International Institute for Applied Systems Analysis
IIASA
www.iiasa.ac.at

Nonstationary Means in a Multinormal Process

Winkler, R.L. and Barry, C.B.

IIASA Research Report
October 1973

RR-73-9

NONSTATIONARY MEANS IN A MULTINORMAL PROCESS

Robert L. Winkler

Christopher B. Barry

October 1973

Nonstationary Means in a Multinormal Process*

Robert L. Winkler** and Christopher B. Barry***

1. Introduction

Multinormal processes have received considerable attention in the statistical literature (e.g. see Johnson and Kotz [8, Chapters 35-36]). Although much of this work has been in the "classical" tradition, the Bayesian approach to multinormal processes is relatively straightforward. Consider a multinormal process of dimension M with unknown mean vector $\tilde{\underset{\sim}{\mu}}$ and known covariance matrix $\underset{\sim}{\Sigma}$. In making inferences about $\tilde{\underset{\sim}{\mu}}$, Bayes' theorem can be expressed in the form

$$f(\underset{\sim}{\mu}|x) = f(\underset{\sim}{\mu})f(\underset{\sim}{x}|\underset{\sim}{\mu},\underset{\sim}{\Sigma})/\int f(\underset{\sim}{\mu})f(\underset{\sim}{x}|\underset{\sim}{\mu},\underset{\sim}{\Sigma})\, d\underset{\sim}{\mu}$$

with the usual abuse of functional notation. That is, assuming that the prior information about $\tilde{\underset{\sim}{\mu}}$ can be expressed in the form of a prior distribution $f(\underset{\sim}{\mu})$, and that sample information from the process, denoted by $\underset{\sim}{x}$, can be summarized (with respect to inferences concerning $\tilde{\underset{\sim}{\mu}}$) by the likelihood function $f(\underset{\sim}{x}|\underset{\sim}{\mu})$, Bayes' theorem revises the prior

---

*This paper was presented at the Institute of Mathematical Statistics Meetings.

**Graduate School of Business, Indiana University, Indiana, U.S.A.; research scholar at the International Institute for Applied Systems Analysis, Laxenburg, Austria.

***University of Florida, U.S.A.

distribution on the basis of the new information, yielding a posterior distribution $f(\underset{\sim}{\mu}|\underset{\sim}{x})$. This provides a framework for inferences about $\underset{\sim}{\tilde{\mu}}$ and for decisions that are related to $\underset{\sim}{\tilde{\mu}}$. For example, if $f(\underset{\sim}{\mu})$ is a multinormal distribution, then it is conjugate to the data-generating process in this instance, and the application of Bayes' theorem results in a posterior distribution that is also a multinormal distribution (e.g. see Raiffa and Schlaifer [9], or DeGroot [6]).

The inferential model presented above is a stationary model. That is, it assumes that $\underset{\sim}{\tilde{\mu}}$ takes on a single value and that $f(\underset{\sim}{\mu})$ and $f(\underset{\sim}{\mu}|\underset{\sim}{x})$ represent uncertainty about what that value is. For example, $\underset{\sim}{\tilde{\mu}}$ could represent the mean rate of change of the prices of M securities, the mean change in the pulse rate of M individuals in response to a particular drug, the mean daily sales at M stores, and so on. In each case, $\underset{\sim}{\tilde{\mu}}$ is assumed to be fixed but unknown.

In many real-world situations, the assumption of stationarity is questionable. For example, security price changes may be well-represented by a nonstationary model; Hsu, Miller, and Wichern [7] claim that a nonstationary normal process is consistent with empirical evidence (also, see Boness, Chen, and Jatusipitak [4]). Production processes may be stationary over short periods of time, but in most cases it would be expected that for a lengthy period, stationarity would be a doubtful assumption.

Despite the apparent existence of nonstationarity in many situations, few Bayesian models for dealing with nonstationary processes have been developed. Bather [2] develops a model in which the mean of a univariate normal process shifts stochastically over time and uses this model in the study of control charts (also, see Carter [5]). Some basic notions underlying this model are treated much more generally in Bather [3].

In this paper, we consider inferences about the mean vector of a multinormal process when the mean vector shifts from period to period, with the shifts governed by an independent multinormal process. This is an extension to the multivariate case of the situation treated in Bather [2]. The model is presented in Section 2, some applications to portfolio analysis are considered in Section 3, and Section 4 contains a brief summary and discussion.

2. The Development of the Model

Consider a data-generating process that generates M-vectors (column vectors) of observations $\tilde{x}_{t1}, \tilde{x}_{t2}, \ldots$ during time period t according to a multinormal process with mean $\tilde{\mu}_t$ and covariance matrix Σ. The covariance matrix Σ is known and does not change over time,[1] whereas $\tilde{\mu}_t$ is not known and may change over time. In particular, values of the mean vector for successive time periods are related as

---

[1] In this paper, all covariance matrices are assumed to be positive-definite and symmetric.

$$\tilde{\mu}_{t+1} = \tilde{\mu}_t + \tilde{\varepsilon}_{t+1} , \quad t = 1,2,\ldots, \tag{2.1}$$

where $\tilde{\varepsilon}_{t+1}$ is a multinormal "random shock" term independent of $\tilde{\mu}_t$ with known mean e and covariance matrix Ω.

If the prior distribution of $\tilde{\mu}_t$ at the beginning of time period t is represented by $f(\mu_t)$, and a sample of size n_t during period t yields $X_t = (x_{t1}, x_{t2},\ldots,x_{tn_t})$, then Bayes' theorem can be used to revise the distribution of $\tilde{\mu}_t$

$$f(\mu_t|X_t) \propto f(\mu_t)f(X_t|\mu_t) \quad . \tag{2.2}$$

In general, this application of Bayes' theorem may be difficult to carry out. If the prior distribution of $\tilde{\mu}_t$ is multinormal, however, it is possible to summarize the revision of the distribution of $\tilde{\mu}_t$ in terms of two simple formulas.

Theorem 1. If the prior distribution of $\tilde{\mu}_t$ is a multinormal distribution with mean m_t' and covariance matrix S_t', and if $\tilde{x}_{ti}$, $i = 1,\ldots,n_t$, are independent and identically distributed conditional upon μ_t, each having a multinormal distribution with mean μ_t and covariance matrix Σ, then the posterior distribution of $\tilde{\mu}_t$ conditional upon $X_t = (x_{t1},\ldots,x_{tn_t})$ is a multinormal distribution with mean m_t'' and covariance matrix S_t'', where

$$m_t'' = S_t''(S_t'^{-1}m_t' + n_t\, \Sigma^{-1}m_t) \tag{2.3}$$

and

$$\underset{\sim}{S}_t'' = (\underset{\sim}{S}_t'^{-1} + n_t\underset{\sim}{\Sigma}^{-1})^{-1} \quad . \tag{2.4}$$

Here $m_t = \sum_{i=1}^{n_t} \underset{\sim}{x}_{ti}/n_t$ is the sample mean vector, and $(n_t, \underset{\sim}{m}_t)$ is sufficient for inferences concerning $\tilde{\underset{\sim}{\mu}}_t$.

Proof. The proof follows directly from DeGroot [6, pp. 175-176]. Factoring out terms in $f(\underset{\sim}{\mu}_t)$ not involving $\underset{\sim}{\mu}_t$, we have

$$f(\underset{\sim}{\mu}_t) \propto \exp[-(1/2)(\underset{\sim}{\mu}_t - \underset{\sim}{m}_t')^T \underset{\sim}{S}_t'^{-1}(\underset{\sim}{\mu}_t - \underset{\sim}{m}_t')] \quad ,$$

where "T" denotes transposition. Furthermore, since the data-generating process is an independent multinormal process, the likelihood function can be written

$$f(\underset{\sim}{X}_t|\underset{\sim}{\mu}_t) \propto \exp[-(1/2) \sum_{i=1}^{n_t} (\underset{\sim}{x}_{ti} - \underset{\sim}{\mu}_t)^T \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{x}_{ti} - \underset{\sim}{\mu}_t)]$$

$$\propto \exp[-(1/2)(\underset{\sim}{m}_t - \underset{\sim}{\mu}_t)^T(n_t\underset{\sim}{\Sigma}^{-1})(\underset{\sim}{m}_t - \underset{\sim}{\mu}_t)] \quad .$$

From this likelihood function, it can be seen that $(n_t, \underset{\sim}{m}_t)$ is sufficient. Applying Bayes' theorem yields

$$f(\underset{\sim}{\mu}_t|\underset{\sim}{X}_t) \propto f(\underset{\sim}{\mu}_t)f(\underset{\sim}{X}_t|\underset{\sim}{\mu}_t)$$

$$= \exp\{-(1/2)[(\underset{\sim}{\mu}_t - \underset{\sim}{m}_t')^T \underset{\sim}{S}_t'^{-1}(\underset{\sim}{\mu}_t - \underset{\sim}{m}_t')$$

$$+ (\underset{\sim}{m}_t - \underset{\sim}{\mu}_t)^T(n_t\underset{\sim}{\Sigma}^{-1})(\underset{\sim}{m}_t - \underset{\sim}{\mu}_t)]\} \quad .$$

Combining terms in the exponent, completing the square on $\underset{\sim}{\mu}_t$, and factoring out terms not involving $\underset{\sim}{\mu}_t$, we have

$$f(\underset{\sim}{\mu}_t | \underset{\sim}{X}_t) \propto \exp[-(1/2)(\underset{\sim}{\mu}_t - \underset{\sim}{m}''_t)^T \underset{\sim}{S}''^{-1}_t(\underset{\sim}{\mu}_t - \underset{\sim}{m}''_t)] \quad , \tag{2.5}$$

where $\underset{\sim}{m}''_t$ and $\underset{\sim}{S}''_t$ are given by (2.3) and (2.4). The distribution in (2.5) is in the form of a multinormal density with mean $\underset{\sim}{m}''_t$ and covariance matrix $\underset{\sim}{S}''_t$.

During the time period, then, the distribution of $\underset{\sim}{\tilde{\mu}}_t$ is revised as new information becomes available. At the end of time period t (the beginning of time period t + 1), the data-generating process is governed by a new mean vector, $\underset{\sim}{\tilde{\mu}}_{t+1}$, so it is necessary to use the posterior distribution of $\underset{\sim}{\tilde{\mu}}_t$ and the relation given by (2.1) to determine the prior distribution of $\underset{\sim}{\tilde{\mu}}_{t+1}$ at the beginning of time period t + 1.

<u>Theorem 2</u>. If the posterior distribution of $\underset{\sim}{\tilde{\mu}}_t$ is as derived in Theorem 1, and the relationship between $\underset{\sim}{\tilde{\mu}}_t$ and $\underset{\sim}{\tilde{\mu}}_{t+1}$ is given by (2.1), where $\underset{\sim}{\tilde{\varepsilon}}_t$ is independent of $\underset{\sim}{\tilde{\mu}}_t$ and multinormal with mean $\underset{\sim}{e}$ and covariance matrix $\underset{\sim}{\Omega}$, then the prior distribution of $\underset{\sim}{\tilde{\mu}}_{t+1}$ is a multinormal distribution with mean $\underset{\sim}{m}'_{t+1}$ and covariance matrix $\underset{\sim}{S}'_{t+1}$, where

$$\underset{\sim}{m}'_{t+1} = \underset{\sim}{m}''_t + \underset{\sim}{e} \tag{2.6}$$

and

$$\underset{\sim}{S}'_{t+1} = \underset{\sim}{S}''_t + \underset{\sim}{\Omega} \quad . \tag{2.7}$$

<u>Proof</u>. Since $\tilde{\mu}_{t+1}$ is a linear combination of independent multinormal random vectors, the result is trivial.

Combining the results of Theorems 1 and 2, we see that prior distributions for mean vectors in successive periods are related as

$$m'_{t+1} = (S'^{-1}_t + n_t\Sigma^{-1})^{-1}(S'^{-1}_t m'_t + n_t\Sigma^{-1} m_t) + e \qquad (2.8)$$

and

$$S'_{t+1} = (S'^{-1}_t + n_t\Sigma^{-1})^{-1} + \Omega \quad . \qquad (2.9)$$

These formulas hold for $t = 1,2,\ldots;$ if the initial prior distribution at the beginning of period one is known, then (2.8) and (2.9) can be applied each period after (n_t, m_t) is observed.

The updating procedure for the model developed in this section is relatively straightforward, but difficulties are encountered in attempting to investigate limiting properties of the model. Starting with m'_1 and S'_1 and repeatedly applying (2.8) and (2.9), it is possible to express m'_t and S'_t as functions of the initial values m'_1 and S'_1, the sample statistics (n_i, m_i), $i = 1,\ldots,t-1$, and the known parameters Σ, Ω, and e. However, these functions are quite complicated, as terms such as $(S'^{-1}_t + n_t\Sigma^{-1})^{-1}$ in (2.9), when applied repeatedly, do not yield simple expressions. In the univariate case, such difficulties are not encountered, because the respective variances can all be expressed as constant multiples of each other.

To avoid the difficulties mentioned in the preceding paragraph, we will investigate a simplified form of the general model. The simplifying assumptions are that $\underset{\sim}{S}_1'$ and $\underset{\sim}{\Omega}$ are constant multiples of $\underset{\sim}{\Sigma}$

$$\underset{\sim}{S}_1' = (n_1')^{-1} \underset{\sim}{\Sigma} \quad , \tag{2.10}$$

and

$$\underset{\sim}{\Omega} = w^{-1} \underset{\sim}{\Sigma} \quad . \tag{2.11}$$

The first assumption, given by (2.10), is frequently encountered in Bayesian work. Essentially, it implies that the prior information at the beginning of period one can be thought as equivalent to the information obtained from a sample of size n_1' from the process. Assumption (2.11) implies that the random shocks that change the mean vector from period to period are such that they do not change the underlying relationship among the elements of the mean vector.

With the inclusion of assumptions (2.10) and (2.11), (2.8) and (2.9) can be expressed in more simplified form:

$$\underset{\sim}{m}_{t+1}' = (n_t' + n_t)^{-1}(n_t'\underset{\sim}{m}_t' + n_t\underset{\sim}{m}_t) + \underset{\sim}{e} \tag{2.12}$$

and

$$\underset{\sim}{S}_{t+1}' = \left[(n_t' + n_t)^{-1} + w^{-1}\right] \underset{\sim}{\Sigma} \quad , \tag{2.13}$$

for $t = 1,2,\ldots$. Moreover, if n_{t+1}' is defined as

$$n_{t+1}' = \left[(n_t' + n_t)^{-1} + w^{-1}\right]^{-1} \quad , \tag{2.14}$$

then

$$\underset{\sim}{S}'_{t+1} = (n'_{t+1})^{-1} \underset{\sim}{\Sigma} \quad . \tag{2.15}$$

From (2.15), it is apparent that the limiting behavior of $\underset{\sim}{S}'_t$ can be studied by investigating the limiting behavior of n'_t. Looking at the special case in which the sample size is the same each period, we can find a limit for n_t.

Theorem 3. If n'_{t+1} is defined as in (2.14) and if $n_t = n$ for $t = 1,2,\ldots$, where n'_1, n, and w are all strictly positive, then

$$\lim_{t\to\infty} n'_t = \frac{n}{2}\left[\left(1 + \frac{4w}{n}\right)^{\frac{1}{2}} - 1\right] \quad . \tag{2.16}$$

Proof. First, if a limit n_L exists, it must satisfy

$$n_L = \left[(n_L + n)^{-1} + w^{-1}\right]^{-1} = \frac{(n_L + n)\, w}{n_L + n + w} \quad ,$$

which simplifies to

$$n_L^2 + nn_L - nw = 0 \quad . \tag{2.17}$$

This equation has exactly one positive root, which is

$$n_L = \frac{-n + (n^2 + 4nw)^{\frac{1}{2}}}{2} = \frac{n}{2}\left[\left(1 + \frac{4w}{n}\right)^{\frac{1}{2}} - 1\right] \quad . \tag{2.18}$$

This is the limit postulated in (2.16); does the sequence $\{n_t'\}$ converge to n_L? Consider

$$n_{t+1}' - n_L = \left[\frac{(n_t' - n)\ w}{n_t' + n + w}\right] - n_L = \frac{(n_t' - n)\ w - n_L(n_t' + n + w)}{n_t' + n + w}$$

$$= \frac{(w - n_L)(n_t' - n_L) - (n_L^2 + nn_L - nw)}{n_t + n + w} \quad .$$

From (2.17), the last term in the numerator is zero, so

$$n_{t+1}' - n_L = \left(\frac{w - n_L}{n_t' + n + w}\right)(n_t' - n_L) \quad .$$

Also, (2.17) implies that $n_L^2 = n(w - n_L)$. Thus,

$$w - n_L > 0 \quad \text{and} \quad [(w - n_L)/(n_t' + n + w)] > 0 \quad ,$$

implying that

$$|n_{t+1}' - n_L| = \left(\frac{w - n_L}{n_t' + n + w}\right)(n_t' - n_L) \leq \left(\frac{w - n_L}{n + w}\right)(n_t' - n_L)$$

Therefore,

$$|n_{t+1}' - n_L| \leq \left(\frac{w - n_L}{w + n}\right)^t |n_1' - n_L| \quad , \qquad (2.19)$$

where it is clear that

$$0 < \left(\frac{w - n_L}{w + n}\right) < 1 \quad .$$

Thus,

$$\lim_{t\to\infty} n_t' = n_L \quad . \tag{2.20}$$

An immediate corollary of Theorem 3 is that under the conditions of the theorem,

$$\lim_{t\to\infty} \underset{\sim}{S}_t' = n_L^{-1}\underset{\sim}{\Sigma} \quad , \tag{2.21}$$

where n_L is given by (2.18). This result follows directly from (2.15) and (2.20). Moreover, it is possible to contrast this result with that of the stationary case. The stationary case can be thought of as a limiting form of the nonstationary case with $\underset{\sim}{e} = 0$ and $w^{-1} = 0$. Thus, from (2.14), we have, for the stationary case,

$$n_{t+1}' = n_t' + n_t \quad .$$

Therefore, assuming that n_t is a positive integer for all t, n_t' increases without bound as t increases, so that, from (2.15), $\underset{\sim}{S}_t'$ approaches a matrix of zeros as t increases. Intuitively, in the stationary case, the distribution of the unknown parameters becomes tighter as we obtain more information. In the nonstationary case, $n_{t+1}' < n_t' + n_t$ because of the additional uncertainty involving the shifts in the mean vector, and the distribution does not necessarily become tighter as t increases. In fact, if n_1', the initial value of n_t', is larger than n_L, the elements of $\underset{\sim}{S}_t'$ will <u>increase</u> as t increases. In this case, initially there is a great deal of information concerning $\underset{\sim}{\tilde{\mu}}_1$. Even though the observations in the first period yield yet further information concerning $\underset{\sim}{\tilde{\mu}}_1$, the random shock at the end of the period is strong enough to imply that there is less information about $\underset{\sim}{\tilde{\mu}}_2$ at the beginning of the second period than there was about $\underset{\sim}{\tilde{\mu}}_1$ at the beginning of the first period. On the other hand, if n_1' is less than n_L, then the information obtained each period "overrides" the uncertainty caused by the random shock, in a sense, and there is more information about $\underset{\sim}{\tilde{\mu}}_2$ at the beginning of the second period than there was about $\underset{\sim}{\tilde{\mu}}_1$ at the beginning of the first period.

Next, we will investigate the behavior of the sequence $\{m_t'\}$. Without loss of generality, assume that $\underset{\sim}{e} = 0$. Then, from (2.12), $\underset{\sim}{m}_{t+1}'$ can be expressed in the form

$$\underset{\sim}{m}_{t+1}' = q_t \underset{\sim}{m}_t' + (1 - q_t)\, \underset{\sim}{m}_t \quad , \tag{2.22}$$

where

$$q_t = n_t'/(n_t' + n_t) \quad . \tag{2.23}$$

Successively applying (2.22) gives $\underset{\sim}{m}_{t+1}'$ as a function of $\underset{\sim}{m}_1'$, the initial mean, and $\underset{\sim}{m}_i$ and q_i for $i = 1,\ldots,t$

$$\underset{\sim}{m}_{t+1}' = \left(\prod_{i=1}^{t} q_i\right) \underset{\sim}{m}_1' + \sum_{j=1}^{t-1} (1 - q_j) \left(\prod_{i=j+1}^{t} q_i\right) \underset{\sim}{m}_j + (1 - q_t)\, \underset{\sim}{m}_t \quad . \tag{2.24}$$

<u>Theorem 4</u>. Under the conditions of Theorems 2 and 3, and with $\underset{\sim}{e} = 0$ and $n_1' = n_L$,

$$\underset{\sim}{m}_{t+1}' = (1 - q) \sum_{i=0}^{t-1} q^i \underset{\sim}{m}_{t-i} + q^t \underset{\sim}{m}_1' \quad , \tag{2.25}$$

where

$$q = n_L/(n_L + n) \quad . \tag{2.26}$$

<u>Proof</u>. From (2.19), $n_1' = n_L$ implies that $n_t' = n_L$ for $t = 2,3,\ldots$; one the process reaches the limit n_L, it remains there. Also, in Theorem 3, it was assumed that $n_t = n$ for all t. Thus, from (2.23), we have

$$q_t = n_L/(n_L + n) = q \quad \text{for all } t.$$

On substituting q for each q_i, $i = 1,\ldots,t$, in (2.24), we get (2.25).

Under the assumptions of Theorem 4, the prior mean vector at the beginning of any period can be expressed as a sum of 1) the initial prior mean vector $\underset{\sim}{m}_1'$, suitably discounted by a factor of q^t and 2) an exponentially weighted sum of the observed sample means. This result seems intuitively appealing; recent observations are weighted more heavily than not-so-recent observations. Observations from a process with a mean that is only "one shock removed" from the current mean receive a weight of (1-q), whereas observations from a process with a mean that is, say, "i shocks removed," receive a weight of $(1-q)q^{i-1}$. Since $0 < q < 1$, the impact of a particular sample mean on future values of $\underset{\sim}{m}_t'$ decreases as t increases.

Theorem 4 utilizes one assumption not previously used: the assumption that $n_1' = n_L$. This assumption implies that at the beginning of the first period, the model is already in steady-state form in the sense that the sequence of variances $\underset{\sim}{S}_t'$ will be a constant sequence. As long as $|n_1' - n_L|$ is not too large, (2.25) will provide a good approximation to the behavior of the sequence $\underset{\sim}{m}_t'$. Furthermore, in any event the approximation will improve as t increases.

3. Application to Portfolio Analysis

One potential area of application of the model discussed in the previous section is portfolio analysis. In portfolio analysis, the process of interest is the process generating changes in security prices, and the decision making objective is to determine an "optimal" portfolio of securities. In Winkler [10], a Bayesian model for forecasting future security prices under the assumption of stationarity is presented, and this model is used in Winkler and Barry [11] in the determination of portfolio selection and revision policies that are optimal in the sense that they maximize the expected utility of the decision maker's wealth at some prespecified future time (i.e. end-of-horizon wealth). In this section, we will sketch briefly the application of the model of Section 2 to allow the determination of optimal portfolios under nonstationarity.

Using the notation of Section 2 suppose that M securities are under consideration for inclusion in the portfolio, and one observation of prices will be made each period, so that $n_t = 1$ for $t = 1,2,\ldots$. The variable of interest in period t is $\tilde{x}_t$ (since $n_t = 1$, we drop the second subscript for convenience), the vector of log price changes of the M securities, which has a multinormal distribution with mean $\tilde{\mu}_t$ and covariance matrix Σ. The process generating successive values of the mean vector at the beginning of period one are just as in Section 2. At the end of time period t, a_t denotes the vector of holdings (in dollars) of the M securities, and the decision maker's wealth at this time is simply $W_t = 1^T a_t$, where 1 is a vector of ones.

A convenient assumption is that the time periods under consideration are short enough that the log price changes are unlikely to differ from zero by a substantial amount. Under this assumption, $\tilde{x}_t$ provides a good approximation to the vector of rates of return, and we will treat $\tilde{x}_t$ as if it were a vector of rates of return. Then the wealth at the end of period t can be written in the form

$$W_t = (1 + x_t)^T (a_{t-1} + p_{t-1} - q_{t-1}) \quad ,$$

where p_{t-1} and q_{t-1} are vectors of the amounts bought and sold, respectively, of the M securities at the end of time period t-1.

To keep matters simple, we will consider only a single-period model, which is a model in which the decision maker's time horizon is always only one period into the future. Thus, at the end of period t-1, the decision maker wants to choose $\underset{\sim}{p}_{t-1}$ and $\underset{\sim}{q}_{t-1}$ to maximize

$$EU(\tilde{W}_t) = EU\left[(1 + \underset{\sim}{\tilde{x}}_t)^T(\underset{\sim}{a}_{t-1} + \underset{\sim}{p}_{t-1} - \underset{\sim}{q}_{t-1})\right] \quad ,$$

subject to

$$\underset{\sim}{1}^T\underset{\sim}{p}_{t-1} = \left(\frac{1 - c}{1 + c}\right)\underset{\sim}{1}^T\underset{\sim}{q}_{t-1} \quad ,$$

$$0 \leq \underset{\sim}{q}_t \leq \underset{\sim}{a}_t \quad ,$$

and

$$\underset{\sim}{p}_t \geq 0 \quad ,$$

where the vector inequalities imply that the inequality holds for each pair of corresponding elements of the vectors, c represents a constant per-unit transaction cost (for both buying and selling), and U represents the decision maker's utility function for W_t. The first constraint reflects the effect of transaction costs, the second constraint prohibits short selling, and the second and third constraints are simply non-negativity constraints.

The uncertainty in the portfolio analysis problem involves $\underset{\sim}{\tilde{x}}_t$. Given some assumptions about the data-generating process and given prior distributions for the underlying parameters of the process, it is possible to determine the distribution of $\underset{\sim}{\tilde{x}}_t$, which is called predictive distribution. For the nonstationary model of Section 2 with $\underset{\sim}{e} = 0$ and with the simplifying assumptions (2.10) and (2.11), the predictive distribution of $\tilde{x}_t$ at the end of time period t-1 is a multinormal distribution with mean $m''_{t-1} = m'_t$ and covariance matrix $[n'_t + n_t)/n'_t]\Sigma$.

Given the predictive distribution of $\tilde{x}_t$, it is easy to compare the stationary and nonstationary models. Suppose that at the beginning of time period t-1, the prior distribution of $\tilde{\mu}_{t-1}$ is the same for the two models. Using (2.14), we have

$$n'_t = [(n'_{t-1} + n_{t-1})^{-1} + w^{-1}]^{-1} \quad .$$

But the stationary model can be thought of as a limiting form of the nonstationary model with $w^{-1} = 0$, so n'_t will be larger for the stationary model than for the nonstationary model. Hence, the elements of the covariance matrix of $\tilde{x}_t$ will be smaller in absolute value for the stationary model.

Given U, one can solve for the optimal portfolio revision at the end of time period t-1. For instance, if U is quadratic, the problem is a quadratic programming problem. For quadratic and exponential utility functions, the optimal solution is found in Barry [1]. Moreover, this solution is compared with the optimal solution to the corresponding stationary model. For a situation with one risky security and one risk-free security, it is found that, all other things being equal, the decision maker using the nonstationary model will hold an amount of the risk-free security greater than or equal to the amount held under the stationary model. This seems intuitively reasonable, since the utility functions imply risk aversion and there is additional uncertainty concerning the mean return from the risky asset in the nonstationary model. The case of two risky securities is also investigated (with similar results) in Barry [1].

The single-period portfolio models allow for the revision of probability distributions and portfolios as new information is received, but they do not take into account the dynamic nature of the portfolio analysis problem. A multiperiod model that does consider the dynamic nature of the situation has been studied in some detail under stationarity, and a nonstationary multiperiod model should also be of considerable interest. For instance, it should be useful to compare the steady-state behavior of the nonstationary multiperiod model (where $n_t' = n_L$, so that n_{t+1}', $n_{t+2}',\ldots$, are all equal to n_L) with the behavior of the stationary multiperiod model (where n_{t+1}, $n_{t+2},\ldots$ form a strictly increasing sequence).

4. Summary and Discussion

In Section 2 a Bayesian model for dealing with a multinormal process with a nonstationary mean vector was discussed. When the model is expressed in its most general form, it appears difficult to make broad statements about the limiting behavior of the model, although formulas for revising the distributions of interest can readily be obtained. With some simplifying assumptions, primarily concerning the structure of the covariance matrices included in the model, it is possible to determine the limiting covariance matrix of $\tilde{\mu}_t$. Unlike the stationary case, the limiting covariance matrix is nonzero, because even though more information is obtained in each period, the mean vector is also shifting stochastically in each period, so uncertainty remains about the value of this mean vector.

Various extensions of the model in Section 2 could be considered. It was assumed that Σ, the covariance matrix of the data-generating process, was known, and this assumption could be relaxed by assessing a joint prior distribution for $\tilde{\mu}_1$ and $\tilde{\Sigma}$ at the beginning of period one and revising this distribution as new information is obtained. If this joint prior distribution is Normal-inverted-Wishart, the extension from the case of known Σ is simple to handle. Similarly, it could be assumed that e and Ω are unknown, although the model could become quite cumbersome if all parameters are assumed unknown. Another possible extension is to assume that the shocks that shift the mean occur stochastically instead of regularly at the beginning of each time period. For example, the shocks might be assumed to be generated by a Poisson process. Carter [5] considered this type of extension for the univariate situation studied by Bather [2].

In Section 3 a very brief outline of the application of the nonstationary model to portfolio analysis was presented.

In the context of Bayesian models of security price changes, a nonstationary model seems more realistic than a stationary model (e.g. it seems reasonable for the variances not to approach zero). For a simple single-period model, the nonstationary model of this paper is compared with a stationary model in Barry [1], and the results indicate that nonstationarity causes some changes in the optimal portfolios. In view of the apparent applicability of nonstationary models in portfolio analysis and in other situations, further work regarding such models seems warranted.

References

[1] Barry, C.B., "A Bayesian Approach to Portfolio Analysis," unpublished doctoral dissertation, Indiana University, 1973.

[2] Bather, J.A., "Control Charts and the Minimization of Costs," Journal of the Royal Statistical Society B, 25 (1963) 49-70.

[3] Bather, J.A. "Invariant Conditional Distributions," Annals of Mathematical Statistics, 36 (1965), 829-846.

[4] Boness, A.J., Chen, A.H., and Jatusipitak, S., "A Hypothesis of Non-Stationary Common Share Price Changes," unpublished manuscript, State University of New York at Buffalo, 1972.

[5] Carter, P.L., "A Bayesian Approach to Quality Control," Management Science, 18 (1972), 647-655

[6] DeGroot, M.H., Optimal Statistical Decisions. New York, McGraw-Hill, 1970.

[7] Hsu, D.-A., Miller, R.B., and Wichern, D.W., "On the Stable Paretian Behavior of Stock Market Prices," Technical Report No. 299, University of Wisconsin, 1972.

[8] Johnson, N.L., and Kotz, S., Distributions in Statistics: Continuous Multivariate Distributions. New York, Wiley, 1972.

[9] Raiffa, H., and Schlaifer, R., Applied Statistical Decision Theory. Boston, Graduate School of Business Administration, Harvard University, 1961.

[10] Winkler, R.L., "Bayesian Models for Forecasting Future Security Prices," Journal of Financial and Quantitative Analysis, 8 (1973), 387-405.

[11] Winkler, R.L., and Barry, C.B., "A Bayesian Model for Portfolio Selection and Revision," unpublished manuscript, Indiana University and University of Florida, 1973.