DRUG DEVELOPMENT AND INDUSTRIAL PHARMACY, 18(11&12), 1281-1309 (1992)

APPLICATION OF KALMAN FILTER STATISTICAL TECHNIQUE IN PHARMACEUTICS

N. R. Bohidar

Philadelphia College of Pharmacy and Science
and
Villanova University

ABSTRACT

Pharmaceutics experiments involving time-dependent or space-dependent observations are generally subjected to single or multiple regression analysis based on the tacit assumption that the regression coefficients are fixed and invariant through time or across space. However, this assumption is invariably not tenable, specially for this type of observations, and the coefficients do indeed vary during the experiment through time or across space. The results of the analysis which does not take this aspect into consideration may be inaccurate. For instance, the results of the standard least squares analysis applied to a multicomponent mixture experiment to determine the concentrations of individual components will be absolutely misleading in the presence of parametric variation. Kalman filter statistical technique is devised to handle such complex problems. A detailed description of the technique is presented in the context of a specific pharmaceutic experiment. For the purpose of data analysis, one has the choice of applying one of the three models, fixed parameter model (standard least

squares), random coefficient model, or variable (stochastic) parameter model. The unique feature of this presentation is the introduction of the three statistical tests of validity which, upon application, can delineate the single model appropriate for the analysis of the available data. Finally, for the purpose of illustration, the numerical results of the statistical analysis of two multicomponent mixture experiments are presented and appropriately interpreted.

INTRODUCTION

The statistical analysis presented in this paper pertains primarily to the pharmaceutics experiments in which the responses are measured over time or across space, such as, stability, solubility, kinetics, dissolution and production process control. Without loss of generality, the description of the methodology is accomplished in the context of a specific pharmaceutics experiment. The presentation depicts, in detail, the statistical analysis of multi-component mixture experiments in which the responses are measured across space, represented by wavelengths, in this case. The determination of the concentration of each individual component in the mixture is of prime interest here. In this context, the data consists of the spectrophotometric absorbance values of the mixture as well as of each of the K components measured across the same N equally spaced wavelengths selected for the study. The data matrix consisting of N rows and K columns is routinely subjected to a multiple regression analysis (1), treating the mixture as the dependent variable and the standards of the K components as the independent variables. The magnitudes of the partial regression coefficients obtained from the least squares statistical basis of this algorithm will be presented here. Futhermore, the original Kalman filter formulations will be extended to derive additional statistical formulas to address some of the practical pharmaceutics objectives. In addition to these extensions, the presentation will also include a description of the three major statistical tests of hypothesis (validity) to be used in the analysis for establishing the presence or absence of one or more of the factors, AC, HS and SPV, in the system, because in the presence of one or more of these factors, the analysis will be much more sophisticated statistically and computationally than the standard multiple regression analysis.

The presentation here will consist of (a) a step-by-step description of the Kalman filter formulation, (b) an extension of the original Kalman filter algorithm for deriving additional statistical formulas, (c) the procedure for estimating all the parameters (including the standard errors) from the available sample data (no parameter will be considered fixed or known), (d) a description of the three statistical tests of validity and (e) the salient results of the statistical analysis of two two-component mixture experiments, one involving Product-BP, and the other Product-MP.

THEORY

In this section a detailed account of the theoretical aspects of the development is presented in the following six separate parts: Part I: Fixed (invariant) parameter regression model, Part II: Stochastic (variable) parameter regression model, Part III: Statistical basis of Kalman filter algorithm, Part IV: Computational aspects of Kalman filter algorithm, Part V: Likelihood function formulation and the maximum analysis represent the respective concentrations of the K components in the mixture (1).

It is clearly recognized, however, that time-dependent or space-dependent data generally tend to exhibit autocorrelation, in the sense that the data values of adjacent time periods or space units are somewhat similar and therefore positively correlated with one another. This phenomenon is known as autocorrelation. It is also conceivable that the signal-to-noise ratio at each wavelength may not necessarily be the same for the entire spectrum. This phenomenon is called heteroscedasticity. Futhermore, implicit in the standard multiple regression analysis is a tacit assumption that the partial regression coefficients remain constant across space or through time (1). However, it would be realistic to consider the possibility that the coefficients do indeed vary during the experimental process due to inherent physico-chemical changes in the system. The analysis which takes this parametric variation into consideration is called "Variable Parameter Regression Analysis" or "Stochastic Parameter Regression Analysis."

It must be noted here that the results of the multiple regression analysis which does not take into consideration all these factors, autocorrelation (AC), heteroscedasticity (HS) and stochastic parameter variation (SPV) in the formulation, will be seriously misleading. On the other hand, when these conditions, AC, HS and SPV, are incorporated into the general least squares framework, one generates a formidable expression which is not amenable to simple and interpretable solutions. The Kalman filter statistical technique (2,3) was devised to provide an unique set of procedures (algorithm) for the solution of such a complex expression. A step-by-step description of the likelihood estimation procedure, and Part VI: Statistical tests of validity for model adequacy.

Part I : Fixed (invariant) Parameter Regression Model (1)

The multiple regression model which is most commonly employed for the statistical analysis of multi-component mixture experiments is explicitly expressed as, $Y_w = C_0X_{0w} + C_1X_{1w} + C_2X_{2w} +$ ---- $C_kX_{kw} + E_w$ where, C_1, C_2, ----C_k are the unknown coefficients (concentrations) to be estimated from the available data, Y_w denotes the measured absorbances of the mixture at wavelength w; X_{1w}, X_{2w}, ---X_{kw} denote the measured absorbances of K standards of components (solutes) at wavelength w, E_w represents the random error associated with Y_w, w = 1,2,---N (N>K), C_0 denotes the Y-intercept since X_{0w} takes the value of one at each wavelength and the E_w's are not correlated with each other. The model depicted above implicitly assumes that the C-coefficients are invariant (fixed) over the entire spectrum. Based on this assumption, the standard least squares estimation procedure provides the estimators of the coefficients, expressed in vector-matrix notation, as $C^* = (X'X)^{-1}X'Y$ derived from the model $Y = XC + E$ in which,

$R^2 = [C^{*\prime}X'Y - R(C_0)]/[Y'Y - R(C_0)]$, $R(C_0) = N\bar{Y}^2$, and

$R_w = (Y_w - Y^*_w) = (Y_w - \Sigma C^*_i X_{iw})$, the residual from the regression at wavelength w. These residual quantities will be used in the computation of the test statistics for the statistical tests of validity, presented later in this section.

Part II: Stochastic (Variable) Parameter Regression Model

The objective here, however, is to consider a different but not unrelated elaboration of the standard multiple regression model and its usual assumptions by incorporating in the model the expressions for the stochastic (variable) regression coefficients, as follows: $Y_w = C_{0w}X_{0w} + C_{1w}X_{1w} + C_{2w}X_{2w} + ----+C_{kw}X_{kw} + E_w$. Now each coefficient is identified by two subscripts, one for the component and one for the wavelength. Note that at wavelength w, the coefficient of the 2nd component is C_{2w}, indicating that at some other point in space a different coefficient (say, C_{2g}) might be appropriate. Since the dependent variable (Y_w) as well as the coefficients (C_w) are random variables, one would need two interconnecting models to depict the process, as follows (4,5):

$$Y_w = X'_w C_w + E_w \qquad \text{EQ - 1}$$

and

$$C_w = C + \Theta(C_{w-1} - C) + A_w \qquad \text{EQ - 2}$$

where, $X'_w = (X_{0w}, X_{1w}, ----X_{kw})$, $C_w = (C_{0w}, C_{1w}--- C_{kw})$, the C_{iw}'s are space-varying parameters autocorrelated with each other across space, the E_w's are random errors assumed to have zero mean, fixed variance σ^2 across space and to be non-autocorrelated, $C' = (C_1, C_2 ---C_k)$ the long range mean of C_w across space, $E(C_{1w}) = C_1, E(C_{2w}) = C_2$, ---- $E(C_{kw}) = C_k$, Θ is a KXK matrix of fixed parameters, denoting the correlation between two consecutive coefficients, and the A_w's are random errors with zero mean and fixed covariance matrix Ω across space and are assumed to be non-autocorrelated with each other and to be uncorrelated with errors in EQ - 1. It may be noted that, EQ - 2 establishes a functional relationship between the coefficient at wavelength w and the coefficient at wavelength (w - 1) via the parameter Θ, so that C_w is predictable from C_{w-1}. EQ - 2 depicts a model which is generally known as the first order autoregressive process. If, however, Θ takes the value of zero, then C_w would randomly fluctuate around the central value C in an unpredictable manner, and EQ -2 would change to $C_w = C + A_w$, called the random coefficient model. It should be noted here that the fixed parameters of the models, C, σ^2, Θ and Ω are called the stationary parameters, since they pertain to all the space units in the spectrum and consequently they do not bear the subscript, w.

The estimation of the various parameters associated with EQ -1 and EQ -2 can not be accomplished either by the standard or the generalized version of the least squares procedure because of the intricately complex (AC, HS, SPV) nature of the error structure as well as of the parameters. The two interconnected models in EQ -1 and EQ - 2, however, fully define the proposed stochastic parameter regression model which is a special case of a general class of models known as state space models (2,3), used quite commonly in engineering. In engineering terminology, EQ -1 is called the measurement equation (2,3) since it relates the observed values of the dependent variable to the independent variables through the unobservable coefficients, C_W and EQ - 2 is called the transition equation (2,3) because it describes the evolution of these unobservable parameters through space.

The interest here will be to apply, for the purpose of parameter estimation and prediction, the state space model and to employ a technique known as the Kalman filter algorithm (KFA) which provides a computationally efficient framework for solving the problem of estimation of the stationary parameters of the model and of prediction of future values. To be specific, KFA is used to solve the following three problems: (1)estimation of the stationary parameters: mean vector C, variance of the error term σ^2 in EQ - 1, variance-covariance matrix Ω of the error term in EQ - 2, and also of the autoregressive parameter matrix Θ in that model. Since the least squares estimation procedure is not applicable here, the well known statistical estimation procedure called the maximum likelihood method (MLM) will be used, and KFA will serve as a computationally efficient means of obtaining the likelihood function (LF) in terms of the stationary parameters of the model. (Note that MLM and LF will be considered in detail later in this section), (ii) estimation of C_w at wavelength w given the observations, Y_1, Y_2, ----Y_{w-1} only, [to be denoted by the notation C(w/w-1)], (iii) estimation of C_w at wavelength w given observations, Y_1, Y_2, ---, Y_w, [to be denoted by notation C(w/w)] and (iv) estimation of the coefficient C_w at each w (w = 1,2, ---- N) based on the observations, Y_1, Y_2 ----Y_N, the entire spectrum [to be denoted by the notation C(W/N)]. It should be noted here that KFA (2,3) has the recursive property in that, it utilizes the past information to generate the future values in discreet steps.

<u>Part III: Statistical Basis of Kalman Filter Algorithm (KFA)</u>

The stochastic parameter models in EQ - 1 and EQ-2 will form the basis for the derivation of the expected value (long-range mean, E) and the variance (V) of the parameters C_w under the assumption that,the error terms in these two models are normally (Gaussian) distributed. The objective of the KFA is then to provide a convenient way of computing the expected values and variances of the parameters given the information available up to space unit W where, W = 1,2,---N. However, prior to reaching that point, one has to consider the following derivations.

A. Distribution of C_w and Y_w given the information up to space unit W - 1, [C(w/w-1), Y (w/w-1)]:

The first step will be to find the joint distribution of C_w and Y_w given the observations Y_1,

Y_2,----Y_{w-1}. Now, taking expectations, conditioned on the available information, EQ-2 yields,

$E(C_w/Y_1,Y_2,---Y_{w-1}) = C + \theta E(C_{w-1}/Y_1, Y_2,---Y_{w-1}) - \theta C + E(A_w)$.

Since the last term on the right hand side of the above equation is zero, one has,

$$C(W/W-1) = \theta C(W-1/W-1) + (I - \theta)C \qquad \text{EQ - 3}$$

where I is a K X K identity matrix.

It further follows from EQ - 2 that the variance-covariance matrix of C_w, given information available up to space unit W - 1, can be expressed as,

$Var(C_w/Y_1,Y_2,---Y_{w-1}) = \theta V(C_{w-1}/Y_1,Y_2,---Y_{w-1})\theta' + \Omega$

where Ω is the variance-covariance matrix of A_w, and Var stands for the variance operator.

Expressed in short notation,

$$V(W/W-1) = \theta V(W-1/w-1)\theta' + \Omega \qquad \text{EQ - 4}$$

The mean of the distribution of Y_w, given information up to space unit W - 1 is obtained by taking conditional expectations in EQ - 1 so that,

$E(Y_w/Y_1,Y_2,---Y_{w-1}) = X'_w E(C_w/Y_1,Y_2,---Y_{w-1}) + E(E_w/Y_1,Y_2,---Y_{w-1})$

Since the last term on the right hand side is zero, the above expression reduces to

$$Y(W/W-1) = X'_w C(W/W-1) \qquad \text{EQ - 5}$$

The expression for the variance of Y_w, given information available up to space unit W-1 is derived from EQ - 1 as follows:

$Var(Y_w/Y_1,Y_2,---Y_{w-1}) = X'_w V(C_w/Y_1,Y_2,---Y_{w-1})X_w + \sigma^2$

where σ^2 is the variance of E_w. In short notation,

$$G_w = G(w/w-1) = X'_w V(w/w-1)X_w + \sigma^2 \qquad \text{EQ - 6}$$

Finally, to complete the derivation of the joint conditional distribution, the covariance between C_w and Y_w given information available up to space unit w-1, as derived from EQ-1 is,

$COV(C_w, Y_w/Y_1, Y_2, --- Y_{w-1}) = Var(C_w/Y_1, Y_2, --- Y_{w-1})X_w = V(w/w-1)X_w$ EQ - 7

where COV stands for the covariance operator.

B. Distribution of C_w and Y_w given the information available up to space unit w, [C(w/w) and Y(w/w)]:

The results of this derivation constitute the central formulation of the Kalman filter algorithm which is primarily based on some of the well known standard statistical properties of the multivariate Gaussian distribution(6). At first the results will be presented in general terms and then they will be expressed in specific terms.

Let Z_1 and Z_2 be a pair of vectors of random variables, whose joint distribution is multivariate Gaussian.

Let $E(Z_1) = \mu_1$, $V(Z_1) = E[(Z_1 - \mu_1)(Z_1 - \mu_1)'] = \Sigma_{11}$

Let $E(Z_2) = \mu_2$, $V(Z_2) = E[(Z_2 - \mu_2)(Z_2 - \mu_2)'] = \Sigma_{22}$

The variance-covariance matrix is denoted by

$$\Sigma_{12} = E[(Z_1 - \mu_1)(Z_2 - \mu_2)']$$

The joint Gaussian distribution can be expressed compactly as,

$$[Z_1, Z_2] \approx N\ [\mu_1, \mu_2], \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \qquad \text{EQ - 8}$$

where N stands for normal or Gaussian distribution.

It is well known from the properties of the multivariate Gaussian distribution that the distribution of Z_1 given Z_2 is also multivariate Gaussian with mean,

$$E(Z_1/Z_2) = \mu_1 + \Sigma_{12}\Sigma^{-1}{}_{22}(Z_2 - \mu_2) \qquad \text{EQ - 9}$$

and the variance-covariance matrix,

$$V(Z_1/Z_2) = \Sigma_{11} - \Sigma_{12}\Sigma^{-1}{}_{22}\Sigma_{21} \qquad \text{EQ - 10}$$

Consider now the joint distribution of $C_w = Z_1$ and $Y_w = Z_2$ given information available up to space unit W-1, that is, given Y_1, Y_2, --- Y_{w-1}. This distribution is multivariate Gaussian, and, using the equations, EQ-3 - EQ-7, it can be expressed in the format given in EQ-8,

$\mu_1 = \theta C(w-1/w-1) + (I - \theta)C$, $\mu_2 = X'_w C(w/w-1)$

$\Sigma_{11} = V(w/w-1)$, $\Sigma_{22} = G_w = G(w/w-1)$,

$\Sigma_{12} = V(w/w-1)X_w \quad \Sigma_{21} = X'_w V(w/w-1)$

Now, after the appropriate substitutions into EQ -9 and EQ - 10, the required expressions are,

$$C(w/w) = \theta C(w-1/w-1) + (I - \theta)C + V(w/w-1)X_w G_w^{-1}[Y_w - X'_w C(w/w-1)] \qquad \text{EQ - 11}$$

and (from EQ - 10 its covariance matrix is)

$$V(w/w) = V(w/w-1) - V(w/w-1)X_w G_w^{-1} X'_w V(w/w-1) \qquad \text{EQ - 12}$$

C. Distribution of C_w and Y_w given <u>all</u> the information available up to space unit N, [C(w/N) and Y(w/N)]:

So far, the appropriate expressions for C(w/w-1), Y(w/w-1), C(w/w) and Y(w/w), and their respective variances have been presented(see EQ-3 - EQ-6, EQ-11, EQ-12). The interest here, however, is to derive appropriate expressions for the stochastic parameters $C_w(w = 1,2,---N)$ based on <u>all</u> the available sample information. The conditional expectation of C_w given the entire sample data is $C(w/N) = E(C_w/Y_1,Y_2,---Y_N)$ and its variance is $V(w/N) = VAR(C_w/Y_1,Y_2,---Y_N)$.

It should be noted here, that the Kalman filter algorithm does not provide the estimates needed here. Rather, that algorithm only generates as functions of the stationary parameters the mean and variance of the stochastic parameter C_w, <u>given only information available up to space unit w.</u> Therefore, the only estimates available up to this point that utilize all the available sample information are C(N/N) and V(N/N), the mean and variance-covariance matrix of the stochastic parameter vector at the last point in the spectrum. To accomplish the stated objective the derivation takes advantage of what is called the fixed interval smoothing algorithm(7). Since the intermediate steps leading to the final results are rather complex and involved, only the final formulas of the conditional mean and variance-covariance matrices of the parameters are presented here, as follows:

$C(w/N) = C(w/w) + H_w[C(w+1/N) - C(w+1/w)]$ and

$V(w/N) = V(w/w) - H_w[V(w+1/w) - V(w+1/N)]H'_w$

where, $H_w = V(w/w)\ \theta' V(w+1/w)^{-1}$

It is appropriate at this point to suggest a computational strategy for the calculation of the above expressions. One should start with w = N-1 and carry out the calculation by setting in turn w = N-2, N-3, N-4--- to obtain point estimates C(w/N) of all the stochastic parameters over the sample spectrum, together with the associated variance-covariance matrices V(w/N). For example, let w = N-1, then C(N-1/N) = C(N-1/N-1 + H_w[C(N/N - C(N/N-1)] where, $H_w = V(N-1/N-1)\theta' V(N/N-1)^{-1}$. Note that the formulas for the terms on the right hand side of the above expressions are already available from the previous derivations given in A and B of Part III. To carry out the calculations of the fixed interval smoothing algorithm equations given above, in principle, it is necessary to know the values of the stationary parameters, C, σ^2, θ, Ω of the model. However, in practice, these unknown true values will be replaced by their maximum likelihood estimates.

<u>Part IV: Computational Aspects of Kalman Filter Algorithm:</u>

At this point it would be appropriate to outline briefly the computational strategies for obtaining the conditional means and variances of the stochastic parameters C_w, <u>for any given values</u> of the space invariant stationary parameters, C, σ^2, θ and Ω of the state space model. The computations are started off by setting w = 1 in EQ-3 - EQ-6, EQ-11, and EQ-12, in turn, to find C(1/0), V(1/0), G_1, C(1/1), V(1/1), C(1/N), V(1/N). Next, using the these results, the computation proceeds by setting w=2 in the same set of equations given above, to find $C(2/1)$, $V(2/1)$, G_2, $C(2/2)$, $V(2/2)$, $C(2/N)$, $V(2/N)$. One can proceed iteratively in this manner, setting in turn, $w = 1,2,$ ---N in those sets of equations given above. The quantities required to perform the calculations at any one state will simply be the results found at the previous stage. For instance, when $w = 1$, $C(1/0)=\Theta C(0/0)+(I - \Theta)C$, $V(1/0) = \Theta V(0/0)\Theta' + \Omega$, $C(1/1) = \Theta C(0/0) + (I - \Theta)C + V(1/0)X_1G_1^{-1}[Y_1-X_1'C(1/0)]$ and $V(1/1) = V(1/0) - V(1/0)X_1G_1^{-1}V(1/0)$. Now, if one substitutes the right hand expression of $C(1/0)$ and $V(1/0)$ into $c(1/1)$ and $V(1/1)$ respectively, it would clearly be observed that the conditional means and variances are indeed the functions of $C(0/0)$ and $V(0/0)$, the initial values, and the space invariant stationary parameters. To compute these two initial values one simply requires the expected value (mean) and variance of EQ -1 given no observations. The expected value of $C_W = C + \Theta(C_{W-1} - C) + A_W$ is actually C by definition and so $C(0/0) = C$. The variance of C_W is $V(0/0) = \Theta V(0/0)\Theta' + \Omega$. For instance, when $k = 1$ (single component), $C(0/0) = C$ and $V(0/0) = \Theta^2\sigma^2c + \sigma^2a$. The above derivation clearly indicates that the estimation of the stationary parameters is extremely critical to the computation of the conditional means and variances. The statistical procedure known as the maximum likelihood estimation method will be used for the purpose.

Part V: Likelihood Function Formulation and Maximum Likelihood Estimation Procedure(8,9,10):

Given a set of observations measured across N wavelengths (space units), the interest here is to estimate the stationary parameters, C, Θ, σ^2 and Ω, associated with the state space model in EQ -1 and EQ -2 by using the statistical estimation method of maximum likelihood. The first step involves the derivation of the likelihood function which is the joint probability distribution of $Y_1, Y_2, ---Y_N$ as a function of the stationary parameters. Now, the joint probability distribution function of $Y_1, Y_2, ---Y_N$ can be expressed as the product of the conditional distribution of Y_w given $Y_1, Y_2, ---Y_{w-1}$. Now, $f(Y_1, Y_2, ---Y_N) = f(Y_1)f(Y_2/Y_1)f(Y_3/Y_1,Y_2)------f(Y_N/Y_1,Y_2,---Y_{N-1})$, based on the following probability relationship, $f(Y_2/Y_1) = f(Y_1,Y_2)/f(Y_1)$, or $f(Y_1,Y_2) = f(Y_1)f(Y_2/Y_1)$ and $f(Y_3/Y_1,Y_2) = f(Y_1,Y_2,Y_3)/f(Y_1,Y_2)$ or $f(Y_1,Y_2,Y_3) = f(Y_1,Y_2)f(Y_3/Y_1,Y_2)$. Now substituting for $f(Y_1,Y_2)$, $f(Y1,Y_2,Y_3) = f(Y_1)f(Y_2/Y_1)f(Y3/Y_1,Y_2)$. This process can be carried on for $w = 1,2,---N$.

Futhermore, the distribution of Y_w given $Y_1, Y_2, ---Y_{w-1}$ is Gaussian with mean equal to $X'_wC(w/w-1)$(see EQ-5) and variance equal to G_w(EQ-6). Explicitly then, the conditional probability function is,

$$f(Y_w/Y_1,Y_2,---Y_{w-1})=(2\pi G_w)^{-\frac{1}{2}}\exp[-(Y_w-X'_wC(w/w-1)^2/2G_w]$$

Now it follows that the likelihood function or the joint probability function of the observations is given by L =

$$(2\pi)^{-N/2} \prod G_w^{-\frac{1}{2}}\exp[-\Sigma(Y_w - X'_wC(w/w-1))^2/2G_w]$$

For computational convenience, however, one uses the log of the likelihood function.

$$\log L = [-N/2\log(2\pi)] - [1/2\Sigma\log G_w] - [1/2\Sigma(Y_w - X'C(w/w-1))^2/G_w]$$

The second step of the procedure is to maximize the above log likelihood function with respect to the parameters of interest. In order to derive the maximum likelihood estimates of the parameters, one requires those values of the parameters for which the log likelihood function is a maximum or equivalently, one needs to minimize the function,

$$\Sigma[\mathrm{Log}\ G_w + (Y_w - X'_wC(w/w-1))^2G_w^{-1}]$$

Notice that, in addition to the observations Y_w and X_w, the expression involves the conditional expectation

C(w/w-1) of the stochastic parameters, and the conditional variance G_w of the dependent variable. Both of these quantities may be conveniently evaluated as functions of the stationary parameters of the model through the Kalman filter algorithm, as shown in Part IV of this section. However, it is not possible to solve this function minimization problem analytically (that is, with a manageable algebraic expression), since C(w/w-1) and G_w will be rather extremely complicated functions of the stationary parameters of the state-space model. Therefore, numerical function optimization algorithms must be employed. The most popular among the numerical function minimization algorithms is the Newton-Raphson method (with its subsequent improvements)(11) which, however, requires the first derivative of the function to be minimized. In fact, as discussed in Pagan(12), it is possible to obtain these quantities analytically. The details of the derivatives are extremely cumbersome. An alternate approach is to evaluate the derivatives numerically.

The numerical maximization of the log likelihood function yields the point estimates of the stationary parameters of the model. The standard errors associated with such estimators can be achieved through Fisher's (13) minimum variance estimator of the maximum likelihood estimates, as follows:

$VAR(B^*)min = -E(\delta^2 logL/\delta B^* \delta B^{*\prime})^{-2}$ where, B* denotes the vector of all the stationary parameters to be estimated, L represents the likelihood function, δ^2 denotes the second derivative of L with respect to B*, the reciprocal of the minimum variance (1/VAR(B*))provides the well known Fisher's information matrix(13), I(B*); E stands for the expected value indicating that the derivatives are evaluated at the estimated values of the parameters, and the square root of the expression VAR(B*) provides the respective standard errors.

It should be noted here that without the help of the high capacity modern computers, the analysis of the stochastic parameter model with multiparameter simultaneous numerical search process will be absolutely prohibitive.

Part VI: Statistical Tests of Validity for Model Adequacy:

A detailed description of the following three models has been presented in Part I and Part II of this section, (a) fixed (invariant) parameter model (FPM), (b) random coefficient model (RCM) and (c) stochastic parameter model (SPM) or sometimes called the Kalman filter model (KFM). Given a set of space-dependent (time-dependent) data for analysis, the interest here is to determine which one of the three models, indeed, adequately describes the data. In other words, if a specific model has been selected for the analysis, it is incumbent upon the selector to demonstrate that the model is truly valid for the purpose of the analysis. Three statistical tests have been developed to accomplish this task. They are called the statistical tests of validity for model adequacy. Since FPM, RCM and KFM constitute a sequence of nested alternatives, in the sense that each generalizes the previous model, the order of the three tests will be, as follows, (i)FPM vs. RCM, (ii) RCM vs. KFM and (iii) FPM vs. KFM.
It should be noted here that the test statistics are extremely complex and consequently the description of the tests will be confined to the models involving only a single independent variable. Indeed, the extension to the case involving more than a single stochastic parameter is extremely difficult. However, a two-variable model should provide a better understanding of the test mechanism and also of the test results.

Consider the following interconnected models,

$$Y_W = C_O + C_W X_W + E_W \qquad w = 1,2,\text{--}N \qquad \text{EQ-13}$$

and

$$C_W = C + \theta(C_{W-1} - C) + A_W \qquad \text{EQ-14}$$

where C_O is the fixed intercept, E_W denotes the error term with zero mean and fixed variance $\sigma_e{}^2$, C is the stationary mean of C_W, $E(C_W) = C$, θ denotes the autoregressive coefficient, that is, correlation of $(C_W, C_{W-1}) = \theta$, the range of θ is, $-1 < \theta < +1$, and A_W is the error term for the stochastic coefficient with zero mean, fixed variance $\sigma_a{}^2$ and is neither autocorrelated or correlated with error term E_W. Note that, for FPM Cw = C which is fixed across space, for RCM the regression coefficient is stochastic but θ is equal to zero and for KFM the autoregressive parameter is not necessarily zero. Based on these properties of the models, a description of the individual tests is presented. It should be noted here that there are two kinds of statistical tests of hypothesis, likelihood ratio test and Lagrangian multiplier test, which will be elaborated in the description of the tests.

TEST - I (FPM vs. RCM)(14,15,16): In this test, the null hypothesis of fixed parameter model is tested against the alternate hypothesis of random coefficient model. In terms of the test notation,

$$H_O: \sigma^2{}_a = 0 \text{ and } H_1: \sigma^2{}_a > 0 \ (\theta = 0)$$

The Lagrangian multiplier test is considered for TEST-I. This test is based on the first derivative of the log likelihood function with respect to $\sigma^2{}_a$, evaluated under the null hypothesis in which $\sigma^2{}_a$ is equal to zero. The derivation involves combining EQ-13 and EQ-14 with θ set to zero,

$$Y_W = C_O + CX_W + E_W + X_W A_W$$

This is essentially a simple regression model with fixed parameters and an error term $(E_W + X_W A_W)$ that has zero mean and a variance, of $VAR(E_W + X_W A_W) = \sigma^2{}_e + X^2{}_W \sigma^2{}_a$.

The likelihood function assuming a Gaussian distribution is as follows:

$L = (2\pi)^{-N/2} \Pi(\sigma^2_e + X^2_w\sigma^2_a)^{-\frac{1}{2}}\exp[-\Sigma(Y_w - C_o - CX_w)^2/2(\sigma^2_e + X^2_w\sigma^2_a)]$

$\mathrm{Log}(L) = -N/2\log(2\pi) - 1/2\Sigma\log(\sigma^2_e + X^2_w\sigma^2_a)^{-\frac{1}{2}}\Sigma[(Y_w - C_o - CX_w)^2/(\sigma^2_e + X^2_w\sigma^2_a)]$

Differentiating the log likelihood function with respect to σ^2_a yields,

$d\log L/d\sigma^2_a = -1/2\Sigma[X^2_w/(\sigma^2_e + X^2_w\sigma^2_a)] + \frac{1}{2}\Sigma[X^2_w(Y_w - C_o - CX_w)^2/(\sigma^2_e + X^2_w\sigma^2_a)^2]$

Now the estimated regression line is $Y_w{*} = c_o + cX_w$, the regression residual is $R_w = (Y_w - Y_w{*})$ and $S^2_e = \Sigma R_w{}^2/N$, residual variance. Substituting these quantities, the above expression becomes,

$d\log L/d\sigma^2_a = -1/2\Sigma[X^2_w/S^2_e) + \frac{1}{2}\Sigma X^2_w R^2_w/S^4_e$

The Lagrangian multiplier test statistic $[T(\theta=0)$ or $T(0)]$ is obtained by dividing the right hand side by its standard error, as follows,

$T(0) = [1/2\Sigma X^2_w(R^2_w/S^2_e - 1)]/[\frac{1}{2}\Sigma X^4_w - (1/2N)(\Sigma X^2_w)^2]^{\frac{1}{2}}$

Computationally then, (i) estimate by the least squares method the standard linear regression residuals, R_w and the residual regression variance, S^2_e, (ii) compute the T(0) statistic above and reject the null hypothesis of FPM in favor of RCM at a nominal significance level of 0.05 when T(0) exceeds the standard normal critical value of 1.645, and (iii) conclude that the regression coefficient is indeed not invariant, that is, the coefficient is of stochastic nature.

<u>TEST - II (RCM vs. KFM)(17)</u>: In this test, the null hypothesis of random coefficient model is tested against the alternate hypothesis of stochastic parameter model obeying a first order autoregressive process with non-zero autoregressive coefficient. In terms of the test notation,

$$H_o: \theta = 0 \text{ and } H_1: \theta \neq 0$$

The likelihood ratio(LR) test is proposed here. In Part V of this section, it is shown that the Kalman filter algorithm can be used to compute the likelihood function of the stochastic parameter regression model by obtaining the maximum likelihood estimates of the stationary parameters of the model. The likelihood ratio test is based on the difference between the maximized log likelihoods under the alternative and null hypothesis. In other words, let L_o denote the highest possible value of the likelihood function under the null hypothesis and L_1 the value of the maximized likelihood under the alternative hypothesis. It is well known (10) that the test statistic,

$$LR = 2\log(L_1/L_0) = 2(\log L_1 - \log L_0)$$

has the chi-square distribution with one degree of freedom, since the test pertains to a single parameter. The test is carried out in the following steps: (i) Find the maximum of the likelihood function given in Part V over the permissible range of the stationary coefficients, $-\infty < C_0 < +\infty$, $-\infty < C < +\infty$, $\sigma^2{}_e > 0$, $\sigma^2{}_a > 0$ and $-1 < \theta < +1$. Denote by L_1, the value of the maximized likelihood, (ii) Repeat the previous step, maximizing the likelihood function over the permissible range of the stationary coefficients, <u>but</u> now with the autoregressive coefficient θ set to zero. Denote by L_o the value of the maximized likelihood subject to this constraint, and (iii) Calculate the test statistic given above and reject the null hypothesis if the calculated value exceeds 3.84, the tabulated chi-square value for one degree of freedom with a significance level of 0.05 in the upper tail of the distribution.

<u>TEST - III (FPM vs. KFM)(14,18)</u>: In this test, the null hypothesis of fixed parameter model is tested against the alternate hypothesis of stochastic parameter model associated with a first-order autoregressive

process. The test notation is the same as that of TEST-I, except that θ is unrestricted, as follows:

H_0: $\sigma^2_a = 0$, $H_1 = \sigma^2_a > 0$ (θ unspecified in the range $-1 < \theta < +1$)

This particular situation involves the additional difficulty that under the null hypothesis θ is undetermined. This difficulty is circumvented through a modification of the Lagrangian multiplier (LM) test by Davies(18). The LM test statistic is

$$T(\theta) = N^*/D^* \quad \text{where,}$$

$N^* = 1/2\Sigma X^2_w(R^2_w/S^2_e - 1) + (1/S^2_e)(\Sigma R_w X_w \Sigma R_i X_i \theta^{w-i})$ and

$D^* = [1/2\Sigma X^4_w + \Sigma X^4_w \Sigma X^2_i \theta^{2(w-2)} - 1/2N(\Sigma X^2_w)^2]^{\frac{1}{2}}$

(For N*, the range of the first sum sign is w=1 to w=N, the range of the second sum sign is w=2 to w=N and the range of the third sum sign is i=1 to i=w-1. For D*, the range of the first sum sign is w=1 to w=N, the range of the second sum sign is w=2 to w=N, the range of the third sum sign is i=1 to i=w-1 and the range of the fourth sum sign is w=1 to w=N.) It is known that the statistic $T(\theta)$ follows a standard normal distribution(18).

Now the alternate hypothesis is not a stochastic parameter model with some particular value of θ, but is rather of such a model where θ might take any value in the range $-1 < \theta < +1$. This suggests that, one has to compute the statistic $T(\theta)$ for a grid of possible values of θ in this region and to base a test of the null hypothesis on the largest value taken by $T(\theta)$ over the range $-1 < \theta < +1$, that is, MAX T (θ). Let M be the critical value, and therefore, MAX $T(\theta)$ is significant when it exceeds M. The calculation of the p-value of the test is extremely complicated and so Davies (18) proposed an upper bound to the p-value of the test as follows,

Prob[Max $T(\theta)$ >M] < F(-M) + Q(M) where F is the cumulative standard Gaussian distribution so that,

$$F(-M) = \int_{-\infty}^{M} (2\pi)^{-\frac{1}{2}} \exp(z^2/2)dz$$

and $Q(M) = (2\pi)^{-1}\exp(-M^2/2) \int_{-1}^{+1}(-B_{11}(U)dU$ EQ - 15

where $B_{11}(U) = [\delta^2/\delta\theta^2 P(\theta,U)]$

That is, $B_{11}(U)$ is the second partial derivative of $P(\theta,U)$ with respect to θ, evaluated at $\theta = U$. Here, $P(\theta,U)$ is defined as, $P(\theta,U) = N^{**}/D^{**}$, where

$N^{**} = P(\theta,U) = 1/2\Sigma X^4{}_w + \Sigma X^2{}_w \underline{\theta}^{w-i}\underline{U}^{w-i} - 1/2N(\Sigma X^2{}_w)^2$

$D^{**} = [P(\theta),P(U)]$ where $P(\theta) = P(\theta,U)$ except substitute $\theta^{2(w-i)}$ for the two underscored terms and $P(U) = P(\theta,U)$ except substitute $U^{2(w-i)}$ for the two underscored terms. In practice, the partial derivative of $B_{11}(U)$ and the integral of Q(M) are easily found by resorting to numerical analysis methods. The test procedure is then as follows: (i) Estimate the regression line, $Y_w^* = c_o + cx_w$, the regression residual $R_w = (Y_w - Y_w^*)$ and $S^2{}_w = \Sigma R^2{}_w/N$, residual variance, (ii) Calculate the statistic $T(\theta)$ for a grid of possible values of θ and hence locate the maximum value $T(\theta)$ takes in the range $-1 < \theta < +1$. Denote this maximum as M, (iii) The quantity F(-M) can be read directly from the tables of the cumulative distribution function of the standard Gaussian distribution and Q(M) is calculated from EQ - 15, (iv) The null hypothesis that the FPM is adequate can be rejected against the alternative of a KFM with unspecified θ, at any significance level greater than F(-M) + Q(M).

It should be noted here that all the three tests must be carried out prior to the interpretation of the outcome. Then, once a model is identified based on the three tests, that model should be considered for the analysis.

MULTICOMPONENT MIXTURE STUDIES: RESULTS AND DISCUSSION

A multicomponent mixture analysis is generally conducted when during the stability determination of drug formulation, one discovers in the sample the presence of not only the intact drug but also its degradation products. Recent advances in spectrophotometers and laboratory computers suggest that, it might be possible to perform such determinations rapidly by resorting to simultaneous multicomponent spectrophotometric analysis on the unseparated mixture of the intact drug and low-level degradation products.

For the purpose of illustration of the statistical methodologies described in the previous section, the data from two recent experiments are considered. A two-component mixture experiment is performed on each of the two drug products, BP and MP. The absorbances of the mixture (Y) and the absorbances of the two standards of components (solutes), X_1 and X_2, are measured at each of the 31 equally spaced wavelengths, ranging from 280-310 nm. The symbol C has been used to represent the concentrations (mg/ml) of the solutes relative to those of the standards in the mixture. The analysis is performed using a HP-8450 UV/VIS diode array spectrophotometer equipped with a 1.00 cm flow cell and an autosampler, yielding a resolution of 1 nm for 200-400 nm and of 2 nm for 400-800 nm. Spectra can be obtained as absorbance, first derivative or second derivative. The method considered here can detect degradates at concentrations lower than 0.05% of the parent compound.

It would be appropriate to note here that, the variation in the concentration of a component within a spectrum may be caused by the presence of unknown components, interferences among the components, inherent variability in instrument performance, and physical and chemical factors such as, solvent quality, temperature, impurities, evaporation and degree of incidence.

Essentially the same conditions depicted above could conceivably bring about autocorrelation and heteroscedastity among the measured responses.

The data collected from the spectrophotometric assay are subjected to the various statistical analyses depicted in the previous section. However, for the purpose of simplification of the interpretation of the results, a component-wise analysis is undertaken, and thus the results presented below pertain primarily to the degradate component of each of the two drug products, BP and MP.

Table-I depicts the estimation aspects of the statistical results associated with BP. The interpretations of the point estimates associated with these three models are as follows: (i) the FPM results indicate that the concentration of the degradate (COD) is estimated to be 0.5760 $\pm$ 0.0170 mg/ml for the entire spectrum under the assumption that the regression coefficient remains invariant. The 95% confidence limits are 0.54 and 0.61. The Durbin-Watson(19) (DW) test statistic (0.2154) indicates that there is a significant serial correlation between the adjacent residuals from regression. (ii) The results of the RCM analysis indicate that the COD varies across the spectrum around a mean value of 0.5465 $\pm$ 0.0114 with two-sigma confidence limits of 0.52 and 0.57, and (iii) For the first order autoregressive stochastic model, however, the COD is estimated to be 0.6142 $\pm$ 0.0713 with with θ = 0.9613, which is the measure of the correlation between two consecutive concentrations across space. The two-sigma confidence limits for the COD are 0.47 and 0.75. The maximized value of the log likelihood function for the three models, FPM, RCM and KFM are -54.7, 5.02 and 24.10 respectively. These are the results of the numerical maximization process. The next

TABLE -I

STATISTICAL RESULTS FOR PRODUCT - BP

Fixed (Invariant) Parameter Model (FPM)

Statistic	c_o	c	S^2_e	R^2
estimates	0.1002	0.5760	0.2031E-3	0.974
Std.Errors	0.3552	0.1703E-1*	(DW = 0.2154)	

Random Coefficient Model (RCM)

Statistic	c_o	c	S^2_e	S^2_a
estimates	0.3662	0.5465	0.8618E-4	0.2752E-2
Std.Errors	0.4366E-2	0.1138E-1	0.6584E-4	0.8164E-3

Stochastic Parameter Regression Model (KFM)

Statistic	c_o	c	S^2_e	S^2_a
estimates	0.3612	0.6142	0.8206E-4	0.5294E-3
Std.Err.	0.7248E-2	0.7131E-1	0.4666E-4	0.1653E-3

θ = 0.9613 θ Std. Err. = 0.3987E-1

* Floating-point decimal S^2_a = Maximum likelihood estimate of σ^2_a

STD = Standard

TABLE -II

ESTIMATES OF CONCENTRATIONS AND STANDARD ERRORS

Wavelength	C(W/N)	$V(W/N)^{\frac{1}{2}}$	Wavelength	C(W/N)	$V(W/N)^{\frac{1}{2}}$
280	0.7279	0.000250	296	0.5171	0.002000
281	0.6513	0.000246	297	0.5260	0.002984
282	0.6019	0.000245	298	0.5377	0.004487
283	0.5699	0.000247	299	0.5439	0.006669
284	0.5477	0.000252	300	0.5611	0.009486
285	0.5327	0.000260	301	0.5935	0.012922
286	0.5224	0.000269	302	0.5927	0.016140
287	0.5156	0.000280	303	0.6176	0.018975
288	0.5107	0.000295	304	0.6202	0.020514
289	0.5079	0.000319	305	0.6089	0.021124
290	0.5067	0.000359	306	0.6177	0.022009
291	0.5063	0.000425	307	0.6125	0.022301
292	0.5058	0.000530	308	0.6113	0.022455
293	0.5076	0.000695	309	0.6164	0.022443
294	0.5100	0.000953	310	0.6018	0.022491
295	0.5122	0.001361			

step will be to examine the results of the three tests of validity for model adequacy, as follows: (i) FPM vs. RCM: The value of the Lagrangian multiplier test statistic, T(O) is 5.358 and the standard normal critical value for a one-sided 5% test is 1.645 which indicates that FPM is rejected ($p < 0.05$) and RCM is accepted. An approximate test (15) can be accomplished by computing $Z_{.05}$ = 0.2752E-2/0.8164E-3) = 3.371 which is significant (see Table-I for column-heading $S^2{}_a$ in RCM). (ii) RCM vs. KFM: The value of the likelihood ratio test statistic is 36.18 and the chi-square value at 5% significance level with one degree of freedom is 3.84 indicating that RCM is rejected ($p < 0.05$) and KFM is accepted and (iii) FPM vs. KFM: The value of the test statistic $T(\theta)$ = 8.1588 which is the maximum $T(\theta)$ for θ = 0.47. The upper bound critical probability or the p-value is F(-8.1588) + Q(8.1588) = 0.2461E-5 implying the fact that FPM is rejected at any numerical significance level, and KFM is accepted. A corollary to this test is the standard normal test (for $\theta \neq 0$) $Z_{.05}$ = (0.9613/0.3987E-1) = 24.11 which is highly significant(See Table-I, last row). The results of the three tests clearly indicate that the stochastic parameter regression model (KFM) is the model to be used ($p \ll 0.01$) for the analysis. It also implies that the COD does indeed vary across the entire spectrum. Table-II demonstrates the estimate of the COD [C(W/N)] and its standard error $[V(W/N)]^{\frac{1}{2}}$ at each of the 31 wavelengths. A cursory examination of the table indicates that the concentrations range from 0.5058 to 0.7279 with the maximum and minimum occurring at wavelengths 280 and 292 respectively. It is interesting to note that this concentration range is totally contained within the two-sigma confidence limits associated with KFM (0.47-0.75) indicating that this model is fully sensitive to the

TABLE -III

STATISTICAL RESULTS FOR PRODUCT - MP

Fixed (Invariant) Parameter Model (FPM)

Statistic	c_o	c	S^2	R^2
estimates	0.1046E-2	0.4965	0.9214E-8	1.000
Std.Errors	0.2651E-4	0.1208E-3*		

Random Coefficient Model (RCM)

Statistic	c_o	c	$S^2{}_e$	$S^2{}_a$
estimates	0.1047E-2	0.4965	0.6607E-8	0.5327E-7
Std.Errors	0.2277E-4	0.1254E-3	0.2438E-8	0.5294E-7

Stochastic Parameter Regression Model (KFM)

Statistic	c_o	c	$S^2{}_e$	$S^2{}_a$	θ
estimates	0.1046E-2	0.4965	0.6593E-8	0.4272E-7	.4878
Std.Err.	0.2308E-4	0.1548E-3	0.2400E-8	0.5428E-7	.5294

* Floating-point decimal $S^2{}_a$ = Maximum likelihood estimate of $\sigma^2{}_a$

STD = Standard

stochastic phenomenon involved. The confidence limits associated with the other two models fail to do this. The reliability of each of the estimates is very high here because of the low magnitudes of the standard errors.

The results of the statistical analysis pertaining to Product-MP are presented in TABLE-III. The results show that not only the value of R^2 in FPM is equal to 1.0 but also the residual variance ($S^2{}_e$) is virtually zero, indicating that the model fits the data perfectly. In this situation, one would expect the three models to perform similarly. Indeed, the C-coefficients as well as their respective standard errors associated with the three models are essentially of the same magnitude.

This also goes for the case of the intercept (C_o). It should be noted here that the maximum value of the log likelihood function for FPM, RCM, and KFM are 219.305, 220.044, and 220.256 respectively. All three functions also attained the same maximum.

The next step is to examine the results of the tests of validity for model adequacy, as follows: (i) FPM vs. RCM: The value of the Lagrangian multiplier test statistic is 1.294 (<1.645) which is not significant ($p > 0.05$) and hence the null hypothesis that FPM=RCM can not be rejected, (ii) RCM vs. KFM: The value of the likelihood ratio statistic is 0.424 (< 3.84) which is not significant ($p > 0.05$) and hence the null hypothesis that RCM=KFM can not be rejected and (iii) FPM vs. KFM: The value of the test statistic $T(\theta)$= 1.498 which is the maximum $T(\theta)$ when θ=0.34 with a upper bound p-value of 0.0696 ($p>0.05$) and hence the null hypothesis FPM=KFM can not be rejected. The results of the above three tests clearly lead to the conclusion that, FPM (standard least squares model) is the model of choice ($p>0.05$) for the analysis.

These two examples (BP and MP) illustrate two different prospectives of the statistical methodology depicted in this presentation.

ACKNOWLEDGEMENTS

Sincere thanks are due to Dr. P.K. Shiromani of Merck and Dr. S.K. Sarkar of SKB for the discussion on spectrophotometer operations and for exhibiting avid interest in statistical applications. Grateful thanks are due to Mrs. B. Mukerji for carrying out the least squares computations as well as other required computations in a PC. As always, special thanks are due to Mrs. B.J. Tomlinson for her meticulous efforts in accomplishing the task of typing the manuscript with utmost rapidity.

REFERENCES

1. N.R. Bohidar and K.E. Peace. Pharmaceutical Formulation Development. Chapter IV. "Biopharmaceutical Statistics for Drug Development". Marcel Dekker, Inc. New York (1988).

2. R.E. Kalman. A New Approach to Linear Filtering and Prediction Theory. Journal of Basic Engineering, Vol. 82, 34-45 (1960).

3. R.E. Kalman and R.S. Bucy. New Results in Linear Filtering and Prediction Theory. Journal of Basic Engineering, Vol. 83, 95-108 (1961).

4. G.E.P. Box and G.M. Jenkins. "Time Series Analysis, Forecasting and Control", Revised Edition, Holden-Day, San Francisco (1976).

5. G.E. Tiao. and G.E.P. Box. Modeling Multiple Time Series with Applications, Journal of American Statistical Association, Vol 76, 802-816 (1981).

6. T.W. Anderson. "An Introduction to Multivariate Statistical Analysis". Wiley Publications, New York (1971).

7. C.F. Ansley. and R. Kohn. A Geometrical Derivation of the Fixed Interval Smoothing Algorithm, Biometrika, Vol 69,486-487 (1982).

8. N. Gupta. and R. K. Mehra. Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations, IEEE Transactions on Automatic Control AC, Vol 19 (6), 774-783 (1974).

9. F.C. Schweppe. Evaluation of Likelihood Functions for Gaussian Signals. IEEE Transactions on Information Theory IT, Vol 11, 61-70 (1965).

10. A.M. Mood. and F.A. Graybill. "Introduction to the Theory of Statistics", 2nd ed. McGraw-Hill, New York (1963).

11. R.H. Pennington. "Introductory Computer Methods and Numerical Analysis". 2nd ed. The Macmillan Company, New York (1970).

12. A.R. Pagan. Some Identification and Estimation Results for Regression Models with Stochastically Varying Coefficients, Journal of Econometrics, Vol. 13, 341-363 (1980).

13. R.L. Anderson. and T.A. Bancroft. "Statistical Theory in Research", McGraw-Hill Book Co. New York (1952).

14. T.S. Breusch. and A.R. Pagan. The Lagrange Multiplier Test and Its Application to Model Specification in Econometrics, Review of Economic Studies, Vol. 47, 239-254 (1980).

15. T.S. Breusch. A Simple Test for Heteroscedasticity and Random Coefficient Variation, Econometrica. Vol. 47, 1287-1294 (1979).

16. D.S. Silvey. The Lagrangean Multiplier Test. Annals of Mathematical Statistics, Vol. 30, 389-407 (1959).

17. D. Chant. On Asymptotic Tests of Composite Hypotheses in Non-standard Conditions, Biometrika, Vol. 61(2), 291-298 (1974).

18. R.B. Davies. Hypothesis Testing When A Nuisance Parameter is Present Only Under the Alternative. Biometrika, Vol. 64(2) 247-254 (1977).

19. J. Durbin and G.S. Watson. Testing for Serial Correlation in Least Squares Regression III. Biometrika, Vol. 58, 1-19 (1971).