(\underline{x}) = x_1$. Then, since $E\{f(\underline{X})\} = E\{X_1\} = \mu_1$, $f(\underline{x})$ is an unbiased estimator of μ_1. Therefore by the Rao-Blackwell Theorem $h(t) = E\{f(\underline{x})|t(\underline{x})\}$ is an unbiased estimator of μ_1, and by the Lehmann-Scheffé Theorem h(t) is the UMV-RUE of μ_1. So, h(t), the conditional expectation of x_1, must be determined.

The conditional probability, given the order statistic, assigns equal probability to each of the k! permutations of $(x_{(1)},\ldots,x_{(k)})$. Then if one is fixed, say $x_{(i)} = x_1$, there remain (k - 1)! permutations with $x_{(i)} = x_1$. Thus

$$P\{X_1 = x_{(i)} | t(\underline{x})\} = \frac{(k-1)!}{k!} = \frac{1}{k}, \quad i = 1,\ldots,k \qquad (C.3.1)$$

and

$$\mu_1^* = h(t) = E\{X_1|t\} = \sum_{i=1}^{k} x_{(i)} \cdot \frac{1}{k} = \frac{1}{k} \sum_{i=1}^{k} x_i = \bar{x} \qquad (C.3.2)$$

Therefore $\mu_1^* = \bar{x}$ is the UMV-RUE of μ_1.

In the body of this thesis UMV-RUE's of several moments and products of moments are used. The UMV-RUE's used are presented in Equation C.4.5. It would be somewhat redundant and serve no useful purpose to present the development of each of these UMV-RUE's. Example C.3.2, however, does present the development of a somewhat typical UMV-RUE.

Example C.3.2 The UMV-RUE of $\mu_4\mu_2$, the Product of the Fourth and Second (Variance) Central Moments of X. μ_4 can be expressed as

$$\mu_4 = E\{[X - \mu_1]^4\} = \alpha_4 - 4\alpha_3\alpha_1 + 6\alpha_2\alpha_1^2 - 3\alpha_1^4 \qquad (C.3.3)$$

where $\alpha_r = E\{X^r\}$ is the rth non-central moment of X.

Similarly

$$\mu_2 = E\{[X - \mu_1]^2\} = \alpha_2 - \alpha_1^2 \qquad (C.3.4)$$

Then

$$\mu_4\mu_2 = \alpha_4\alpha_2 - \alpha_4\alpha_1^2 - 4\alpha_3\alpha_2\alpha_1 + 4\alpha_3\alpha_1^3 + 6\alpha_2^2\alpha_1^2 - 9\alpha_2\alpha_1^4 + 3\alpha_1^6 \qquad (C.3.5)$$

Let

$$f(\underline{x}) = x_1^4x_2^2 - x_1^4x_2x_3 - 4x_1^3x_2^2x_3 + 4x_1^3x_2x_3x_4 + 6x_1^2x_2^2x_3x_4 - 9x_1^2x_2x_3x_4x_5 + 3x_1x_2x_3x_4x_5x_6 \qquad (C.3.6)$$

$$E\{f(\underline{X})\} = \alpha_4\alpha_2 - \alpha_4\alpha_1^2 - 4\alpha_3\alpha_2\alpha_1 + 4\alpha_3\alpha_1^3 + 6\alpha_2^2\alpha_1^2 - 9\alpha_2\alpha_1^4 + 3\alpha_1^6 = \mu_4\mu_2 \qquad (C.3.7)$$

Therefore $f(\underline{x})$ is an unbiased estimator of $\mu_4\mu_2$. The conditional expectation, given the order statistic, of $f(\underline{x})$ is the UMV-RUE of $\mu_4\mu_2$. Proceeding as in Example C.3.1 but fixing six elements of $t(\underline{x})$, say $x_{(e)} = x_1$, $x_{(f)} = x_2$, $x_{(g)} = x_3$, $x_{(h)} = x$, $x_{(i)} = x_5$, $x_{(j)} = x_6$. There remain $(k - 6)!$ permutations of $t(\underline{x})$. Thus

$$P\{X_1=x_{(e)}, X_2=x_{(f)}, X_3=x_{(g)}, X_4=x_{(h)}, X_5=x_{(i)}, X_6=x_{(j)} \mid t(\underline{x})\} = \frac{(k-6)!}{k!} = \frac{1}{k(k-1)(k-2)(k-3)(k-4)(k-5)}, \quad e,f,g,h,i,j = 1,\ldots,k, \; e\neq f\neq g\neq h\neq i\neq j \qquad (C.3.8)$$

and

$$(\mu_4\mu_2)^* = h(t) = E\{f(\underline{X})\,|\,t\}$$

$$= \frac{1}{k(k-1)(k-2)(k-3)(k-4)(k-5)} \sum_{e\neq f\neq g\neq h\neq i\neq j}^{k}\sum\sum\sum\sum\sum [x_e^4x_f^2 - x_e^4x_fx_g - 4x_e^3x_f^2x_g$$

$$+4x_e^3x_fx_gx_h + 6x_e^2x_f^2x_gx_h - 9x_e^2x_fx_gx_hx_i + 3x_ex_fx_gx_hx_ix_j] \qquad (C.3.9)$$

which will reduce to

$$(\mu_4\mu_2)^* = \frac{1}{(k-1)\ldots(k-5)}[-(k^4-4k^3+11k^2-8k)a_6 + k(k^4-9k^3+53k^2-135k+120)a_4a_2$$

$$+k(4k^3-28k^2+80k-80)a_3^2 - k^2(6k^2-27k+30)a_2^3 + k(6k^3-24k^2+66k-48)a_5a_1$$

$$-k^2(4k^3-12k^2+44k-60)a_3a_2a_1 - k^2(k^3+6k^2-7k+30)a_4a_1^2 + k^3(6k^2-15)a_2^2a_1^2$$

$$+k^3(4k^2+20)a_3a_1^3 - 9k^5a_2a_1^4 + 3k^5a_1^6] \qquad (C.3.10)$$

where $a_r = \frac{1}{k}\sum_{i=1}^{k} x_i^r$.

$(\mu_4\mu_2)^*$ can be reduced further to

$$(\mu_4\mu_2)^* = \frac{1}{(k-1)\ldots(k-5)}[-(k^4-4k^3+11k^2-8k)m_6 + k(k^4-9k^3+53k^2-135k+120)m_4m_2$$

$$+k(4k^3-28k^2+80k-80)m_3^2 - k^2(6k^2-27k+30)m_2^3] \qquad (C.3.11)$$

where $m_r = \frac{1}{k}\sum_{i=1}^{k}(x_i - \mu_1^*)$, $r = 2,3,\ldots$

Verification of the UMV-RUE's was performed by taking the expectation of the UMV-RUE's. This was accomplished by using Equations C.4.5 to express UMV-RUE's in terms of sample moments. Then Equations C.4.1, C.4.2, and C.4.3 were used to determine the expectations of the sample moments and thus the expectations of the UMV-RUE's.

C.4 Some Relationships Helpful in the Development of UMV-RUE's.

m_r in terms of μ_r'. In order to determine $E\{m_r\}$, the expectations of the sample moments, it is helpful to first express the sample moments,

$$m_1 = \frac{1}{k}\sum_{i=1}^{k} x_i, \quad m_r = \frac{1}{k}\sum_{i=1}^{k}(x_i - m_1)^r, \qquad r = 2,3,\ldots,$$

in terms of μ_r', where

$$\mu_r' = \frac{1}{k}\sum_{i=1}^{k}(x_i - \mu_1)^r, \qquad r = 1,2,\ldots$$

The sample moments in terms of μ_r', $r = 1,2,\ldots,$ are given by

$$m_1 = \mu_1' + \mu_1 \tag{C.4.1a}$$

$$\begin{bmatrix} m_2 \\ m_1^2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_2' \\ (\mu_1')^2 \end{bmatrix} + \mu_1 \begin{bmatrix} 0 \\ 2\mu_1' \end{bmatrix} + \mu_1^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{C.4.1b}$$

$$\begin{bmatrix} m_3 \\ m_2 m_1 \\ m_1^3 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu_3' \\ \mu_2'\mu_1' \\ (\mu_1')^3 \end{bmatrix} + \mu_1 \begin{bmatrix} 0 \\ m_2 \\ 3(\mu_1')^2 \end{bmatrix} + \mu_1^2 \begin{bmatrix} 0 \\ 0 \\ 3\mu_1' \end{bmatrix} + \mu_1^3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{C.4.1c}$$

$$\begin{bmatrix} m_4 \\ m_3 m_1 \\ m_2^2 \\ m_2 m_1^2 \end{bmatrix} = \begin{bmatrix} 1 & -4 & 0 & 6 & -3 \\ 0 & 1 & 0 & -3 & 2 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} \mu_4' \\ \mu_3'\mu_1' \\ (\mu_2')^2 \\ \mu_2'(\mu_1')^2 \\ (\mu_1')^4 \end{bmatrix} + \mu_1 \begin{bmatrix} 0 \\ m_3 \\ 0 \\ \mu_2'\mu_1' - (\mu_1')^3 + m_2 m_1 \end{bmatrix} \tag{C.4.1d}$$

$$\begin{bmatrix} m_5 \\ m_4 m_1 \\ m_3 m_2 \\ m_3 m_1^2 \\ m_2^2 m_1 \\ m_2 m_1^3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 10 & 0 & -10 & 4 \\ 0 & 1 & 0 & -4 & 0 & 6 & -3 \\ 0 & 0 & 1 & -1 & -3 & 5 & -2 \\ 0 & 0 & 0 & 1 & 0 & -3 & 2 \\ 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_5' \\ \mu_4'\mu_1' \\ \mu_3'\mu_2' \\ \mu_3'(\mu_1')^2 \\ (\mu_2')^2\mu_1' \\ \mu_2'(\mu_1')^3 \\ (\mu_1')^5 \end{bmatrix}$$

$$+ \mu_1 \begin{bmatrix} 0 \\ m_4 \\ 0 \\ \mu_3'\mu_1' - 3\mu_2'(\mu_1')^2 + 2(\mu_1')^4 + m_3 m_1 \\ m_2^2 \\ 2\mu_2'(\mu_1')^2 - 2(\mu_1')^4 + m_2 m_1^2 \end{bmatrix} + \mu_1^2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \mu_2'\mu_1' - (\mu_1')^3 \end{bmatrix} \qquad \text{(C.4.1e)}$$

With $\underline{H}$ given by

$$\underline{H} = \begin{bmatrix} 1 & -6 & 0 & 15 & 0 & 0 & -20 & 0 & 0 & 15 & -5 \\ 0 & 1 & 0 & -5 & 0 & 0 & 10 & 0 & 0 & -10 & 4 \\ 0 & 0 & 1 & -1 & 0 & -4 & 4 & 0 & 6 & -9 & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & -4 & 0 & 0 & 6 & -3 \\ 0 & 0 & 0 & 0 & 1 & -6 & 4 & 0 & 9 & -12 & 4 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & -3 & 5 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -3 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -3 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} m_6 \\ m_5 m_1 \\ m_4 m_2 \\ m_4 m_1^2 \\ m_3^2 \\ m_3 m_2 m_1 \\ m_3 m_1^3 \\ m_2^3 \\ m_2^2 m_1^2 \\ m_2 m_1^4 \end{bmatrix} = \underline{H} \begin{bmatrix} \mu_6' \\ \mu_5'\mu_1' \\ \mu_4'\mu_2' \\ \mu_4'(\mu_1')^2 \\ (\mu_3')^2 \\ \mu_3'\mu_2'\mu_1' \\ \mu_3'(\mu_1')^3 \\ (\mu_2')^3 \\ (\mu_2')^2(\mu_1')^2 \\ \mu_2'(\mu_1')^4 \\ (\mu_1')^6 \end{bmatrix} + \mu_1 \begin{bmatrix} 0 \\ m_5 \\ 0 \\ \mu_4'\mu_1' - 4\mu_3'(\mu_1')^2 + 6\mu_2'(\mu_1')^3 - 3(\mu_1')^5 + m_4 m_1 \\ 0 \\ m_3 m_2 \\ 2\mu_3'(\mu_1')^2 - 6\mu_2'(\mu_1')^3 + 4(\mu_1')^5 + m_3 m_1^2 \\ 0 \\ (\mu_2')^2\mu_1' - 2\mu_2'(\mu_1')^3 + (\mu_1')^5 + m_2^2 m_1 \\ 3\mu_2'(\mu_1')^3 - 3(\mu_1')^5 + m_2 m_1^3 \end{bmatrix}$$

$$+ \mu_1^2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \mu_3'\mu_1' - 3\mu_2'(\mu_1')^2 + 2(\mu_1')^4 \\ 0 \\ 0 \\ 3\mu_2'(\mu_1')^2 - 3(\mu_1')^4 \end{bmatrix} + \mu_1^3 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \mu_2'\mu_1' - (\mu_1')^3 \end{bmatrix} \qquad \text{(C.4.1f)}$$

$\underline{E\{\mu_r'\}}$. The expectations of μ_r' in terms of the moments μ_r are given by

$$E\{\begin{bmatrix} \mu_1' \\ \mu_2' \\ (\mu_1')^2 \\ \mu_3' \\ \mu_2'\mu_1' \\ (\mu_1')^3 \end{bmatrix}\} = \begin{bmatrix} 0 \\ \mu_2 \\ \frac{1}{k}\mu_2 \\ \mu_3 \\ \frac{1}{k}\mu_3 \\ \frac{1}{k^2}\mu_3 \end{bmatrix} \qquad \text{(C.4.2a)}$$

$$E\left\{\begin{bmatrix} \mu_4' \\ \mu_3'\mu_1' \\ (\mu_2')^2 \\ \mu_2'(\mu_1')^2 \\ (\mu_1')^4 \end{bmatrix}\right\} = \frac{1}{k^3}\begin{bmatrix} k^3 & 0 \\ k^2 & 0 \\ k^2 & k^2(k-1) \\ k & k(k-1) \\ 1 & 3(k-1) \end{bmatrix}\begin{bmatrix} \mu_4 \\ \mu_2^2 \end{bmatrix} \qquad (C.4.2b)$$

$$E\left\{\begin{bmatrix} \mu_5' \\ \mu_4'\mu_1' \\ \mu_3'\mu_2' \\ \mu_3'(\mu_1')^2 \\ (\mu_2')^2\mu_1' \\ \mu_2'(\mu_1')^3 \\ (\mu_1')^5 \end{bmatrix}\right\} = \frac{1}{k^4}\begin{bmatrix} k^4 & 0 \\ k^3 & 0 \\ k^3 & k^3(k-1) \\ k^2 & k^2(k-1) \\ k^2 & 2k^2(k-1) \\ k & 4k(k-1) \\ 1 & 10(k-1) \end{bmatrix}\begin{bmatrix} \mu_5 \\ \mu_3\mu_2 \end{bmatrix} \qquad (C.4.2c)$$

$$E\left\{\begin{bmatrix} \mu_6' \\ \mu_5'\mu_1' \\ \mu_4'\mu_2' \\ \mu_4'(\mu_1')^2 \\ (\mu_3')^2 \\ \mu_3'\mu_2'\mu_1' \\ \mu_3'(\mu_1')^3 \\ (\mu_2')^3 \\ (\mu_2')^2(\mu_1')^2 \\ \mu_2'(\mu_1')^4 \\ (\mu_1')^6 \end{bmatrix}\right\} = \frac{1}{k^5}\begin{bmatrix} k^5 & 0 & 0 & 0 \\ k^4 & 0 & 0 & 0 \\ k^4 & k^4(k-1) & 0 & 0 \\ k^3 & k^3(k-1) & 0 & 0 \\ k^4 & 0 & k^4(k-1) & 0 \\ k^3 & k^3(k-1) & k^3(k-1) & 0 \\ k^2 & 3k^2(k-1) & k^2(k-1) & 0 \\ k^3 & 3k^3(k-1) & 0 & k^3(k-1)(k-2) \\ k^2 & 3k^2(k-1) & 2k^2(k-1) & k^2(k-1)(k-2) \\ k & 7k(k-1) & 4k(k-1) & 3k(k-1)(k-2) \\ 1 & 15(k-1) & 10(k-1) & 15(k-1)(k-2) \end{bmatrix}\begin{bmatrix} \mu_6 \\ \mu_4\mu_2 \\ \mu_3^2 \\ \mu_2^3 \end{bmatrix} \qquad (C.4.2d)$$

The computations involved in the tedious task of determining Equations C.4.2 were eased somewhat by the use of several relations developed by Tchouproff (12). Note that $E\{\mu_1'\} = 0$ and $E\{\mu_r'\} = \mu_r$, $r = 2, 3, \ldots$.

<u>$E\{m_r\}$.</u> $E\{m_r\}$, the expectations of the sample moments, are determined by using $E\{\mu_r'\}$ of Equations C.4.2 in the expectations of m_r of Equations C.4.1. These results are

$$E\{m_1\} = \mu_1 \tag{C.4.3a}$$

$$E\left\{\begin{bmatrix} m_2 \\ m_1^2 \end{bmatrix}\right\} = \frac{1}{k}\begin{bmatrix} k-1 & 0 \\ 1 & k \end{bmatrix}\begin{bmatrix} \mu_2 \\ \mu_1^2 \end{bmatrix} \tag{C.4.3b}$$

$$E\left\{\begin{bmatrix} m_3 \\ m_2 m_1 \\ m_1^3 \end{bmatrix}\right\} = \frac{1}{k^2}\begin{bmatrix} (k-1)(k-2) & 0 & 0 \\ (k-1) & k(k-1) & 0 \\ 1 & 3k & k^2 \end{bmatrix}\begin{bmatrix} \mu_3 \\ \mu_2\mu_1 \\ \mu_1^3 \end{bmatrix} \tag{C.4.3c}$$

$$E\left\{\begin{bmatrix} m_4 \\ m_3 m_1 \\ m_2^2 \\ m_2 m_1^2 \end{bmatrix}\right\}$$

$$= \frac{1}{k^3}\begin{bmatrix} (k-1)(k^2-3k+3) & 0 & 3(k-1)(2k-3) & 0 \\ (k-1)(k-2) & k(k-1)(k-2) & -3(k-1)(k-2) & 0 \\ (k-1)^2 & 0 & (k-1)(k^2-2k+3) & 0 \\ (k-1) & 2k(k-1) & (k-1)(k-3) & k^2(k-1) \end{bmatrix}\begin{bmatrix} \mu_4 \\ \mu_3\mu_1 \\ \mu_2^2 \\ \mu_2\mu_1^2 \end{bmatrix} \tag{C.4.3d}$$

$$E\left\{\begin{bmatrix} m_5 \\ m_4 m_1 \\ m_3 m_2 \\ m_3 m_1^2 \\ m_2^2 m_1 \\ m_2 m_1^3 \end{bmatrix}\right\} = \frac{1}{k^4} K_{m\mu}^5 \begin{bmatrix} \mu_5 \\ \mu_4 \mu_1 \\ \mu_3 \mu_2 \\ \mu_3 \mu_1^2 \\ \mu_2^2 \mu_1 \\ \mu_2 \mu_1^3 \end{bmatrix} \qquad (C.4.3e)$$

$$E\left\{\begin{bmatrix} m_6 \\ m_5 m_1 \\ m_4 m_2 \\ m_4 m_1^2 \\ m_3^2 \\ m_3 m_2 m_1 \\ m_3 m_1^3 \\ m_2^3 \\ m_2^2 m_1^2 \\ m_2 m_1^4 \end{bmatrix}\right\} = \frac{1}{k^5} K_{m\mu}^6 \begin{bmatrix} \mu_6 \\ \mu_5 \mu_1 \\ \mu_4 \mu_2 \\ \mu_4 \mu_1^2 \\ \mu_3^2 \\ \mu_3 \mu_2 \mu_1 \\ \mu_3 \mu_1^3 \\ \mu_2^3 \\ \mu_2^2 \mu_1^2 \\ \mu_2 \mu_1^4 \end{bmatrix} \qquad (C.4.3f)$$

where $K_{m\mu}^5$ and $K_{m\mu}^6$ are given in Figure 15.

$\underline{\Psi^{*'}}$. The 21 distinct elements of $\Psi^{*'}$, the covariance matrix of $\underline{\mu}^*$, where $\underline{\mu}^* = [\mu_1^*, \mu_2^*, \mu_3^*, \mu_1^{2*}, \mu_1^{3*}, (\mu_2\mu_1)^*]^T$, are given by

$$\sigma^2_{\mu_1^*} = \frac{1}{k}\mu_2 \qquad (C.4.4a)$$

$$\begin{bmatrix} \sigma_{\mu_1^*\mu_2^*} \\ \sigma_{\mu_1^*\mu_1^{2*}} \end{bmatrix} = \frac{1}{k} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \mu_3 \\ \mu_2\mu_1 \end{bmatrix} \qquad (C.4.4b)$$

$$
K^5_{m\mu} = \begin{bmatrix}
(k-1)(k-2)(k^2-2k+2) & 0 & 10(k-1)(k-2)^2 & 0 & 0 & 0 \\
(k-1)(k^2-3k+3) & k(k-1)(k^2-3k+3) & -2(k-1)(2k^2-12k+15) & 0 & 3k(k-1)(2k-3) & 0 \\
(k-1)^2(k-2) & 0 & (k-1)(k-2)(k^2-5k+10) & 0 & 0 & 0 \\
(k-1)(k-2) & 2k(k-1)(k-2) & (k-1)(k-2)(k-10) & k^2(k-1)(k-2) & -6k(k-1)(k-2) & 0 \\
(k-1)^2 & k(k-1)^2 & 2(k-1)(k^2-4k+5) & 0 & k(k-1)(k^2-2k+3) & 0 \\
(k-1) & 3k(k-1) & 2(k-1)(2k-5) & 3k^2(k-1) & 3k(k-1)(k-3) & k^3(k-1)
\end{bmatrix}
$$

$$
K^6_{m\mu} = \begin{bmatrix}
(k-1)(k^4-5k^3+10k^2-10k+5) & 0 & 15(k-1)(k^3-4k^2+7k-5) & 0 & -10(k-1)(2k^2-6k+5) & 0 & 0 & 15(k-1)(k-2)(3k-5) & 0 & 0 \\
(k-1)(k-2)(k^2-2k+2) & k(k-1)(k-2)(k^2-2k+2) & -5(k-1)(k-2)(k^2-4k+6) & 0 & 10(k-1)(k-2)^2 & 10k(k-1)(k-2)^2 & 0 & -30(k-1)(k-2)^2 & 0 & 0 \\
(k-1)^2(k^2-3k+3) & 0 & (k-1)(k^4-5k^3+30k^2-63k+45) & 0 & -2(k-1)(2k^3-8k^2+18k-15) & 0 & 0 & 3(k-1)(k-2)(2k^2-9k+15) & 0 & 0 \\
(k-1)(k^2-3k+3) & 2k(k-1)(k^2-3k+3) & (k-1)(k-3)(k^2-9k+15) & k^2(k-1)(k^2-3k+3) & -2(k-1)(2k^2-12k+15) & -4k(k-1)(2k^2-12k+15) & 0 & 9(k-1)(k-2)(2k-5) & 3k^2(k-1)(2k-3) & 0 \\
(k-1)^2(k-2)^2 & 0 & -3(k-1)(k-2)^2(2k-5) & 0 & (k-1)(k-2)^2(k^2-2k+10) & 0 & 0 & 3(k-1)(k-2)(3k^2-12k+20) & 0 & 0 \\
(k-1)^2(k-2) & k(k-1)^2(k-2) & (k-1)(k-2)(k^2-10k+15) & 0 & (k-1)(k-2)(k^2-5k+10) & k(k-1)(k-2)(k^2-5k+10) & 0 & -3(k-1)(k-2)(k^2-5k+10) & 0 & 0 \\
(k-1)(k-2) & 3k(k-1)(k-2) & 3(k-1)(k-2)(k-5) & 3k^2(k-1)(k-2) & (k-1)(k-2)(k-10) & 3k(k-1)(k-2)(k-10) & k^3(k-1)(k-2) & -3(k-1)(k-2)(3k-10) & -9k^2(k-1)(k-2) & 0 \\
(k-1)^3 & 0 & 3(k-1)^2(k^2-2k+5) & 0 & -2(k-1)(3k^2-6k+5) & 0 & 0 & (k-1)(k-2)(k^3-3k^2+9k-15) & 0 & 0 \\
(k-1)^2 & 2k(k-1)^2 & (k-1)(k-3)(3k-5) & k^2(k-1)^2 & 2(k-1)(k^2-4k+5) & 4k(k-1)(k^2-4k+5) & 0 & (k-1)(k-2)(k^2-6k+15) & k^2(k-1)(k^2-2k+3) & 0 \\
(k-1) & 4k(k-1) & (k-1)(7k-15) & 6k^2(k-1) & 2(k-1)(2k-5) & 8k(k-1)(2k-5) & 4k^3(k-1) & 3(k-1)(k-2)(k-5) & 6k^2(k-1)(k-3) & k^4(k-1)
\end{bmatrix}
$$

Figure 15. $K^5_{m\mu}$ and $K^6_{m\mu}$

$$\begin{bmatrix} \sigma_{\mu_1^*\mu_3^*} \\ \sigma_{\mu_1^*\mu_1^{3*}} \\ \sigma_{\mu_1^*(\mu_2\mu_1)^*} \\ \sigma^2_{\mu_2^*} \\ \sigma_{\mu_2^*\mu_1^{2*}} \\ \sigma^2_{\mu_1^{2*}} \end{bmatrix} = \frac{1}{k(k-1)} \begin{bmatrix} k-1 & 0 & -3(k-1) & 0 \\ 0 & 0 & 0 & 3(k-1) \\ 0 & k-1 & k-1 & 0 \\ k-1 & 0 & -(k-3) & 0 \\ 0 & 2(k-1) & -2 & 0 \\ 0 & 0 & 2 & 4(k-1) \end{bmatrix} \begin{bmatrix} \mu_4 \\ \mu_3\mu_1 \\ \mu_2^2 \\ \mu_2\mu_1^2 \end{bmatrix} \qquad (C.4.4c)$$

$$\begin{bmatrix} \sigma_{\mu_2^*\mu_3^*} \\ \sigma_{\mu_2^*\mu_1^{3*}} \\ \sigma_{\mu_2^*(\mu_2\mu_1)^*} \\ \sigma_{\mu_3^*\mu_1^{2*}} \\ \sigma_{\mu_1^{2*}\mu_1^{3*}} \\ \sigma_{\mu_1^{2*}(\mu_2\mu_1)^*} \end{bmatrix} = \frac{1}{k(k-1)} K^5_{\sigma\mu} \begin{bmatrix} \mu_5 \\ \mu_4\mu_1 \\ \mu_3\mu_2 \\ \mu_3\mu_1^2 \\ \mu_2^2\mu_1 \\ \mu_2\mu_1^3 \end{bmatrix} \qquad (C.4.4d)$$

$$\begin{bmatrix} \sigma^2_{\mu_3^*} \\ \sigma_{\mu_3^*\mu_1^{3*}} \\ \sigma_{\mu_3^*(\mu_2\mu_1)^*} \\ \sigma^2_{\mu_1^{3*}} \\ \sigma_{\mu_1^{3*}(\mu_2\mu_1)^*} \\ \sigma^2_{(\mu_2\mu_1)^*} \end{bmatrix} = \frac{1}{k(k-1)(k-2)} K^6_{\sigma\mu} \begin{bmatrix} \mu_6 \\ \mu_5\mu_1 \\ \mu_4\mu_2 \\ \mu_4\mu_1^2 \\ \mu_3^2 \\ \mu_3\mu_2\mu_1 \\ \mu_3\mu_1^3 \\ \mu_2^3 \\ \mu_2^2\mu_1^2 \\ \mu_2\mu_1^4 \end{bmatrix} \qquad \text{(C.4.4e)}$$

where $K^5_{\sigma\mu}$ and $K^6_{\sigma\mu}$ are given in Figure 16.

<u>UMV-RUE's of Moments and Products of Moments.</u> The UMV-RUE's of $\underline{\mu}$, where $\underline{\mu} = [\mu_1, \mu_2, \mu_3, \mu_1^2, \mu_1^3, \mu_2\mu_1]^T$, and the UMV-RUE's required to determine $\hat{\Psi}^*$, the UMV-RUE of Ψ^*, are given in terms of the sample moments by

$$\mu_1^* = m_1 \qquad \text{(C.4.5a)}$$

$$\begin{bmatrix} \mu_2^* \\ \mu_1^{2*} \end{bmatrix} = \frac{1}{k-1} \begin{bmatrix} k & 0 \\ -1 & k-1 \end{bmatrix} \begin{bmatrix} m_2 \\ m_1^2 \end{bmatrix} \qquad \text{(C.4.5b)}$$

$$K_{\sigma\mu}^{5} = \begin{bmatrix} (k-1) & 0 & -2(2k-5) & 0 & 0 & 0 \\ 0 & 0 & 0 & 3(k-1) & -6 & 0 \\ 0 & (k-1) & (k-3) & 0 & -(k-3) & 0 \\ 0 & 2(k-1) & -6 & 0 & -6(k-1) & 0 \\ 0 & 0 & 0 & 0 & 6 & 6(k-1) \\ 0 & 0 & 2 & 2(k-1) & 2(k-2) & 0 \end{bmatrix}$$

$$K_{\sigma\mu}^{6} = \begin{bmatrix} (k-1)(k-2) & 0 & -3(k-2)(2k-5) & 0 & -(k-2)(k-10) & 0 & 0 & 3(3k^2-12k+20) & 0 & 0 \\ 0 & 0 & 0 & 3(k-1)(k-2) & 0 & -18(k-2) & 0 & 12 & -9(k-1)(k-2) & 0 \\ 0 & (k-1)(k-2) & (k-2)(k-4) & 0 & -3(k-2) & -(k-2)(4k-10) & 0 & -3(k^2-4k+8) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 18(k-2) & 9(k-1)(k-2) \\ 0 & 0 & 0 & 0 & 0 & 6(k-2) & 3(k-1)(k-2) & -6 & 3(k-2)(k-3) & 0 \\ 0 & 0 & (k-2) & (k-1)(k-2) & (k-2) & 2(k-2)(k-3) & 0 & (k^2-4k+10) & -(k-2)(k-3) & 0 \end{bmatrix}$$

Figure 16. $K_{\sigma\mu}^{5}$ **and** $K_{\sigma\mu}^{6}$

$$\begin{bmatrix} \mu_3^* \\ (\mu_2\mu_1)^* \\ \mu_1^{3*} \end{bmatrix} = \frac{1}{(k-1)(k-2)} \begin{bmatrix} k^2 & 0 & 0 \\ -k & k(k-2) & 0 \\ 2 & -3(k-2) & (k-1)(k-2) \end{bmatrix} \begin{bmatrix} m_3 \\ m_2 m_1 \\ m_1^3 \end{bmatrix} \quad \text{(C.4.5c)}$$

$$\begin{bmatrix} \mu_4^* \\ (\mu_3\mu_1)^* \\ \mu_2^{2*} \\ (\mu_2\mu_1^2)^* \end{bmatrix} = \frac{k}{(k-1)(k-2)(k-3)} K_{\mu m}^4 \begin{bmatrix} m_4 \\ m_3 m_1 \\ m_2^2 \\ m_2 m_1^2 \end{bmatrix} \quad \text{(C.4.5d)}$$

$$\begin{bmatrix} \mu_5^* \\ (\mu_4\mu_1)^* \\ (\mu_3\mu_2)^* \\ (\mu_3\mu_1^2)^* \\ (\mu_2^2\mu_1)^* \\ (\mu_2\mu_1^3)^* \end{bmatrix} = \frac{k}{(k-1)(k-2)(k-3)(k-4)} K_{\mu m}^5 \begin{bmatrix} m_5 \\ m_4 m_1 \\ m_3 m_2 \\ m_3 m_1^2 \\ m_2^2 m_1 \\ m_2 m_1^3 \end{bmatrix} \quad \text{(C.4.5e)}$$

$$\begin{bmatrix} \mu_6^* \\ (\mu_5\mu_1)^* \\ (\mu_4\mu_2)^* \\ (\mu_4\mu_1^2)^* \\ \mu_3^{2*} \\ (\mu_3\mu_2\mu_1)^* \\ (\mu_3\mu_1^3)^* \\ \mu_2^{3*} \\ (\mu_2^2\mu_1^2)^* \\ (\mu_2\mu_1^4)^* \end{bmatrix} = \frac{k}{(k-1)(k-2)(k-3)(k-4)(k-5)} K_{\mu m}^6 \begin{bmatrix} m_6 \\ m_5 m_1 \\ m_4 m_2 \\ m_4 m_1^2 \\ m_3^2 \\ m_3 m_2 m_1 \\ m_3 m_1^3 \\ m_2^3 \\ m_2^2 m_1^2 \\ m_2 m_1^4 \end{bmatrix} \quad \text{(C.4.5f)}$$

where $K_{\mu m}^4$, $K_{\mu m}^5$, and $K_{\mu m}^6$ are given in Figure 17.

$$K^4_{\mu m} = \begin{bmatrix} k^2-2k+3 & 0 & -3(2k-3) & 0 \\ -(k+1) & k(k-3) & 3(k-1) & 0 \\ -(k-1) & 0 & k^2-3k+3 & 0 \\ 2 & -2(k-3) & -k & (k-2)(k-3) \end{bmatrix}$$

$$K^5_{\mu m} = \begin{bmatrix} k(k^2-5k+10) & 0 & -10k(k-2) & 0 & 0 & 0 \\ -(k^2-k+6) & (k-4)(k^2-2k+3) & 2(2k^2+k-6) & 0 & -3(2k-3)(k-4) & 0 \\ -k(k-1) & 0 & k(k^2-2k+2) & 0 & 0 & 0 \\ 2(k+2) & -2(k+1)(k-4) & -(k^2+8k-8) & k(k-3)(k-4) & 6(k-1)(k-4) & 0 \\ 2(k-1) & -(k-1)(k-4) & -2(k^2-2k+2) & 0 & (k-4)(k^2-3k+3) & 0 \\ -6 & 6(k-4) & 5k & -3(k-3)(k-4) & -3k(k-4) & (k-2)(k-3)(k-4) \end{bmatrix}$$

$$K^6_{\mu m} = \begin{bmatrix} k^4-9k^3+31k^2-35k+40 & 0 & -15(k^3-8k^2+29k-40) & 0 & -40(k^2-6k+10) & 0 & 0 & 15k(3k-10) & 0 & 0 \\ -(k^3-4k^2+11k+16) & k(k-5)(k^2-5k+10) & 5(k^3-6k^2+29k-48) & 0 & 20(k^2-5k+8) & -10k(k-2)(k-5) & 0 & -30k(k-3) & 0 & 0 \\ -(k-1)(k^2-3k+8) & 0 & k^4-9k^3+53k^2-135k+120 & 0 & 4(k-2)(k^2-5k+10) & 0 & 0 & -3k(k-2)(2k-5) & 0 & 0 \\ 2(k^2+11) & -2(k^3-6k^2+11k-30) & -(k^3+47k-90) & (k-4)(k-5)(k^2-2k+3) & -8(k^2-3k+5) & 4(k-5)(k+2)(2k-3) & 0 & 9k(2k-5) & -3(k-4)(k-5)(2k-3) & 0 \\ -(k-1)(k^2-k+4) & 0 & 3(2k^3-5k^2-5k+20) & 0 & k^4-8k^3+25k^2-10k-40 & 0 & 0 & -3k(3k^2-15k+20) & 0 & 0 \\ 2(k+1)(k-1) & -k(k-1)(k-5) & -(k^3+6k^2-25k+30) & 0 & -(k^3-4k^2+15k-20) & k(k-5)(k^2-2k+2) & 0 & 3k(k^2-4k+5) & 0 & 0 \\ -6(k+3) & 6(k-5)(k+2) & 3(k^2+9k-10) & -3(k-5)(k-4)(k+1) & 2k(k+3) & -3(k-5)(k^2+8k-8) & k(k-5)(k-4)(k-3) & -3k(3k-5) & 9(k-5)(k-4)(k-1) & 0 \\ 2(k-1)(k-2) & 0 & -3(k-2)(k^2-5k+10) & 0 & -2(3k^2-15k+20) & 0 & 0 & k(k-2)(k^2-7k+15) & 0 & 0 \\ -6(k-1) & 4(k-5)(k-1) & 5k^2-9k+10 & -(k-5)(k-4)(k-1) & 2k(k-1) & -4(k-5)(k^2-2k+2) & 0 & -k(k^2-3k+5) & (k-5)(k-4)(k^2-3k+3) & 0 \\ 24 & -24(k-5) & -18k & 12(k-5)(k-4) & -3k & 20k(k-5) & -4(k-5)(k-4)(k-3) & 3k^2 & -6k(k-5)(k-4) & (k-5)(k-4)(k-3)(k-2) \end{bmatrix}$$

Figure 17. $K^4_{\mu m}$, $K^5_{\mu m}$, and $K^6_{\mu m}$

APPENDIX D

BAYESIAN ESTIMATION

D.1 Introduction. This appendix is concerned with the development of a procedure whereby the prediction estimate, $\underline{\mu}_n'$ and the observation or data estimate, $\underline{\mu}_n^*$, are combined to produce the estimate, $\hat{\underline{\mu}}_n$. This development is based on the use of Bayes' Rule in what is commonly called Bayes' learning.

D.2 Bayes' Learning. Let $\underline{\theta}$ be a vector valued random variable (an unknown parameter set modeled as a vector valued random variable) and $\underline{Y}$ a vector valued random variable statistically related to θ. The a posteriori density function of $\underline{\theta}$ given $\underline{Y}$ according to Bayes' Rule is given by

$$f_{\underline{\theta}|\underline{Y}} = \frac{f_{\underline{Y}|\underline{\theta}}\, f_{\underline{\theta}}}{f_{\underline{Y}}} \qquad (D.2.1)$$

where $f_{\underline{\theta}}$ is the a priori density function of $\underline{\theta}$, $f_{\underline{Y}|\underline{\theta}}$ is the conditional density function of $\underline{Y}$ given $\underline{\theta}$ and f_Y is given by

$$f_{\underline{Y}} = \int f_{Y|\underline{\theta}}\, d\, F_{\underline{\theta}}$$

An iterative, or recursive, approach to the computation of an a posteriori density function can also use Bayes' Rule. Let $\underline{\theta}_n$ be a vector valued random variable and $\underline{Y}_0, \underline{Y}_1, \ldots, \underline{Y}_n$ vector valued random variables statistically related to $\underline{\theta}_n$. The a posteriori density function of $\underline{\theta}_n$ given $\underline{Y}_0, \underline{Y}_1, \ldots, \underline{Y}_n$, where $\underline{Y}_0, \underline{Y}_1, \ldots, \underline{Y}_n$ are conditional independent given $\underline{\theta}_n$, is

$$f_{\underline{\theta}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_n} = \frac{f_{\underline{Y}_n|\underline{\theta}_n} \, f_{\underline{\theta}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}}}{f_{\underline{Y}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_n}} \qquad (D.2.2)$$

where $f_{\underline{\theta}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}}$ is the a priori density function of $\underline{\theta}_n$ given $\underline{Y}_0, \underline{Y}_1, \ldots, \underline{Y}_{n-1}$, $f_{\underline{Y}_n|\underline{\theta}_n}$ is the conditional density of $\underline{Y}_n$ given $\underline{\theta}_n$ and is referred to as the liklihood of $\underline{Y}_n$, and $f_{\underline{Y}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}}$ is given by

$$f_{\underline{Y}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}} = \int f_{\underline{Y}_n|\underline{\theta}_n} \, dF_{\underline{\theta}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}}$$

If in Equation D.2.1 $f_{\underline{\theta}|\underline{Y}_n}$ is of the same family of density functions as $f_{\underline{\theta}}$ and in Equation D.2.2 $f_{\underline{\theta}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_n}$ is of the same family of density functions as $f_{\underline{\theta}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}}$, then $f_{\underline{\theta}}$ and $f_{\underline{\theta}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}}$ are said to be reproducing a priori density functions (11).

When Bayes' Rule is used as in Equations D.2.1 and D.2.2 to estimate or learn the parameter set $\underline{\theta}$ or $\underline{\theta}_n$, respectively, and the a priori density functions are reproducing densities, the estimation or learning process is called Bayes' learning.

Ideally the use of Bayes' learning to estimate $\underline{\mu}_n$ would be to determine the density of $\underline{\mu}_n$ given $\underline{\mu}_0', \underline{X}_0, \underline{X}_1, \ldots, \underline{X}_n$ by

$$f_{\underline{\mu}_n|\underline{\mu}_0',\underline{X}_0,\underline{X}_1,\ldots,\underline{X}_n} = \frac{f_{\underline{X}_n|\underline{\mu}_n} \, f_{\underline{\mu}_n|\underline{\mu}_0',\underline{X}_0,\underline{X}_1,\ldots,\underline{X}_{n-1}}}{f_{\underline{X}_n|\underline{\mu}_0',\underline{X}_0,\underline{X}_1,\ldots,\underline{X}_{n-1}}} \qquad (D.2.3)$$

where $\underline{X}_n = \{X_{i,n}\}$, $i = 1, \ldots, k$, $n = 0, 1, \ldots,$ is a vector valued random variable for each n representing the k observations of the random variable X_n and $\underline{\mu}_0'$ is the initial estimate of $\underline{\mu}_0$. The difficulty in using Equation D.2.3 is that $f_{\underline{X}_n|\underline{\mu}_n}$ is unknown. Since the k observations of X_n are considered to be independent, $f_{\underline{X}_n|\underline{\mu}_n} = \prod_{i=1}^{k} f_{X_{i,n}|\underline{\mu}_n}$. Then, since $f_{X_{i,n}|\underline{\mu}_n}$ is the unknown density function which is to be approximated with estimates of its moments (See Section 2.2 and 3.1), $f_{\underline{X}_n|\underline{\mu}_n}$ is also unknown.

Instead, the approach here is to assume that $\underline{\mu}_n^*$ is a normally (Gaussian) distributed random vector and to use Bayes' learning to estimate the parameters of its Gaussian distribution. From these estimates an estimate of $\underline{\mu}_n$ is formed.

By making the assumption that $\underline{\mu}_n^*$ is a normally distributed random vector some obvious contradictions are overlooked. It is highly unlikely that in any particular case the elements of $\underline{\mu}_n^*$ will ever be jointly normally distributed. Certainly this is not generally true. For instance, consider the case where X_n is normally distributed. $\underline{\mu}_n^*$ is formed from the k samples of X_n. The estimate μ_{1n}^*, the sample mean, is normally distributed but the estimate μ_{2n}^*, the unbiased sample variance, is chi-square distributed; so that $\underline{\mu}_n^*$ can not be a normally distributed random vector. However the assumption here is that the normal distribution will yield a good approximation to the density of $\underline{\mu}_n^*$.

D.3 Gaussian-Wishart: A Reproducing Density Function. If the likelihood, $f_{\underline{Y}|\underline{\theta}}$, of Equation D.2.1 is the Gaussian density function with $\underline{\theta}$ the unknown parameter set composed of the mean vector, $\underline{M}$, and the inverted covariance matrix, $\underline{P}$, i.e., $\underline{Y} \sim N(\underline{M},\underline{P}^{-1})$, then Keehn (8) has shown that the reproducing a priori density function f_{θ}, for

$\underline{\theta} = (\underline{M},\underline{P})$ is the composite Gaussian-Wishart density function, G.W.$(w,v,\underline{R},\underline{Q})$.

If $\underline{Y}$ is a r-dimensional vector which is normally distributed then

$$f_{\underline{Y}|\underline{M},\underline{P}}(\underline{y}|\underline{M},\underline{P}) = N(\underline{M},\underline{P}^{-1})$$
$$= (2\pi)^{-\frac{r}{2}} |\underline{P}|^{\frac{1}{2}} \exp[-\frac{1}{2} (\underline{y}-\underline{M})^T \underline{P}(\underline{y}-\underline{M})] \tag{D.3.1}$$

where $\underline{M}$ is the r-dimensional mean vector and $\underline{P}$ is the (r x r) inverted covariance matrix. The composite Gaussian-Wishart density function on $(\underline{M},\underline{P})$ is

$$f_{\underline{M},\underline{P}}(\underline{m},\underline{p}) = G.W.(w',v',\underline{R}',\underline{Q}')$$
$$= (2\pi)^{-\frac{r}{2}} \left|w'\underline{p}\right|^{\frac{1}{2}} \exp[-\frac{1}{2} (\underline{m}-\underline{R}')^T w'\underline{p}(\underline{m}-\underline{R}')]$$
$$\cdot\ C_{r,v'} \left|\frac{v'}{2}\underline{Q}'\right|^{\frac{v'-1}{2}} \left|\underline{p}\right|^{\frac{v'-r-2}{2}} \exp[-\frac{1}{2}\,\mathrm{tr}\ v'\underline{Q}'\underline{p}] \tag{D.3.2}$$

where $\underline{R}'$ is a r-dimensional vector, $\underline{Q}'$ is a (r x r) positive definite matrix, w' and v' are real numbers associated with $\underline{R}'$ and $\underline{Q}'$, respectively, such that $w' > 0$ and $v' > r + 2$, $C_{k,v'}$ is given by

$$C_{r,v'} = \frac{1}{\pi^{\frac{r(r-1)}{4}} \prod_{a=1}^{r} \Gamma(\frac{v'-a}{2})}$$

and "tr ·" represents the trace of " · ".

The Gaussian-Wishart density implies that the random covariance matrix $\underline{P}^{-1}$ is distributed according to the inverted Wishart law with parameters v' and $\underline{Q}'$ where $\underline{Q}'$ is a covariance matrix and v' is a confidence factor which measures how concentrated the inverted Wishart law is about $\underline{Q}$. The concentration is greater when v' is larger. The random mean vector $\underline{M}$ is then distributed according to the Gaussian law with mean $\underline{R}'$ and covariance matrix $\frac{1}{w'}\,\underline{P}^{-1}$ where w' is a confidence factor which measures how concentrated the Gaussian law is about $\underline{R}'$. The concentration is greater when w' is larger. w' and v' can be thought of as constants reflecting the confidence that $\underline{R}'$ and $\underline{Q}'$ are the true mean vector and covariance matrix, respectively, of the Gaussian distributed random vector $\underline{Y}$ (8).

Since the Gaussian-Wishart density function is a reproducing a priori density with respect to the Gaussian density function with unknwon mean vector and covariance matrix, the a posteriori density function is also a Gaussian-Wishart density function. If the a priori density function is given by Equation D.3.2 then the a posteriori density function is of the same form as Equation D.3.2 with different parameters. Thus $f_{\underline{M},\underline{P}|\underline{Y}}$ is given by

$$f_{\underline{M},\underline{P}|\underline{Y}}(\underline{m},\underline{p}|\underline{y}) = G.W.(w,v,\underline{R},\underline{Q}) \qquad (D.3.3)$$

where, from Keehn (8),

$$w = w' + 1 \;, \qquad v = v' + 1,$$

$$\underline{R} = \frac{w'\underline{R}' + \underline{y}}{w'+1} \;, \qquad (D.3.4)$$

$$Q = \frac{1}{v'+1}\left[v'\underline{Q}' + w'\underline{R}'\underline{R}'^T + \underline{y}\,\underline{y}^T - w\,\underline{R}\,\underline{R}^T\right] \;,$$

and $\underline{y}$ is the observation of the random vector $\underline{Y}$.

In the iterative form of Bayes' Rule, Equation D.2.2, if the likelihood, $f_{\underline{Y}_n|\underline{\theta}_n}$, is Gaussian, $\underline{Y}_n \sim N(\underline{M}_n,\underline{P}_n^{-1})$, then the reproducing a priori density of $\underline{\theta}_n$ given $\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}$ is

$$f_{\underline{M}_n,\underline{P}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}}(\underline{m},\underline{p}|\underline{y}_0,\underline{y}_1,\ldots,\underline{y}_{n-1})$$

$$= G.W.(w_n', v_n', \underline{R}_n', \underline{Q}_n') \tag{D.3.5}$$

and the a posteriori density of $\underline{\theta}_n$ given $\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_{n-1}$ is

$$f_{\underline{M}_n,\underline{P}_n|\underline{Y}_0,\underline{Y}_1,\ldots,\underline{Y}_n}(\underline{m},\underline{p}|\underline{y}_0,\underline{y}_1,\ldots,\underline{y}_n) = G.W.(w_n, v_n, \underline{R}_n, \underline{Q}_n) \tag{D.3.6}$$

where

$$w_n = w_n' + 1 \quad , \quad v_n = v_n' + 1 \quad ,$$

$$\underline{R}_n = \frac{w_n'\underline{R}_n' + \underline{y}_n}{w_n' + 1} \quad ,$$

$$\underline{Q}_n = \frac{1}{v_n' + 1}\left[v_n'\underline{Q}_n' + w_n'\underline{R}_n'\underline{R}_n'^T + \underline{y}_n\underline{y}_n^T - w_n\underline{R}_n\underline{R}_n^T\right] \quad ,$$

and $\underline{y}_n$ is the observation of $\underline{Y}_n$.

<u>D.4 Learning the Augmented Moment Vector, $\underline{\mu}_n$.</u> As indicated in Section D.2 the approach here is to assume that $\underline{\mu}_n^*$ is a normally distributed random vector and to use Bayes' learning to estimate the parameters of the normal distribution and then form an estimate of $\underline{\mu}_n$.

Assuming that $\underline{\mu}_n^* \sim N(\underline{M}_n, \underline{P}_n^{-1})$ where, recalling that $\underline{\mu}_n^*$ is a 6-dimensional vector, $\underline{M}_n$ is the 6-dimensional mean vector and $\underline{P}_n^{-1}$ is the (6 x 6) inverted covariance matrix of $\underline{\mu}_n^*$, Bayes' Rule for the density function of $(\underline{M}_n, \underline{P}_n)$ is

$$f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_n} = \frac{f_{\underline{\mu}_n^*|\underline{M}_n,\underline{P}_n} \; f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_{n-1}}}{f_{\underline{\mu}_n^*|\mathcal{E}_{n-1}}} \tag{D.4.1}$$

where $\mathcal{E}_n = (\underline{\mu}_0', \underline{\mu}_0^*, \underline{\mu}_1^*, \ldots, \underline{\mu}_n^*)$ with $\underline{\mu}_0'$ the initial estimate of $\underline{\mu}_0$ and

$\underline{\mu}_i^*$ the observation of $\underline{\mu}_i$, $i = 0,1,\ldots,n$. Then the reproducing a priori density of $(\underline{M}_n, \underline{P}_n^{-1})$ is

$$f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_{n-1}}(\underline{m},\underline{p}) = \text{G.W.}(w_n', v_n', \underline{R}_n', \underline{Q}_n') \qquad \text{(D.4.2)}$$

and the a posteriori density of $(\underline{M}_n, \underline{P}_n^{-1})$ is

$$f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_n}(\underline{m},\underline{p}) = \text{G.W.}(w_n, v_n, \underline{R}_n, \underline{Q}_n) \qquad \text{(D.4.3)}$$

where, as in Equations D.3.4,

$$w_n = w_n' + 1 \quad , \quad v_n = v_n' + 1$$

$$\underline{R}_n = \frac{w_n' \underline{R}_n' + \underline{\mu}_n^*}{w_n' + 1} \quad , \qquad \text{(D.4.4)}$$

$$\underline{Q}_n = \frac{1}{v_n' + 1}\left[v_n'\underline{Q}_n' + w_n'\underline{R}_n'\underline{R}_n'^T + \underline{\mu}_n^* \underline{\mu}_n^{*T} - w_n\underline{R}_n\underline{R}_n^T\right]$$

Recall that the form of the Gaussian-Wishart (G.W.) density function is given in Equation D.3.2.

From the discussion of Section D.3 $\underline{R}_n$ and $\underline{Q}_n$ are estimates of the mean vector and covariance matrix of $\underline{\mu}_n^*$ given $\mathcal{E}_n$ with w_n and v_n reflecting the confidence in $\underline{R}_n$ and $\underline{Q}_n$, respectively. However the objective is to determine an estimate of $\underline{\mu}_n$ not estimates of the mean vector and covariance matrix of $\underline{\mu}_n^*$. The desired estimate is taken to be $\hat{\underline{\mu}}_n$ such that $\hat{\underline{\mu}}_n = g(\underline{\mu}_0', \underline{\mu}_0^*, \underline{\mu}_1^*, \ldots, \underline{\mu}_n^*)$ and the mean squared error between $\hat{\underline{\mu}}_n$ and $\underline{\mu}_n$ is minimized. The estimate which minimizes the mean squared error is the conditional expectation.

Therefore $\hat{\underline{\mu}}_n = E\{\underline{\mu}_n \mid \mathcal{E}_n\}$.

In Chapter III $\underline{\mu}_n^*$ is developed as an unbiased estimate of $\underline{\mu}_n$, i.e., $E\{\underline{\mu}_n^* | \underline{\mu}_{na}\} = \underline{\mu}_n$. Then

$$E\{\underline{\mu}_n^*\} = E\{E\{\underline{\mu}_n^* \mid \underline{\mu}_{na}\}\} = E\{\underline{\mu}_n\}$$

Now to show that $E\{\underline{\mu}_n^* \mid \mathcal{E}_n\} = E\{\underline{\mu}_n \mid \mathcal{E}_n\}$

$$E\{\underline{\mu}_n^* \mid \mathcal{E}_n\} = E\{\underline{\mu}_n^* \mid \underline{\mu}_0', \underline{\mu}^*(0), \underline{\mu}^*(1), \ldots, \underline{\mu}^*(n)\}$$

where $\underline{\mu}^*(n)$ is the observation of $\underline{\mu}_n$, i.e., the random variable $\underline{\mu}_n^*$ is observed to have the value $\underline{\mu}^*(n)$. Then using the properties of conditional expectation

$E\{\underline{\mu}_n^* \mid \mathcal{E}_n\}$

$$= E\{E\{\underline{\mu}_n^* \mid \underline{\mu}_{na} \underline{\mu}_0', \mu^*(0), \underline{\mu}^*(1), \ldots, \underline{\mu}^*(n)\} \mid \underline{\mu}_0', \underline{\mu}^*(0), \underline{\mu}^*(1), \ldots, \underline{\mu}^*(n)\}$$

and since $\underline{\mu}_n^*$ given $\underline{\mu}_{na}$ is conditionally independent of the initial estimate $\underline{\mu}_0'$ and the observations $\underline{\mu}^*(0), \ldots, \underline{\mu}^*(n)$

$$E\{\underline{\mu}_n^* \mid \mathcal{E}_n\} = E\{E\{\underline{\mu}_n^* \mid \underline{\mu}_{na}\} \mid \underline{\mu}_0, \underline{\mu}^*(0), \ldots, \underline{\mu}^*(n)\}$$

$$= E\{\underline{\mu}_n \mid \underline{\mu}_0', \underline{\mu}^*(0), \ldots, \underline{\mu}^*(n)\}$$

$$= E\{\underline{\mu}_n \mid \mathcal{E}_n\}$$

Therefore

$$\hat{\underline{\mu}}_n = E\{\underline{\mu}_n \mid \mathcal{E}_n\} = E\{\underline{\mu}_n^* \mid \mathcal{E}_n\}$$

$$= \int_{\underline{\mu}} \underline{\mu} \; f_{\underline{\mu}_n^* \mid \mathcal{E}_n}(\underline{\mu}) \; d\underline{\mu} \qquad \text{(D.4.5)}$$

Equation D.4.5 indicates that $\hat{\underline{\mu}}_n$ is the mean vector of the conditional density function, $f_{\underline{\mu}_n^* \mid \mathcal{E}_n}$. The conditional density function, $f_{\underline{\mu}_n^* \mid \mathcal{E}_n}$, is called the post-sampling density of $\underline{\mu}_n^*$ and is determined from

$$f_{\underline{\mu}_n^*|\mathcal{E}_n} = \int_{\underline{m}}\int_{\underline{p}} f_{\underline{\mu}_n^*|\underline{M}_n,\underline{P}_n,\mathcal{E}_n}\, f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_n}\, d\underline{m}\, d\underline{p} \qquad (D.4.6)$$

where $f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_n}$ is the a posteriori density of $(\underline{M}_n,\underline{P}_n)$ given by Equation D.4.3 and $f_{\underline{\mu}_n^*|\underline{M}_n,\underline{P}_n,\mathcal{E}_n}$ is the conditional density of $\underline{\mu}_n^*$ given its mean vector, $\underline{M}_n$, and inverted covariance matrix, $\underline{P}_n$ and $\mathcal{E}_n$. Since $\underline{\mu}_n^*$ is assumed to be normally distributed given $\underline{M}_n$ and $\underline{P}_n$

$$f_{\underline{\mu}_n^*|\underline{M}_n,\underline{P}_n,\mathcal{E}_n} = N(\underline{M}_n,\underline{P}_n^{-1}) = f_{\underline{\mu}_n^*|\underline{M}_n,\underline{P}_n}$$

to that Equation D.4.6 becomes

$$f_{\underline{\mu}_n^*|\mathcal{E}_n} = \int_{\underline{m}}\int_{\underline{p}} f_{\underline{\mu}_n^*|\underline{M}_n,\underline{P}_n}\, f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_n}\, d\underline{m}\, d\underline{p} \qquad (D.4.7)$$

The integration in Equation D.4.7 can be performed using the properties of the Gaussian-Wishart density as presented by Cramér (2). Upon performing the integration the resulting density function is

$$f_{\underline{\mu}_n^*\,\mathcal{E}_n}(\underline{\mu}|\underline{\mu}_0',\underline{\mu}_0^*,\ldots,\mu_n^*) = \frac{(2\pi)^{-\frac{r}{2}}\left(\frac{w_n}{w_n+1}\right)^{\frac{r}{2}}\frac{\Gamma(\frac{v_n}{2})}{\Gamma(\frac{v_n-r}{2})}\left(\frac{v_n}{2}\right)^{-\frac{r}{2}}\left|\underline{Q}_n\right|^{-\frac{1}{2}}}{\left[1+\frac{w_n}{(w_n+1)v_n}(\underline{\mu}-\underline{R}_n)^T\,\underline{Q}_n^{-1}\,(\underline{\mu}-\underline{R}_n)\right]^{\frac{v_n}{2}}} \qquad (D.4.8)$$

where $r = 6$. Equation D.4.8 corresponds to the post likelihood developed by Keehn (8).

The mean, which is somewhat obvious from inspection of Equation D.4.8 but which can be verified by performing the onerous integrations, of the post-sampling density, $f_{\underline{\mu}_n^*|\mathcal{E}_n}$, is $\underline{R}_n$. Therefore the best estimate of $\underline{\mu}_n$ given $\mathcal{E}_n$ is

$$\hat{\underline{\mu}}_n = \underline{R}_n \qquad (D.4.9)$$

From Equations D.4.4

$$\underline{\hat{\mu}}_n = \underline{R}_n = \frac{1}{w_n' + 1} [w_n' \underline{R}_n' + \underline{\mu}_n^*]$$

$$= \frac{w_n'}{w_n} \underline{R}_n' + \frac{1}{w_n} \underline{\mu}_n^* \qquad \text{(D.4.10)}$$

To complete the development of the Bayesian estimate the relationship between $\underline{R}_n'$ and the prediction estimate $\underline{\mu}_n'$ must be established and the value of w_n' must be determined. It will be shown here that $\underline{R}_n' = \underline{\mu}_n'$ and $w_n' = w_{n-1}$.

The estimate which minimizes the mean squared error is the conditional espectation of $\underline{\mu}_n$ given $\mathcal{E}_{n-1}$, $E\{\underline{\mu}_n | \mathcal{E}_{n-1}\}$.

Using the augmented moment model, Equation 2.4.5,

$$E\{\underline{\mu}_n | \mathcal{E}_{n-1}\} = E\{[\underline{A}_n \underline{\mu}_{n-1} + \underline{\mu}_{S_n}] | \mathcal{E}_{n-1}\}$$

$$= \underline{A}_n E\{\underline{\mu}_{n-1} | \mathcal{E}_{n-1}\} + \underline{\mu}_{S_n}$$

$$= \underline{A}_n \underline{\hat{\mu}}_{n-1} + \underline{\mu}_{S_n}$$

$$= \underline{\mu}_n'$$

Therefore since $E\{\underline{\mu}_n^* | \mathcal{E}_{n-1}\} = E\{\underline{\mu}_n | \mathcal{E}_{n-1}\}$ (as in the development of $E\{\underline{\mu}_n^* | \mathcal{E}_n\} = E\{\underline{\mu}_n | \mathcal{E}_n\}$)

$$E\{\underline{\mu}_n^* | \mathcal{E}_{n-1}\} = \int_{\underline{\mu}} \underline{\mu} \, f_{\underline{\mu}_n^* | \mathcal{E}_{n-1}}(\underline{\mu}) \, d\underline{\mu} = \underline{\mu}_n' \qquad \text{(D.4.11)}$$

As in the development of $f_{\underline{\mu}_n^* | \mathcal{E}_n}$, $f_{\underline{\mu}_n^* | \mathcal{E}_{n-1}}$, the pre-sampling density of $\underline{\mu}_n^*$, is determined from

$$f_{\underline{\mu}_n^* | \mathcal{E}_{n-1}} = \int_{\underline{m}} \int_{\underline{p}} f_{\underline{\mu}_n^* | \underline{M}_n, \underline{P}_n} \, f_{\underline{M}_n, \underline{P}_n | \mathcal{E}_{n-1}} \, d\underline{m} \, d\underline{p} \qquad \text{(D.4.12)}$$

where $f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_{n-1}}$ is the a priori density of $(\underline{M}_n,\underline{P}_n)$ given by Equation D.4.2 and $f_{\underline{\mu}_n^*|\underline{M}_n,\underline{P}_n} = N(\underline{M}_n,\underline{P}_n)$. Upon performing the integration $f_{\underline{\mu}_n^*|\mathcal{E}_{n-1}}$ is of the same form as Equation D.4.8 with $w_n,v_n,\underline{R}_n,\underline{Q}_n$ replaced by $w_n',v_n',\underline{R}_n',\underline{Q}_n'$. Thus since the mean of $f_{\underline{\mu}_n^*|\mathcal{E}_n}$ is $\underline{R}_n$ the mean of the pre-sampling density, $f_{\underline{\mu}_n^*|\mathcal{E}_{n-1}}$, is $\underline{R}_n'$. Therefore from Equation D.4.11,

$$\underline{R}_n' = \underline{\mu}_n' \tag{D.4.13}$$

The constant w_n' comes from the a priori density of $(\underline{M}_n,\underline{P}_n^{-1})$, Equation D.4.2, G.W.$(w_n',v_n',\underline{R}_n',\underline{Q}_n')$.

Since

$$\underline{\mu}_n' = \underline{A}_n\hat{\underline{\mu}}_{n-1} + \underline{\mu}_{S_n}$$

then

$$\underline{R}_n' = \underline{A}_n\underline{R}_{n-1} + \underline{\mu}_{S_n} \tag{D.4.14}$$

Similarly $\underline{Q}_n'$ is given by

$$\underline{Q}_n' = \underline{A}_n\underline{Q}_{n-1}\underline{A}_n^T \tag{D.4.15}$$

The a priori density of $(\underline{M}_n,\underline{P}_n^{-1})$ is the a posteriori density of $(\underline{M}_{n-1},\underline{P}_{n-1}^{-1})$ with the parameters $\underline{R}_{n-1}$ and $\underline{Q}_{n-1}$ replaced by $\underline{R}_n'$ and $\underline{Q}_n'$ respectively (See Section D.5).

$$f_{\underline{M}_n,\underline{P}_n|\mathcal{E}_{n-1}} = \text{G.W.}(w_n',v_n',\underline{R}_n',\underline{Q}_n') \tag{D.4.16}$$

where

$$w_n' = w_{n-1} \quad , \quad v_n' = v_{n-1} \; ,$$

$$\underline{R}_n' = \underline{A}_n\,\underline{R}_{n-1} + \underline{\mu}_{S_n} \; , \quad \underline{Q}_n' = \underline{A}_n\,\underline{Q}_{n-1}\,\underline{A}_n^T \; .$$

Thus using Equations D.4.13, D.4.16, and D.4.10 $\hat{\underline{\mu}}_n$ becomes

$$\hat{\underline{\mu}}_n = \frac{w_n'}{w_n}\,\underline{\mu}_n' + \frac{1}{w_n}\,\underline{\mu}_n^* \qquad \text{(D.4.17)}$$

where $w_n' = w_{n-1}$ and $w_n = w_{n-1}'$

Equation D.4.17 is then the Bayesian estimate of $\underline{\mu}_n$.

It should be noted that Equation D.4.17 does not use the error covariance matrices, Ψ_n' or $\hat{\Psi}_n^*$, of $\underline{\mu}_n'$ and $\underline{\mu}_n^*$, respectively, in determining $\hat{\underline{\mu}}_n$. Therefore there is no need to determine a relationship between Ψ_n' and $\underline{Q}_n'$ or $\hat{\Psi}_n$ and $\underline{Q}_n$. Thus the selections of $\underline{Q}_n'$ and v_n' used in the a priori Gaussian-Wishart densities of Equations D.4.2 and D.4.16 are arbitrary and useful only to the theoretical development of $\hat{\underline{\mu}}_n$.

Equation D.4.17 can be developed in a simpler manner by assuming that the a priori density on $\underline{M}_n$, $f_{\underline{M}_n|\mathcal{E}_{n-1}}$, is $N(\underline{R}_n', \frac{1}{w_n'} I)$ and that the likelihood of $\underline{\mu}_n^*$, $f_{\underline{\mu}_n^*|\underline{M}_n}$, is $N(\underline{M}_n, I)$. Under these assumptions Bayes' Rule, Equation D.4.1 becomes

$$f_{\underline{M}_n|\mathcal{E}_n} = \frac{f_{\underline{\mu}_n^*|\underline{M}_n}\, f_{\underline{M}_n|\mathcal{E}_{n-1}}}{f_{\underline{\mu}_n^*|\mathcal{E}_{n-1}}}$$

where the a priori density is the reproducing normal density function. The a posteriori density on $\underline{M}_n$, $f_{\underline{M}_n|\mathcal{E}_n}$, then becomes $N(\underline{R}_n, \frac{1}{w_n} I)$. $\underline{R}_{n-1}$, w_{n-1}, $\underline{R}_n'$, w_n', $\underline{R}_n$, and w_n are still defined and related by Equations D.4.4 and D.4.16.

The assumption is made here that the projection of $\frac{1}{w_{n-1}} I$ is $\frac{1}{w_n'} I$, so that $w_n' = w_{n-1}$. This is not the actual case unless $\underline{A}_n$ is an orghogonal matrix. The projection of $\frac{1}{w_n'} I$ is given by

$$\underline{A}_n \frac{1}{w'_n} I \underline{A}_n^T = \frac{1}{w'_n} \underline{A}_n \underline{A}_n^T$$

which is not equal to $\frac{1}{w'_n}$ I unless $\underline{A}_n^{-1} = \underline{A}_n^T$, i.e., $\underline{A}_n$ is an orthogonal matrix. Section D.5 shows that such an assumption is not necessary in the Gaussian-Wishart formulation.

This simpler procedure is adapted from what is sometimes referred to as learning the mean vector of normal patterns (9).

D.5 The Density Function of the Projection of $(\underline{M}_{n-1}, \underline{P}_{n-1})$. In this section it is shown that if $(\underline{M}_{n-1}, \underline{P}_{n-1}^{-1})$ is projected to $(\underline{M}_n, \underline{P}_n^{-1})$ according to

$$\underline{M}_n = \underline{A}_n \underline{M}_{n-1} + \underline{\mu}_{S_n} \tag{D.5.1}$$

and

$$\underline{P}_n^{-1} = \underline{A}_n \underline{P}_{n-1}^{-1} \underline{A}_n^T \tag{D.5.2}$$

and if the density function of $(\underline{M}_{n-1}, \underline{P}_{n-1})$ is the Gaussian-Wishart density function then the density function of $(\underline{M}_n, \underline{P}_n)$ is also the Gaussian-Wishart density function.

In order to ease the presentation the notation is simplified. Let $\underline{A}_n = \underline{A}$, $\underline{M}_{n-1} = \underline{M}$, $\underline{\mu}_{S_n} = \underline{B}$, $\underline{M}_n = \underline{M}_1$, $\underline{P}_{n-1} = \underline{P}$, and $\underline{P}_n = \underline{P}_1$. With this simplified notation it is shown that if

$$(\underline{M}, \underline{P}) \sim \text{G.W.}(w, v, \underline{R}, \underline{Q}), \tag{D.5.3}$$

$$\underline{M}_1 = \underline{A}\,\underline{M} + \underline{B} \tag{D.5.4}$$

and

$$\underline{P}_1^{-1} = \underline{A}\ \underline{P}^{-1}\ \underline{A}^T \tag{D.5.5}$$

then

$$(\underline{M}_1, \underline{P}_1) \sim G.W.(w_1, v_1, \underline{R}_1, \underline{Q}_1) \tag{D.5.6}$$

where

$$w_1 = w \quad , \qquad v_1 = v \quad ,$$

$$\underline{R}_1 = \underline{A}\ \underline{R} + \underline{B} \quad , \tag{D.5.7}$$

$$\underline{Q}_1 = \underline{A}\ \underline{Q}\ \underline{A}^T$$

Note that if $\underline{P}_1^{-1} = \underline{A}\ \underline{P}^{-1}\ \underline{A}^T$ then

$$\underline{P}_1 = (\underline{A}^T)^{-1}\ \underline{P}\ \underline{A}^{-1} \tag{D.5.8}$$

As indicated in Section D.3

$$f_{\underline{M},\underline{P}}(\underline{m},\underline{p}) = f_{\underline{M}|\underline{P}}(\underline{m}|\underline{p})\ f_{\underline{P}}(\underline{p}) \tag{D.5.9}$$

where

$$\begin{aligned} f_{\underline{M}|\underline{P}}(\underline{m}|\underline{p}) &= N(\underline{R}, \tfrac{1}{w}\ \underline{p}^{-1}) \\ &= (2\pi)^{-\frac{r}{2}} \left|w\ \underline{p}\right|^{\frac{1}{2}} \exp[-\tfrac{1}{2}\ (\underline{m} - \underline{R})^T\ w\ \underline{p}\ (\underline{m} - \underline{R})] \end{aligned} \tag{D.5.10}$$

and

$$f_{\underline{P}}(\underline{p}) = C_{r,v} \left|\tfrac{v}{2}\ \underline{Q}\right|^{\frac{v-1}{2}} \left|\underline{p}\right|^{\frac{v-r-2}{2}} \exp[-\tfrac{1}{2}\ \mathrm{tr}\ v\ \underline{Q}\ \underline{p}] \tag{D.5.11}$$

where the constant $C_{r,v}$ is defined in Section D.3.

Since Equations D.5.4 and D.5.8 are linear transformations of $\underline{M}$ and $\underline{P}$ the density function $f_{\underline{M}_1,\underline{P}_1}(\underline{m}_1, \underline{p}_1)$ can be expressed as

$$f_{\underline{M}_1,\underline{P}_1}(\underline{m}_1,\underline{p}_1) = \frac{1}{|J|} f_{\underline{M},\underline{P}}(\underline{m} = \underline{A}^{-1}[\underline{m}_1-\underline{B}], \underline{p} = \underline{A}^T\underline{p}_1\underline{A})$$

$$= \frac{1}{|J|} f_{\underline{M}|\underline{P}}(\underline{m} = \underline{A}^{-1}[\underline{m}_1-\underline{B}]|\underline{p} = \underline{A}^T\underline{p}_1\underline{A}) f_{\underline{P}}(\underline{p} = \underline{A}^T\underline{p}_1\underline{A}) \qquad \text{(D.5.12)}$$

First consider $f_{\underline{P}}(\underline{p} = \underline{A}^T \underline{p}_1 \underline{A})$.

$$f_{\underline{P}}(\underline{p} = \underline{A}^T \underline{p}_1 \underline{A}) = C_{r,v} \left|\frac{v}{2} \underline{Q}\right|^{\frac{v-1}{2}} \left|\underline{A}^T \underline{p}_1 \underline{A}\right|^{\frac{v-r-2}{2}} \exp[-\tfrac{1}{2} \operatorname{tr} \underline{Q}\, \underline{A}^T \underline{p}_1 \underline{A}] \qquad \text{(D.5.13)}$$

From the properties of the trace of a matrix

$$\operatorname{tr} v\, \underline{Q}\, \underline{A}^T \underline{p}_1 \underline{A} = \operatorname{tr} v\, \underline{A}\, \underline{Q}\, \underline{A}^T \underline{p}_1 \qquad \text{(D.5.14)}$$

From the properties of determinants

$$\left|\underline{A}^T \underline{p}_1 \underline{A}\right| = \left|\underline{A}^T\right|\left|\underline{p}_1\right|\left|\underline{A}\right| = \left|\underline{A}^T\right|\left|\underline{A}\right|\left|\underline{p}_1\right| = \left|\underline{A}^T \underline{A}\right|\left|\underline{p}_1\right| \qquad \text{(D.5.15)}$$

and

$$\left|\underline{A}^T \underline{p}_1 \underline{A}\right|^{\frac{v-r-2}{2}} = [\left|\underline{A}^T \underline{A}\right|\left|\underline{p}_1\right|]^{\frac{v-r-2}{2}} = \left|\underline{A}^T \underline{A}\right|^{\frac{v-r-2}{2}} \left|\underline{p}_1\right|^{\frac{v-r-2}{2}}$$

$$= \left|\underline{A}^T \underline{A}\right|^{\frac{v-1}{2}} \left|\underline{A}^T \underline{A}\right|^{\frac{-r-1}{2}} \left|\underline{p}_1\right|^{\frac{v-r-2}{2}} \qquad \text{(D.5.16)}$$

then

$$\left|\frac{v}{2} \underline{Q}\right|^{\frac{v-1}{2}} \left|\underline{A}^T \underline{p}_1 \underline{A}\right|^{\frac{v-r-2}{2}} = \left|\frac{v}{2} \underline{Q}\right|^{\frac{v-1}{2}} \left|\underline{A}^T \underline{A}\right|^{\frac{v-1}{2}} \left|\underline{A}^T \underline{A}\right|^{\frac{-r-1}{2}} \left|\underline{p}_1\right|^{\frac{v-r-2}{2}}$$

$$= [\left|\frac{v}{2} \underline{Q}\right|\left|\underline{A}^T\right|\left|\underline{A}\right|]^{\frac{v-1}{2}} \left|\underline{A}^T \underline{A}\right|^{\frac{-r-1}{2}} \left|\underline{p}_1\right|^{\frac{v-r-2}{2}}$$

$$= \left|\frac{v}{2} \underline{A}\, \underline{Q}\, \underline{A}^T\right|^{\frac{v-1}{2}} \left|\underline{A}^T \underline{A}\right|^{\frac{-r-1}{2}} \left|\underline{p}_1\right|^{\frac{v-r-2}{2}} \qquad \text{(D.5.17)}$$

Using Equations D.5.14 and D.5.17 in Equation D.5.13

$$f_{\underline{P}}(\underline{p} = \underline{A}^T \underline{p}_1 \underline{A})$$

$$= \left|\underline{A}^T \underline{A}\right|^{\frac{-r-1}{2}} c_{r,v}\left|\frac{v}{2} \underline{A} \underline{Q} \underline{A}^T\right|^{\frac{v-1}{2}} \left|\underline{p}_1\right|^{\frac{v-r-2}{2}} \exp\left[- \frac{1}{2} \text{tr } v \underline{A} \underline{Q} \underline{A}^T \underline{p}_1\right] \tag{D.5.18}$$

With $\underline{Q}_1$ as given in Equation D.5.7 $f_{\underline{P}}(\underline{p} = \underline{A}^T \underline{p}_1 \underline{A})$ becomes

$$f_{\underline{P}}(\underline{p} = \underline{A}^T \underline{p}_1 \underline{A})$$

$$= \left|\underline{A}^T \underline{A}\right|^{\frac{-r-1}{2}} c_{r,v}\left|\frac{v}{2} \underline{Q}_1\right|^{\frac{v-1}{2}} \left|\underline{p}_1\right|^{\frac{v-r-2}{2}} \exp\left[- \frac{1}{2} \text{tr } v \underline{Q}_1 \underline{p}_1\right] \tag{D.5.19}$$

which, except for the constant $\left|\underline{A}^T \underline{A}\right|^{\frac{-r-1}{2}}$, is of the same form as Equation D.5.11.

Now consider $f_{\underline{M}|\underline{P}}(\underline{m} = \underline{A}^{-1} [\underline{m}_1 - \underline{B}] \mid \underline{p} = \underline{A}^T \underline{p}_1 \underline{A})$.

$$f_{\underline{M}|\underline{P}}(\underline{m} = \underline{A}^{-1}[\underline{m}_1-\underline{B}] \mid \underline{p} = \underline{A}^T\underline{p}_1\underline{A})$$

$$= (2\pi)^{\frac{-r}{2}} \left|w\underline{A}^T\underline{p}_1\underline{A}\right|^{\frac{1}{2}} \exp\{- \tfrac{1}{2}(\underline{A}^{-1}[\underline{m}_1-\underline{B}]-\underline{R})^T w\underline{A}^T\underline{p}_1\underline{A}(\underline{A}^{-1}[\underline{m}_1-\underline{B}]-\underline{R})\} \tag{D.5.20}$$

Again from the properties of determinants

$$\left|w \underline{A}^T \underline{p}_1 \underline{A}\right|^{\frac{1}{2}} = \left|\underline{A}^T \underline{A}\right|^{\frac{1}{2}} \left|w \underline{p}_1\right|^{\frac{1}{2}} \tag{D.5.21}$$

The exponential of Equation D.5.20 contains

$$(\underline{A}^{-1}[\underline{m}_1 - \underline{B}] - \underline{R})^T w \underline{A}^T \underline{p}_1 \underline{A}(\underline{A}^{-1}[\underline{m}_1 - \underline{B}] - \underline{R})$$

$$= ([\underline{m}_1^T - \underline{B}^T](\underline{A}^{-1})^T - \underline{R}^T) \underline{A}^T w \underline{p}_1 \underline{A}(\underline{A}^{-1}[\underline{m}_1 - \underline{B}] - \underline{R})$$

$$= (\underline{m}_1^T - \underline{B}^T - \underline{R}^T \underline{A}^T) w \underline{p}_1 (\underline{m}_1 - \underline{B} - \underline{A} \underline{R})$$

$$= (\underline{m}_1 - \underline{A} \underline{R} - \underline{B})^T w \underline{p}_1 (\underline{m}_1 - \underline{A} \underline{R} - \underline{B}) \tag{D.5.22}$$

Using Equations D.5.21 and D.5.22 in Equation D.5.20

$$f_{\underline{M}|\underline{P}}(\underline{m} = \underline{A}^{-1}[\underline{m}_1-\underline{B}] \mid \underline{p} = \underline{A}^T\underline{p}_1\underline{A})$$
$$= |\underline{A}^T\underline{A}|^{\frac{1}{2}} (2\pi)^{\frac{-r}{2}} |w\ \underline{p}_1|^{\frac{1}{2}} \exp\{-\tfrac{1}{2}(\underline{m}_1-\underline{AR}-\underline{B})^T\ w\ \underline{p}_1(\underline{m}_1-\underline{AR}-\underline{B})\} \quad (D.5.23)$$

With $\underline{R}_1$ as given in Equations D.5.7, Equation D.5.23 becomes

$$f_{\underline{M}|\underline{P}}(\underline{m} = \underline{A}^{-1}[\underline{m}_1 - \underline{B}] \mid \underline{p} = \underline{A}^T\ \underline{p}_1\ \underline{A})$$
$$= |\underline{A}^T\ \underline{A}|^{\frac{1}{2}} (2\pi)^{\frac{-r}{2}} |w\ \underline{p}_1|^{\frac{1}{2}} \exp[-\tfrac{1}{2}(\underline{m} - \underline{R}_1)^T\ w\ \underline{p}_1(\underline{m}_1 - \underline{R}_1)]$$
(D.5.24)

which, except for the constant $|\underline{A}^T\ \underline{A}|^{\frac{1}{2}}$, is of the same form as Equation D.5.10.

Using Equations D.5.19 and D.5.24 in Equation D.5.12 the joint density function of $\underline{M}_1$ and $\underline{P}_1$ becomes

$$f_{\underline{M}_1,\underline{P}_1}(\underline{m}_1,\underline{p}_1) = \frac{1}{|J|}\ |\underline{A}^T\underline{A}|^{\frac{-r}{2}}(2\pi)^{\frac{-r}{2}}|w\ \underline{p}_1|^{\frac{1}{2}} \exp[-\tfrac{1}{2}(\underline{m}_1-\underline{R}_1)^T w\ \underline{p}_1(\underline{m}_1-\underline{R}_1)]$$
$$\cdot\ C_{r,v}\left|\tfrac{v}{2}\ \underline{Q}_1\right|^{\frac{v-1}{2}} |\underline{p}_1|^{\frac{v-r-2}{2}} \exp[-\tfrac{1}{2}\ tr\ v\ \underline{Q}_1\ \underline{p}_1] \quad (D.5.25)$$

and since $f_{\underline{M}_1,\underline{P}_1}(\underline{m}_1,\underline{p}_1)$ must be a density function

$$|A^T\ A|^{\frac{-r}{2}} = |J|$$

Therefore

$$f_{\underline{M}_1,\underline{P}_1}(\underline{m}_1,\underline{p}_1) = f_{\underline{M}_1|\underline{P}_1}(\underline{m}_1|\,\underline{p}_1)f_{\underline{P}_1}(\underline{p}_1) \quad (D.5.26)$$

where

$$f_{\underline{M}_1|\underline{P}_1}(\underline{m}_1|\,\underline{p}_1) = (2\pi)^{\frac{-r}{2}} |w_1\ \underline{p}_1|^{\frac{1}{2}} \exp[-\tfrac{1}{2}(\underline{m}_1 - \underline{R}_1)^T\ w_1\ \underline{p}_1(\underline{m}_1 - \underline{R}_1)]$$
(D.5.27)

and

$$f_{\underline{P}_1}(\underline{p}_1) = C_{r,v_1} \left|\frac{v_1}{2}\underline{Q}_1\right|^{\frac{v_1-1}{2}} \left|\underline{p}_1\right|^{\frac{v_1-r-2}{2}} \exp\left[-\frac{1}{2}\operatorname{tr} v_1 \underline{Q}_1 \underline{p}_1\right] \tag{D.5.28}$$

with $w_1, v_1, \underline{R}_1$, and $\underline{Q}_1$ given in Equations D.5.7.

Equation D.5.26 with Equations D.5.27 and D.5.28 is the Gaussian-Wishart density function, so that

$$(\underline{M}_1, \underline{P}_1^{-1}) \sim G.W.(w_1, v_1, \underline{R}_1, \underline{Q}_1)$$

Returning to the original notation it is concluded that if

$$(\underline{M}_{n-1}, \underline{P}_{n-1}) \sim G.W.(\underline{w}_{n-1}, \underline{v}_{n-1}, \underline{R}_{n-1}, \underline{Q}_{n-1})$$

and

$$\underline{M}_n = \underline{A}_n \underline{M}_{n-1} + \mu_{S_n}$$

$$\underline{P}_n^{-1} = \underline{A}_n \underline{P}_{n-1}^{-1} \underline{A}_n^T$$

then

$$(\underline{M}_n, \underline{P}_n) \sim G.W.(w_n', v_n', \underline{R}_n', \underline{Q}_n')$$

where

$$w_n' = w_{n-1} \quad , \quad v_n' = v_{n-1}$$

$$\underline{R}_n' = \underline{A}_n \underline{R}_{n-1} + \mu_{S_n}$$

$$\underline{Q}_n' = \underline{A}_n \underline{Q}_{n-1} \underline{A}_n^T$$

D.6 Bayesian Recursive Moment Estimation Algorithm and Summary.

In order to summarize the results developed in this Appendix a Bayesian recursive moment estimation algorithm is presented and some comments are offered on the inadequacies of Bayesian moment estimation.

The Bayesian Recursive Moment Estimation Algorithm:

(1) Determine the prediction estimate, $\underline{\mu}_n'$, from $\hat{\underline{\mu}}_{n-1}$,

$$\underline{\mu}_n' = \underline{A}_n \hat{\underline{\mu}}_{n-1} + \underline{\mu}_{S_n} \qquad (D.6.1)$$

and w_n', the confidence factor in $\underline{\mu}_n'$, from w_{n-1}, the confidence factor in $\hat{\underline{\mu}}_{n-1}$,

$$w_n' = w_{n-1} \qquad (D.6.2)$$

(2) From the observations of X_n, $\underline{\mu}_n^*$, the data estimate, or observation of $\underline{\mu}_n$, is computed by Equations 3.2.4 and 3.2.5.

(3) Using Equation D.4.17 the Bayesian estimate, $\hat{\underline{\mu}}_n$, is determined from $\underline{\mu}_n'$, $\underline{\mu}_n^*$, and w_n',

$$\hat{\underline{\mu}}_n = \frac{w_n'}{w_n} \underline{\mu}_n' + \frac{1}{w_n} \underline{\mu}_n^* \qquad (D.6.3)$$

where $w_n = w_n' + 1$

Then the algorithm begins again with $\hat{\underline{\mu}}_n$ projected to $\underline{\mu}_{n+1}'$, etc.

Actually Equation D.6.3 is an average of $\underline{\mu}_n^*$ with the projections of all the previous observations and the initial estimate of $\underline{\mu}_0$, $\underline{\mu}_0', \underline{\mu}_0^*, \underline{\mu}_1^*, \ldots, \underline{\mu}_n^*$. For example let $w_0' = 1$, then

$$\hat{\underline{\mu}}_0 = \frac{1}{2} \underline{\mu}_0' + \frac{1}{2} \underline{\mu}_0^*$$

From Equation D.6.1

$$\underline{\mu}_1' = \underline{A}_1 \hat{\underline{\mu}}_0 + \underline{\mu}_{S_1} = \frac{1}{2} \underline{A}_1 \underline{\mu}_0' + \frac{1}{2} \underline{A}_1 \underline{\mu}_0^* + \underline{\mu}_{S_1}$$

then using Equation D.6.3

$$\underline{\hat{\mu}}_1 = \frac{2}{3}\underline{\mu}_1' + \frac{1}{3}\underline{\mu}_1^* = \frac{1}{3}\underline{A}_1\underline{\mu}_0' + \frac{1}{3}\underline{A}_1\underline{\mu}_0^* + \frac{1}{3}\underline{\mu}_1^* + \frac{2}{3}\underline{\mu}_{S_1}$$

Similarly

$$\underline{\mu}_2' = \underline{A}_2\underline{\hat{\mu}}_1 + \underline{\mu}_{S_2} = \frac{1}{3}\underline{A}_1\ \underline{A}_2\underline{\mu}_0' + \frac{1}{3}\underline{A}_1\underline{A}_2\underline{\mu}_0^* + \frac{1}{3}\underline{A}_2\underline{\mu}_1^* + \frac{2}{3}\underline{A}_2\underline{\mu}_{S_1} + \underline{\mu}_{S_2}$$

$$\underline{\hat{\mu}}_2 = \frac{3}{4}\underline{\mu}_2' + \frac{1}{4}\underline{\mu}_2^* = \frac{1}{4}\underline{A}_1\underline{A}_2\underline{\mu}_0' + \frac{1}{4}\underline{A}_1\underline{A}_2\underline{\mu}_0^* + \frac{1}{4}\underline{A}_2\underline{\mu}_1^* + \frac{1}{4}\underline{\mu}_2^*$$

$$+ \frac{2}{4}\underline{A}_2\underline{\mu}_{S_1} + \frac{3}{4}\underline{\mu}_{S_2}$$

$$\vdots$$

$$\underline{\hat{\mu}}_n = \frac{n-1}{n}\underline{\mu}_n' + \frac{1}{n}\underline{\mu}_n^*$$

$$= \frac{1}{n}\underline{A}_1\cdots\underline{A}_n\underline{\mu}_0' + \frac{1}{n}\underline{A}_1\cdots\underline{A}_n\underline{\mu}_0^* + \frac{1}{n}\underline{A}_2\cdots\underline{A}_n\underline{\mu}_1^* + \cdots + \frac{1}{n}\underline{\mu}_n^*$$

$$+ \frac{2}{n}\underline{A}_2\cdots\underline{A}_n\underline{\mu}_{S_1} + \frac{3}{n}\underline{A}_3\cdots\underline{A}_n\underline{\mu}_{S_2} + \cdots + \frac{n-1}{n}\underline{\mu}_{S_n} \qquad (D.6.4)$$

There are some obvious deficiencies in determining $\underline{\hat{\mu}}_n$ in the manner of Equations D.6.3 or D.6.4. The weight attached to $\underline{\mu}_n^*$ is always $\frac{1}{w_n' + 1}$. Although in this study the number of observations of X_n used to determine $\underline{\mu}_n^*$ is implicitly considered to be a constant k for each n, k could vary with n. In either case the weight attached to $\underline{\mu}_n^*$ would be a better measure of how good $\underline{\mu}_n^*$ is as an estimate of $\underline{\mu}_n$ if it was a function of k. If k is large then $\frac{1}{w_n' + 1}$ should be large. If k is small then $\frac{1}{w_n' + 1}$ should be small.

As indicated at the end of Section D.4 neither Ψ_n' or $\hat{\Psi}_n^*$ as developed in Chapter III enters into the actual determination of $\underline{\hat{\mu}}_n$. Since Ψ_n' and $\hat{\Psi}_n^*$ are measures of the goodness of $\underline{\mu}_n'$ and $\underline{\mu}_n^*$, respectively, it would be desirable for the weights attached to $\underline{\mu}_n'$ and $\underline{\mu}_n^*$ to be functions of Ψ_n' and $\hat{\Psi}_n^*$. $\hat{\Psi}_n^*$ is a function of k so that use of $\hat{\Psi}_n^*$ in the weight of $\underline{\mu}_n^*$ would make use of k also, which is desired as indicated above.

The recursive moment estimation scheme developed in Chapter III possesses these desirable properties.

VITA

John Cleveland Hooker

Candidate for the Degree of

Doctor of Philosophy

Thesis: APPROXIMATING THE DISTRIBUTION FUNCTION OF TIME-VARYING EQUIPMENT OUTPUTS BY RECURSIVE MOMENT ESTIMATION

Major Field: Electrical Engineering

Biographical:

Personal Data: Born on May 10, 1939, in Terrell, Texas, the son of J. Loy and Lena H. Hooker.

Education: Attended primary and secondary schools in Dallas, Texas, and graduated from Sunset High School in May, 1957; received the Bachelor of Science degree in Electrical Engineering from the University of Texas at Arlington in June, 1962; received the Master of Science degree in Electrical Engineering from Oklahoma State University in May, 1965; completed requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1968.

Professional Experience: Employed by Ling-Temco-Vought, Inc., Dallas, Texas, as an engineering co-operative student from June, 1957 to June, 1962. Employed by LTV Electrosystems, Garland, Texas, as a guidance and controls engineer from June, 1962 to September, 1963 and during the summer of 1964. Employed by the School of Electrical Engineering of Oklahoma State University as a graduate research assistant from January, 1965 to February, 1968.

Professional Organizations: Member of Eta Kappa Nu, Pi Mu Epsilon, and the Institute of Electrical and Electronic Engineers.