Journal of Applied Probability & Statistics Vol. **1**, No. 1, pp. 15-29

Prediction Distribution of Future Regression and Residual Sum of Squares Matrices for Multivariate Simple Regression Model with Correlated Normal Responses

Shahjahan Khan

Department of Mathematics & Computing, Australian Centre for Sustainable Catchments, University of Southern Queensland, Toowoomba, Queensland, Australia. Email: khans@usq.edu.au

Abstract

This paper considers multivariate simple regression model under normally distributed errors, for both realized and future responses, with unknown regression parameters ($\boldsymbol{\beta}$) and covariance matrix ($\boldsymbol{\Sigma}$). The prediction distributions of the future regression matrix (FRM) and future residual sum of squares matrix (FRSSM) for the future regression model are obtained. Conditional on the realized responses, the FRM follows a matrix T distribution whose shape parameter depends on the sample size and the dimension of the regression parameters in the model, and the FRSSM follows a scaled generalized beta distribution. The same results have been obtained by both the classical and Bayesian methods under uniform prior.

Keywords: Multivariate simple regression model, invariant differentials, uniform prior, predictive inference, future regression matrix and residual sum of squares matrix, and matrix normal, matrix T and generalized beta distributions.

2000 Mathematics Subject Classification: 62H10, 62J05.

1 Introduction

When several response variables are associated with one single value of an explanatory variable the use of the simple or multiple regression model is inappropriate. Then a multivariate simple regression model can be used to analyze the data. The multivariate simple regression model is different from both the simple and multiple regression models. It is a more general model than the commonly used simple linear regression model. In the simple regression model there is only one response variable corresponding to a specific value of the explanatory variable. Whereas in a multivariate simple regression model there are a set of values of several ($p \geq 2$) response variables corresponding to a single value of the explanatory variable. Thus the simple regression model is a special case of the multivariate simple regression model when $p = 1$. It is used to analyze data from studies where there are more than one response variables for a particular value of the explanatory variable. For example, if patients are given the same dose of a medicine to observe responses on $p \geq 2$

characteristics of every subject, then, for one particular value of the explanatory variable, there will be p different values on p different response variables. The model can also be applied to any other experimental or observational studies where several response variables are generated for one particular value of the independent variable. Khan (2005) studied different improved estimators for the multivariate simple regression model. For details on the multivariate simple regression model see Saleh (2006).

Contrary to the commonly available prediction of observables, following Khan (2004), here we propose the prediction distribution of the FRM and FRSSM for the multivariate simple regression model. Although traditionally predictive inference is directed towards inference involving the observables rather than the parameters, Khan (2002, 2004) proposed predictive inference for the future parameters. In general, predictive inference uses the realized responses from the *performed experiment* to make inferences about the behavior of the unobserved future responses of the *future experiment* (cf. Aitchison and Dunsmore (1975, p.1)). The outcomes of the two experiments are connected through the same structure of the model and indexed by the common set of parameters. The prediction distribution forms the basis of all predictive inferences. For details on the predictive inference methods and wide range of applications of prediction distribution interested readers may refer to Aitchison and Sculthorpe (1965) and Geisser (1993). Predictive inference for a set of future responses of a model, conditional on the realized responses from the same model, has been derived by many authors including Fraser and Haq (1969), Aitchison and Dunsmore (1975), and Haq and Khan (1990). The prediction distribution of a set of future responses from the model has been used by Haq and Rinco (1976) to derive β-expectation tolerance region. Guttman (1970) and Aitchison and Dunsmore (1975) obtained different tolerance regions from the prediction distribution.

There has been many studies in the area of predictive inference mainly for multiple regression models with independent and normal errors. The pioneering work in this area includes Fraser and Haq (1969), Guttman (1970), and Haq and Khan (1990). The Bayesian works include Aitchison (1964), and Aitchison and Sculthorpe (1965). Aitchison and Dunsmore (1975) provide an excellent account of the theory and application of the prediction methods. Fraser and Haq (1969) and Khan and Haq (1994) obtained prediction distribution for the multivariate normal and Student-t models, respectively, by using the structural distribution approach. Haq (1982) used the structural relations, rather than the structural density function, to derive the prediction distribution. Geisser (1993) discussed the Bayesian approach to predictive inference and discussed a wide range of real-life applications in many areas. This includes model selection, discordancy, perturbation analysis, classification, regulation, screening and interim analysis.

This paper considers the widely used multivariate simple regression model with correlated normal responses for the realized as well as the future models. The two sets of responses are connected through the common set of regression and scale parameter matrices. Following Khan (2002, 2004), we pursue the predictive approach to derive the distribution of the FRM and FRSSM for the future responses, conditional on a set of realized responses. The distribution of the FRM and FRSSM of the future responses, conditional on the realized responses, are obtained. The predictive distribution of the FRM follows a matrix T distribution, and the FRSSM of the future regression follows a scaled generalized beta distribution. The distribution of the statistics for the future regression model, conditional on the realized responses, are dependent, and hence their joint density can't be factorized. Identical prediction distributions, for both the FRM and ERSSM, are obtained by both the classical and Bayesian approaches under uniform prior. For a good comparison of Bayesian and classical prediction see Ren et al. (2004).

The multivariate simple regression model with independent normal errors has been introduced in Section 2. Some notations and preliminaries are provided in Section 3. The multivariate simple regression model for the future responses is discussed in Section 4. The predictive distributions of the FRM and FRSSM, conditional on the realized sample, are derived in Section 5 by the classical approach. In Section 6, the same prediction distributions are obtained by Bayesian method under uniform prior. Some concluding remarks are included in Section 7.

2 The Multivariate Simple Regression Model

Let $\boldsymbol{y}_j$ be a p-dimensional column vector of the values of the jth response on a set of p dependent variables associated with a single value of the explanatory variable x_j from a multivariate simple regression model. Then $\boldsymbol{y}_j$ can be represented by the set of linear equations

$$\boldsymbol{y}_j = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 x_j + \Gamma \boldsymbol{e}_j \quad \text{for } j = 1, 2, \ldots, n \tag{2.1}$$

where $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$ are the p-dimensional intercept and slope parameters respectively, Γ is a $p \times p$ non-singular scale parameter matrix, and $\boldsymbol{e}_j$ is the vector of error variables associated with the responses $\boldsymbol{y}_j$. Assume that each component of the error vector, $\boldsymbol{e}_j$, is identical and independently distributed as a normal variable with location 0 and scale 1, that is, $\boldsymbol{e}_j \sim N(\boldsymbol{0}, I_p)$ in which I_p is an identity matrix of order p. The equation in (2.1) can be written in a more precise form as

$$\boldsymbol{y}_j = \boldsymbol{\beta} \boldsymbol{z}_j + \Gamma \boldsymbol{e}_j \tag{2.2}$$

where $\boldsymbol{\beta} = [\boldsymbol{\beta}_0, \boldsymbol{\beta}_1]$, a $p \times 2$ dimensional matrix of regression parameters; and $\boldsymbol{z}_j = [1 \; x_j]'$, a 2×1 dimensional design matrix of known values of the regressor for $j = 1, 2, \ldots, n$. Therefore, the joint density function of the error vector $\boldsymbol{e}_j$ can be written as

$$f(\boldsymbol{e}_j) = [2\pi]^{-\frac{p}{2}} e^{-\frac{1}{2}\boldsymbol{e}_j'\boldsymbol{e}_j}. \tag{2.3}$$

and that of the response vector $\boldsymbol{y}_j$ as

$$f(\boldsymbol{y}_j|\boldsymbol{\beta},\Gamma) = [2\pi]^{-\frac{p}{2}}|\boldsymbol{\Sigma}|^{-1}e^{-\frac{1}{2}(\boldsymbol{y}_j-\boldsymbol{\beta}\boldsymbol{z}_j)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}_j-\boldsymbol{\beta}\boldsymbol{z}_j)} \tag{2.4}$$

where $\boldsymbol{\Sigma} = \Gamma\Gamma'$, the covariance matrix of the response vector $\boldsymbol{y}_j$. Now, a set of $n > p$ responses, $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n]$, from the above multivariate simple regression model can be expressed as

$$\boldsymbol{Y} = \boldsymbol{\beta}\boldsymbol{Z} + \Gamma\boldsymbol{E} \tag{2.5}$$

where $\boldsymbol{Y}$ is the response matrix of order $p \times n$; $\boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n]$ is a $2 \times n$ dimensional design matrix of known values of the regressor; and $\boldsymbol{E}$ is a $p \times n$ dimensional matrix of the random error components associated with the response matrix $\boldsymbol{Y}$. It may be noted here that the nonconventional representation of responses as row vectors, rather than column vectors, is adopted in line with Fraser (1968 and 1979) to facilitate straightforward comparison of results.

Since each of the p-dimensional response column vector, $\boldsymbol{y}_j$, follows a multivariate normal distribution, the joint density function of the $p \times n$ order response matrix $\boldsymbol{Y}$ follows a *matrix normal* distribution with the density function

$$f(\boldsymbol{Y}|\boldsymbol{\beta},\boldsymbol{\Sigma}) = [2\pi]^{-\frac{pn}{2}}|\boldsymbol{\Sigma}|^{-\frac{n}{2}}e^{-\frac{1}{2}tr\left\{\boldsymbol{\Sigma}^{-1}(\boldsymbol{Y}-\boldsymbol{\beta}\boldsymbol{Z})(\boldsymbol{Y}-\boldsymbol{\beta}\boldsymbol{Z})'\right\}} \tag{2.6}$$

where $tr(\Omega)$ is the trace of the matrix Ω. Although each column of $\boldsymbol{E}$ is independent of other columns, the columns of $\boldsymbol{Y}$ are not independent. The above multivariate simple regression model represents a set of responses as the independent realizations of the *performed experiment*. We term the responses from the performed experiment as the realized responses. Our aim is to derive the prediction distribution of the FRM and FRSSM based on a set of n_f unrealized future responses from the *future experiment*, conditional on the realized responses.

3 Some Notations and Preliminaries

In this Section we introduce some useful notations to facilitate the derivation of the results in the forthcoming Sections. Let the regression matrix of $\boldsymbol{E}$ on $\boldsymbol{Z}$ be $\boldsymbol{B}_E$ and the residual sum of squares matrix of the error regression be S_E. Then we have

$$\boldsymbol{B}_E = \boldsymbol{E}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1} \text{ and } \boldsymbol{S}_E = [\boldsymbol{E} - \boldsymbol{B}_E\boldsymbol{Z}][\boldsymbol{E} - \boldsymbol{B}_E\boldsymbol{Z}]'. \tag{3.1}$$

Let C_E be a positive definite nonsingular matrix such that $\boldsymbol{S}_E = \boldsymbol{C}_E\boldsymbol{C}_E'$. Then define standardized residuals matrix, based on the error regression, as $\boldsymbol{R}_E = \boldsymbol{C}_E^{-1}[\boldsymbol{E} - \boldsymbol{B}_E\boldsymbol{Z}]$. So, we can write the error matrix, $\boldsymbol{E}$, as a function of $\boldsymbol{B}_E$ and $\boldsymbol{C}_E$ in the following way:

$$\boldsymbol{E} = \boldsymbol{B}_E\boldsymbol{Z} + \boldsymbol{C}_E\boldsymbol{R}_E \text{ and hence } \boldsymbol{E}\boldsymbol{E}' = \boldsymbol{B}_E\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{B}_E' + \boldsymbol{S}_E \tag{3.2}$$

since $\boldsymbol{R}_E\boldsymbol{R}_E' = I_p$, inner product of two orthonormal matrices, and $\boldsymbol{Z}\boldsymbol{R}_E' = \boldsymbol{0}$, as $\boldsymbol{Z}$ and $\boldsymbol{R}_E$ are orthogonal matrices. In the next Section, we define similar regression matrix and residual sum of squares matrix for the future regression model.

For the multivariate simple regression model, the density of the error matrix in (2.5) can be written as

$$f(\boldsymbol{E}) = [2\pi]^{-\frac{pn}{2}} e^{-\frac{1}{2}tr\{\boldsymbol{E}\boldsymbol{E}'\}}. \tag{3.3}$$

The abode density represents a matrix normal distribution, that is, $\boldsymbol{E}$ follows a $p \times n$ dimensional matrix normal distribution with location $\boldsymbol{0}$ and covariance matrix $I_p \bigotimes I_n$. Each row (or column) of $\boldsymbol{E}$ independently follows a multivariate normal distribution of appropriate dimension.

To find the joint distribution of $\boldsymbol{B}_E$ and $\boldsymbol{S}_E$ from the distribution of the error matrix $\boldsymbol{E}$ we note the following differential relation (cf. Fraser 1979, p.114 or Eaton, 1983, p.194-204)

$$d\boldsymbol{E} = |\boldsymbol{S}_E|^{\frac{n-p-2-1}{2}} d\boldsymbol{B}_E d\boldsymbol{S}_E d\boldsymbol{R}_E. \tag{3.4}$$

This invariant differential relation allows us to derive the joint distribution of the error statistics $\boldsymbol{B}_E$ and $\boldsymbol{S}_E$ from the joint distribution of the error matrix $\boldsymbol{E}$.

3.1 Distribution of $\boldsymbol{B}_E$ and $\boldsymbol{S}_E$

Applying the relation in (3.2) and the differential in (3.4), from the above join density of the realized error matrix, the density function of $\boldsymbol{B}_E$ and $\boldsymbol{S}_E$, conditional on $\boldsymbol{R}_E(\cdot)$, becomes

$$f(\boldsymbol{B}_E, \boldsymbol{S}_E|\boldsymbol{R}_E(\cdot)) \propto |\boldsymbol{S}_E|^{\frac{n-p-2-1}{2}} e^{-\frac{1}{2}tr\{\boldsymbol{B}_E\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{B}_E' + \boldsymbol{S}_E\}}. \tag{3.5}$$

Since the above density does not depend on $\boldsymbol{R}_E(\cdot)$, the conditional distribution is the same as the unconditional distribution. Also, the joint density of $\boldsymbol{B}_E$ and $\boldsymbol{S}_E$ factors and hence $\boldsymbol{B}_E$ and $\boldsymbol{S}_E$ are independently distributed. The marginal distribution of the error regression matrix is

$$f(\boldsymbol{B}_E)) = [2\pi]^{-\frac{2p}{2}} \left|\boldsymbol{Z}\boldsymbol{Z}'\right|^{-1} e^{-\frac{1}{2}tr\{\boldsymbol{B}_E\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{B}_E'\}}. \tag{3.6}$$

Thus $\boldsymbol{B}_E$ follows a $p \times 2$ dimensional matrix variate normal distribution with location matrix $\boldsymbol{0}$ and variance-covariance matrix $[\boldsymbol{Z}\boldsymbol{Z}']^{-1} \bigotimes I_2$ in which $\bigotimes$ is the Kronecker product of two matrices and I_2 is the unit matrix of order 2. Note that each column of $\boldsymbol{B}_E$ follows a p-variate normal distribution. Similarly, the marginal distribution of the residual sum of squares matrix of the error regression is

$$f(\boldsymbol{S}_E) = \frac{1}{[2]^{\frac{(n-2)p}{2}} \Gamma_p\left(\frac{n-2}{2}\right)} |\boldsymbol{S}_E|^{\frac{n-p-2-1}{2}} e^{-\frac{1}{2}tr\{\boldsymbol{S}_E\}} \tag{3.7}$$

where $\Gamma_p(a)$ is the generalized gamma function defined as

$$\Gamma_p(a) = [\pi]^{\frac{a(a-1)}{4}} \prod_{i=1}^{a} \Gamma\left(a - \frac{1}{2}[i-1]\right). \tag{3.8}$$

Clearly $\boldsymbol{S}_E$ follows a p-dimensional Wishart distribution with $(n-2)$ degrees of freedom and variance I_p.

3.2 Distribution of $\boldsymbol{B}_Y$ and $\boldsymbol{S}_Y$

To find the distribution of the sample regression matrix (SRM), $\boldsymbol{B}_Y$, and sample residual sum of squares matrix (SRSSM), $\boldsymbol{S}_Y$, of the response regression we use the relations (cf. Fraser, 1979, p.115)

$$\boldsymbol{B}_E = \Gamma^{-1}(\boldsymbol{B}_Y - \boldsymbol{\beta}), \text{ and } \boldsymbol{S}_E = \boldsymbol{\Sigma}^{-1}\boldsymbol{S}_Y, \tag{3.9}$$

where $\boldsymbol{B}_Y = \boldsymbol{Y}\boldsymbol{Z}'(\boldsymbol{Z}\boldsymbol{Z}')^{-1}$ and $\boldsymbol{S}_Y = [\boldsymbol{Y} - \boldsymbol{B}_Y\boldsymbol{Z}][\boldsymbol{Y} - \boldsymbol{B}_Y\boldsymbol{Z}]'$ are the sample regression matrix of $\boldsymbol{Y}$ on $\boldsymbol{Z}$, and the residual sum of squares matrix of the regression based on the realized responses respectively. It may be mentioned here that both $\boldsymbol{S}_E$ and $\boldsymbol{S}_Y$ have the same structure from the definitions of $\boldsymbol{S}_E$ in (3.1) and that of $\boldsymbol{S}_Y$ above. It can easily be shown that $\boldsymbol{R}_E = \boldsymbol{C}_Y^{-1}[\boldsymbol{Y} - \boldsymbol{B}_Y\boldsymbol{Z}] = \boldsymbol{R}_Y$ where $\boldsymbol{C}_Y$ is such that $\boldsymbol{S}_Y = \boldsymbol{C}_Y\boldsymbol{C}_Y'$.

Now using the relations in (3.9) along with the associated differentials

$$d\boldsymbol{B}_E = |\boldsymbol{\Sigma}^{-1}|d\boldsymbol{B}_Y \text{ and } d\boldsymbol{S}_E = |\boldsymbol{\Sigma}^{-1}|^{\frac{p+1}{2}}d\boldsymbol{S}_Y \tag{3.10}$$

the joint density function of $\boldsymbol{B}_Y$ and $\boldsymbol{S}_Y$ becomes

$$f(\boldsymbol{B}_Y, \boldsymbol{S}_Y) \propto |\boldsymbol{\Sigma}|^{-\frac{n}{2}} e^{-\frac{1}{2}tr\boldsymbol{\Sigma}^{-1}[(\boldsymbol{B}_Y - \boldsymbol{\beta})\boldsymbol{Z}\boldsymbol{Z}'(\boldsymbol{B}_Y - \boldsymbol{\beta})' + \boldsymbol{S}_Y]}. \tag{3.11}$$

Once again, for the normal model under study, the above joint distribution of $\boldsymbol{B}_Y$ and $\boldsymbol{S}_Y$ can be factored, and hence the two statistics, $\boldsymbol{B}_Y$ and $\boldsymbol{S}_Y$, are independently distributed. However, this is not so for the same statistics based on the unobserved future responses, conditional on realized responses (cf. Khan, 2004). The marginal distributions of $\boldsymbol{B}_Y$ and $\boldsymbol{S}_Y$ are given, respectively, by

$$f(\boldsymbol{B}_Y) = [2\pi]^{-\frac{2p}{2}} |\boldsymbol{\Sigma}|^{-1} |\boldsymbol{Z}\boldsymbol{Z}'|^{\frac{p}{2}} e^{-\frac{1}{2}tr\boldsymbol{\Sigma}^{-1}[(\boldsymbol{B}_Y - \boldsymbol{\beta})\boldsymbol{Z}\boldsymbol{Z}'(\boldsymbol{B}_Y - \boldsymbol{\beta})']} \tag{3.12}$$

$$f(\boldsymbol{S}_Y) = \frac{1}{[2]^{\frac{(n-2)p}{2}} \Gamma_p(\frac{n-2}{2}) |\boldsymbol{\Sigma}|^{\frac{n-2}{2}}} |\boldsymbol{S}_Y|^{\frac{n-p-2-1}{2}} e^{-\frac{1}{2}tr\boldsymbol{\Sigma}^{-1}[\boldsymbol{S}_Y]}. \tag{3.13}$$

Therefore, $\boldsymbol{B}_Y$ follows a $p \times 2$ dimensional matrix normal distribution with location matrix $\boldsymbol{\beta}$ and scale matrices $\boldsymbol{\Sigma}$ and $[\boldsymbol{Z}\boldsymbol{Z}']^{-1}$, and independently $\boldsymbol{S}_Y$ follows a p-variate Wishart distribution with $(n-2)$ degrees of freedom and variance $\boldsymbol{\Sigma}$. However, as it is shown later in the paper, the prediction distribution of the future regression matrix and the future residual sum of squares matrix, conditional on the realized responses, are not independently distributed.

4 The Future Model

In this Section we introduce the multivariate simple regression model for the unrealized future responses from the *future experiment*, and use both the realized sample and unrealized future sample to derive the prediction distribution of the FRM and FRSSM. First, consider a set of n_f unrealized future responses, $\boldsymbol{Y}_f = (\boldsymbol{y}_{f1}, \boldsymbol{y}_{f2}, \cdots, \boldsymbol{y}_{fn_f})$, from the same multivariate simple regression model in (2.1) with the same regression and scale parameter matrices as defined in Section 2. Such a set of future responses can be expressed as

$$\boldsymbol{Y}_f = \boldsymbol{\beta}\boldsymbol{Z}_f + \Gamma\boldsymbol{E}_f \tag{4.1}$$

where $\boldsymbol{Z}_f$ is a $2\times n_f$ dimensional matrix of the values of regressors that generate the $p\times n_f$ dimensional future response matrix $\boldsymbol{Y}_f$, and $\boldsymbol{E}_f$ is the matrix of future error components. Similar to the error matrix of the realized model the future error matrix from the *future experiment* follows independent and identical normal distribution. Our aim is to find the prediction distribution of the FRM and FRSSM, conditional on the realized responses, by the classical method as well as by the Bayesian method under the uniform prior.

Assuming that the future errors follow the same matrix normal distribution as the realized errors, the density function of the future error matrix is given by

$$f(\boldsymbol{E}_f) = [2\pi]^{-\frac{pn_f}{2}} e^{-\frac{1}{2}tr\{\boldsymbol{E}_f\boldsymbol{E}_f'\}}. \tag{4.2}$$

From the specifications of the model, the future sample is independent of the realized sample. Thus the join density function of the combined error matrix, that is, the errors associated with the realized and that of the future responses, $(\boldsymbol{E}, \boldsymbol{E}_f)$, can be expressed as

$$f(\boldsymbol{E}, \boldsymbol{E}_f) = [2\pi]^{-\frac{p(n+n_f)}{2}} e^{-\frac{1}{2}tr\{\boldsymbol{E}\boldsymbol{E}'+\boldsymbol{E}_f\boldsymbol{E}_f'\}}. \tag{4.3}$$

This joint density function of the combined errors is used to derive the prediction distributions of the FRM and FRSSM based on a set of future responses of the future model in the next Section.

5 Predictive Distribution of FRM and FRSSM

In this Section we derive the predictive distributions of the future regression matrix (FRM) and future residual sum of squares matrix (FRSSM) based on the future multivariate simple regression model, conditional on the realized responses.

The joint density function of the error statistics $\boldsymbol{B}_E$, $\boldsymbol{S}_E$, $\boldsymbol{B}_{E_f}$ and $\boldsymbol{S}_{E_f}$, given $\boldsymbol{R}(\cdot)$, is derived from the above joint density of the combined error matrix in (4.3) by applying the properties of invariant differentials (see Eaton, 1983, p.194-206) as follows:

$$p\Big(\boldsymbol{B}_E, \boldsymbol{S}_E, \boldsymbol{B}_{E_f}, \boldsymbol{S}_{E_f} | \boldsymbol{R}(\cdot)\Big) \propto |\boldsymbol{S}_E|^{\frac{n-p-2-1}{2}} |\boldsymbol{S}_{E_f}|^{\frac{n_f-p-2-1}{2}} \times e^{-\frac{1}{2}tr\{h_1(\boldsymbol{B}_E,\boldsymbol{Z})+h_2(\boldsymbol{B}_f,\boldsymbol{Z}_f)\}} \quad (5.1)$$

where $h_1(\boldsymbol{B}_E, \boldsymbol{Z}) = \boldsymbol{B}_E\boldsymbol{Z}\boldsymbol{Z}'\boldsymbol{B}'_E$; $h_2(\boldsymbol{B}_f, \boldsymbol{Z}_f) = \boldsymbol{B}_{E_f}\boldsymbol{Z}_f\boldsymbol{Z}'_f\boldsymbol{B}'_{E_f}$. Note that the above joint density does not depend on $\boldsymbol{R}(\cdot)$ (cf. Fraser, 1968, p.132) so the conditional distribution is the same as the unconditional distribution. Here $\boldsymbol{B}_E$, $\boldsymbol{S}_E$, $\boldsymbol{B}_{E_f}$ and $\boldsymbol{S}_{E_f}$ are independently distributed. So, like $\boldsymbol{B}_E$, $\boldsymbol{B}_{E_f}$ follows a matrix normal distribution. Similarly, $\boldsymbol{S}_{E_f}$ follows a Wishart distribution.

The joint distribution of $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{B}_{E_f}$, and $\boldsymbol{S}_{E_f}$ is then obtained by using the Jacobian of the transformation,

$$J\Big\{[\boldsymbol{B}_E, \boldsymbol{S}_E] \to [\boldsymbol{\beta}, \boldsymbol{\Sigma}^{-1}]\Big\} \propto |\boldsymbol{\Sigma}^{-1}|, \quad (5.2)$$

as follows

$$p\Big(\boldsymbol{\beta}, \boldsymbol{\Sigma}^{-1}, \boldsymbol{B}_{E_f}, \boldsymbol{S}_{E_f}\Big) \propto |\boldsymbol{S}|^{\frac{n-p-2-1}{2}} |\boldsymbol{S}_{Y_f}|^{\frac{n_f-p-2-1}{2}} |\boldsymbol{\Sigma}^{-1}|^{\frac{n-p-1}{2}} e^{-\frac{1}{2}tr\left[\boldsymbol{\Sigma}^{-1}\{\xi(\boldsymbol{B}_Y,\boldsymbol{\beta})+\boldsymbol{S}\}+\eta(\boldsymbol{B}_{E_f},\boldsymbol{\beta})+\boldsymbol{S}_{E_f}\right]} \quad (5.3)$$

where $\xi(\boldsymbol{B}_Y, \boldsymbol{\beta}) = (\boldsymbol{B}_Y - \boldsymbol{\beta})\boldsymbol{Z}\boldsymbol{Z}'(\boldsymbol{B}_Y - \boldsymbol{\beta})'$; $\eta(\boldsymbol{B}_{E_f}, \boldsymbol{\beta}) = \boldsymbol{B}_{E_f}\boldsymbol{Z}_f\boldsymbol{Z}'_f\boldsymbol{B}'_{E_f}$; $\boldsymbol{B} = \boldsymbol{B}_Y$ and $\boldsymbol{S} = \boldsymbol{S}_Y$.

5.1 Prediction Distribution of $\boldsymbol{B}_{Y_f}$

We are interested in the distributions of $\boldsymbol{B}_{Y_f}$ and $\boldsymbol{S}_{Y_f}$, the future regression matrix and future residual sum of squares matrix for the future regression, respectively, conditional on the realized responses.

To derive the joint distribution of $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{B}_{Y_f}$, and $\boldsymbol{S}_{Y_f}$ from the above joint density of $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{B}_{E_f}$, and $\boldsymbol{S}_{E_f}$, note that from the structure of the future regression equation we have

$$\boldsymbol{B}_{E_f} = \boldsymbol{\Sigma}^{-\frac{1}{2}}[\boldsymbol{B}_{Y_f} - \boldsymbol{\beta}] \text{ and } \boldsymbol{S}_{E_f} = \boldsymbol{\Sigma}^{-1}\boldsymbol{S}_{Y_f} \quad (5.4)$$

where $\boldsymbol{B}_{Y_f} = \boldsymbol{Y}\boldsymbol{Z}'_f(\boldsymbol{Z}_f\boldsymbol{Z}'_f)^{-1}$, $\boldsymbol{S}_{Y_f} = [\boldsymbol{Y}_f - \boldsymbol{B}_{Y_f}\boldsymbol{Z}'_f][\boldsymbol{Y}_f - \boldsymbol{B}_{Y_f}\boldsymbol{Z}_f]'$. Therefore, the Jacobian of the transformations become

$$J\Big\{[\boldsymbol{B}_{E_f}, \boldsymbol{S}_{E_f}] \to [\boldsymbol{B}_{Y_f}, \boldsymbol{S}_{Y_f}]\Big\} = |\boldsymbol{\Sigma}^{-1}|^{\frac{p+2+1}{2}}. \quad (5.5)$$

So, the joint density of $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{B}_{Y_f}$ and $\boldsymbol{S}_{Y_f}$ is obtained as

$$p\Big(\boldsymbol{\beta}, \boldsymbol{\Sigma}^{-1}, \boldsymbol{B}_f, \boldsymbol{S}_f\Big) \propto |\boldsymbol{S}|^{\frac{n-p-2-1}{2}} |\boldsymbol{S}_{Y_f}|^{\frac{n_f-p-2-1}{2}} |\boldsymbol{\Sigma}|^{-\frac{n+n_f-p-1}{2}} e^{-\frac{1}{2}tr\left\{\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{Q}_Y+\boldsymbol{Q}_{Y_f}+\boldsymbol{S}+\boldsymbol{S}_f\right]\right\}} \quad (5.6)$$

where $\boldsymbol{Q}_Y = (\boldsymbol{B}_Y - \boldsymbol{\beta})\boldsymbol{Z}\boldsymbol{Z}'(\boldsymbol{B}_Y - \boldsymbol{\beta})'$, $\boldsymbol{Q}_{Y_f} = (\boldsymbol{B}_f - \boldsymbol{\beta})\boldsymbol{Z}_f\boldsymbol{Z}_f'(\boldsymbol{B}_f - \boldsymbol{\beta})'$, $\boldsymbol{B}_f = \boldsymbol{B}_{Y_f}$ and $\boldsymbol{S}_f = \boldsymbol{S}_{Y_f}$ for notational convenience.

Such results can also be obtained by using the Bayesian approach (see Section 6). In particular, the Bayes posterior density can be obtained by assuming uniform prior for the regression and scale parameters of the model. However, the final results of this Section will be the same as that obtained by the Bayesian approach under uniform prior. Interested readers may refer to Fraser and Haq (1969) for further details.

To find the prediction distribution of the FRM and FRSSM we need to integrate out $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}^{-1}$ from the above joint density in (5.6). First integrating out $\boldsymbol{\Sigma}^{-1}$ from the above joint density, the marginal density function of $\boldsymbol{\beta}$, $\boldsymbol{B}_{Y_f}$ and $\boldsymbol{S}_{Y_f}$ can be written as

$$p\Big(\boldsymbol{\beta}, \boldsymbol{B}_f, \boldsymbol{S}_f\Big) \propto \left[\boldsymbol{S}_{Y_f}\right]^{\frac{n_f-p-2-1}{2}} \int_{\boldsymbol{\Sigma}^{-1}} |\boldsymbol{\Sigma}^{-1}|^{\frac{n+n_f-p-1}{2}} e^{-\frac{1}{2}tr\boldsymbol{\Sigma}^{-1}\left\{\boldsymbol{Q}_Y + \boldsymbol{S} + \boldsymbol{Q}_{Y_f} + \boldsymbol{S}_f\right\}} d\boldsymbol{\Sigma}^{-1}. \tag{5.7}$$

Note that $\boldsymbol{\Sigma}^{-1}$ follows a p variate Wishart distribution, that is, $\boldsymbol{\Sigma}^{-1} \sim W_p(n + n_f,\ \boldsymbol{Q}_Y + \boldsymbol{S} + \boldsymbol{Q}_{Y_f} + \boldsymbol{S}_f)$. Completion of the integration leads to the join density function of $\boldsymbol{\beta}$, $\boldsymbol{B}_{Y_f}$ and $\boldsymbol{S}_{Y_f}$ to be

$$p\Big(\boldsymbol{\beta}, \boldsymbol{B}_f, \boldsymbol{S}_f\Big) \quad \propto \quad \left|\boldsymbol{S}_{Y_f}\right|^{\frac{n_f-p-2-1}{2}} \left|\boldsymbol{Q} + \boldsymbol{S} + \boldsymbol{S}_f\right|^{-\frac{n+n_f}{2}} \tag{5.8}$$

in which $\boldsymbol{Q} = \boldsymbol{Q}_Y + \boldsymbol{Q}_{Y_f}$. From the above density function, we derive the joint distribution of the FRM and FRSSM, conditional on the realized responses, by integrating out $\boldsymbol{\beta}$. To facilitate such an integration, the terms involving the regression parameter matrix $\boldsymbol{\beta}$ in $\boldsymbol{Q}$ can be expressed as follows:

$$\begin{aligned} \boldsymbol{Q} &= (\boldsymbol{B} - \boldsymbol{\beta})\boldsymbol{Z}\boldsymbol{Z}'(\boldsymbol{B} - \boldsymbol{\beta})' + (\boldsymbol{B}_f - \boldsymbol{\beta})\boldsymbol{Z}_f\boldsymbol{Z}_f'(\boldsymbol{B}_f - \boldsymbol{\beta})' \\ &= (\boldsymbol{\beta} - \boldsymbol{F}\boldsymbol{A}^{-1})\boldsymbol{A}(\boldsymbol{\beta} - \boldsymbol{F}\boldsymbol{A}^{-1})' + (\boldsymbol{B}_f - \boldsymbol{B})\boldsymbol{H}^{-1}(\boldsymbol{B}_f - \boldsymbol{B})' \end{aligned} \tag{5.9}$$

where

$$\boldsymbol{F} = \boldsymbol{B}\boldsymbol{Z}\boldsymbol{Z}' + \boldsymbol{B}_f\boldsymbol{Z}_f\boldsymbol{Z}_f',\ \boldsymbol{A} = \boldsymbol{Z}\boldsymbol{Z}' + \boldsymbol{Z}_f\boldsymbol{Z}_f',\ \text{and } \boldsymbol{H} = [\boldsymbol{Z}\boldsymbol{Z}']^{-1} + [\boldsymbol{Z}_f\boldsymbol{Z}_f']^{-1}. \tag{5.10}$$

The marginal density of $\boldsymbol{B}_f$ and $\boldsymbol{S}_f$ is derived by using the representation in (5.9) and integrating out $\boldsymbol{\beta}$ from the joint density in (5.8). Thus, we have

$$\begin{aligned} p\Big(\boldsymbol{B}_f, \boldsymbol{S}_f\Big) \quad \propto \quad & |\boldsymbol{S}_f|^{\frac{n_f-p-2-1}{2}} \int_{\boldsymbol{\beta}} \big|\boldsymbol{S} + \boldsymbol{S}_f + (\boldsymbol{B}_f - \boldsymbol{B})\boldsymbol{H}^{-1}(\boldsymbol{B}_f - \boldsymbol{B})' \\ & + (\boldsymbol{\beta} - \boldsymbol{F}\boldsymbol{A}^{-1})\boldsymbol{A}(\boldsymbol{\beta} - \boldsymbol{F}\boldsymbol{A}^{-1})'\big|^{-\frac{n+n_f}{2}} d\boldsymbol{\beta}. \end{aligned} \tag{5.11}$$

Note that $\boldsymbol{\beta}$ follows $p \times 2$ dimensional matrix T distribution, that is,

$$\boldsymbol{\beta} \sim T_{p\times 2}\left[n + n_f - p - 1, \boldsymbol{F}\boldsymbol{A}^{-1}, \boldsymbol{A}, \boldsymbol{S} + \boldsymbol{S}_f + (\boldsymbol{B}_f - \boldsymbol{B})\boldsymbol{H}^{-1}(\boldsymbol{B}_f - \boldsymbol{B})'\right]. \tag{5.12}$$

On completion of the integration using the matrix T integral, the marginal density of $\boldsymbol{B}_f$ and $\boldsymbol{S}_f$ is obtained as

$$p\Big(\boldsymbol{B}_f, \boldsymbol{S}_f\Big) = \Psi_{12} \times |\boldsymbol{S}_f|^{\frac{n_f-p-2-1}{2}} \times \left|\boldsymbol{S} + \boldsymbol{S}_f + (\boldsymbol{B}_f - \boldsymbol{B})\boldsymbol{H}^{-1}(\boldsymbol{B}_f - \boldsymbol{B})'\right|^{-\frac{n+n_f-p-1}{2}} \quad (5.13)$$

where $\Psi_{12} = \Big\{|\boldsymbol{H}|^{-\frac{p}{2}}\Gamma_p(\frac{n+n_f-p-1}{2})|\boldsymbol{S}|^{\frac{n-p-1}{2}}\Big\}\Big\{(\pi)^{\frac{2p}{2}}\Gamma_p(\frac{n-p-1}{2})\Gamma_p(\frac{n_f-2}{2})\Big\}^{-1}$ is the normalizing constant. This is the joint prediction distribution of the FRM and FRSSM of the future responses. Unlike the join density of SRM, $\boldsymbol{B}_Y$, and SRSSM, $\boldsymbol{S}_Y$, in (3.11), the above join density can't be factored, and hence $\boldsymbol{B}_f$ and $\boldsymbol{S}_f$ are dependent.

The prediction distribution of the future regression matrix, $\boldsymbol{B}_f = \boldsymbol{B}_{Y_f}$, can now be obtained by integrating out $\boldsymbol{S}_f$ from (5.13). The integration yields

$$p\Big(\boldsymbol{B}_f\Big|\boldsymbol{Y}\Big) = \Psi_1 \times \left|\boldsymbol{S} + (\boldsymbol{B}_f - \boldsymbol{B})\boldsymbol{H}^{-1}(\boldsymbol{B}_f - \boldsymbol{B})'\right|^{-\frac{n+m-p-1}{2}} \quad (5.14)$$

where $\Psi_1 = \Psi_{12} \times B_p^{-1}\Big(\frac{n_f-2}{2}, \frac{n-p-1}{2}\Big)$. On simplification the normalizing constant becomes $\Psi_1 = \{\Gamma_p(\frac{n+m-p-1}{2})|\boldsymbol{S}|^{\frac{n-p-1}{2}}\}\{(\pi)^{\frac{2p}{2}}\Gamma_p(\frac{n-p-1}{2})|\boldsymbol{H}|^{\frac{p}{2}}\}^{-1}$. The prediction distribution of $\boldsymbol{B}_f$ can be written in the usual matrix T distribution form as follows:

$$p\Big(\boldsymbol{B}_f\Big|\boldsymbol{Y}\Big) = \Psi_6 \times \left|I_p + (\boldsymbol{B}_f - \boldsymbol{B})[\boldsymbol{S}\boldsymbol{H}]^{-1}(\boldsymbol{B}_f - \boldsymbol{B})'\right|^{-\frac{n+2-p-1}{2}}. \quad (5.15)$$

Since the density in (5.15) is a matrix T density, the prediction distribution of the future regression matrix, $\boldsymbol{B}_f$, conditional on the realized responses, follows a multivariate matrix T distribution of dimension $p \times 2$, and with $(n-2p)$ degrees of freedom. Thus, $[\boldsymbol{B}_f|\boldsymbol{Y}] \sim T_p\,(n-2p, \boldsymbol{B}, \boldsymbol{H}, \boldsymbol{S})$ where $\boldsymbol{B}$ is the sample regression matrix of realized responses and $\boldsymbol{H}$ and $\boldsymbol{S}$ are the scale matrices. It is observed that the degrees of freedom parameter of the prediction distribution of the future regression matrix $\boldsymbol{B}_f$ depends on the size of the realized sample and the dimension of the regression parameter matrix of the model. Khan (2001) obtained a similar result for the multiple regression model with normal errors.

5.2 Prediction Distribution of $\boldsymbol{S}_{Y_f}$

The prediction distribution of the FRSSM for the future regression model, $\boldsymbol{S}_{Y_f}$, conditional on the realized responses, $\boldsymbol{Y}$, is obtained by integrating out $\boldsymbol{B}_f$ from (5.13). Since $\boldsymbol{B}_f$ follows a matrix T distribution, using the matrix T integral the prediction distribution of the FRSSM becomes

$$p\Big(\boldsymbol{S}_{Y_f}\Big|\boldsymbol{Y}\Big) \propto \Psi_2 \times |\boldsymbol{S}_{Y_f}|^{\frac{n_f-p-2-1}{2}}\left|\boldsymbol{S} + \boldsymbol{S}_{Y_f}\right|^{-\frac{n+n_f-p-m-1}{2}}. \quad (5.16)$$

The density function in (5.16) can be written in the usual form of the generalized beta distribution as follows

$$p\Big(\boldsymbol{S}_f\Big|\boldsymbol{Y}\Big) = \Psi_2 \times [\boldsymbol{S}_f]^{\frac{n_f-p-2-1}{2}}\left|I_p + \boldsymbol{S}^{-1}\boldsymbol{S}_f\right|^{-\frac{n+n_f-p-m-1}{2}} \quad (5.17)$$

where $\Psi_2 = \left\{\Gamma_p(\frac{n+n_f-p-m-1}{2})|\boldsymbol{S}|^{-\frac{n-p-1}{2}}\right\}\left\{\Gamma_p(\frac{n-p-1}{2})\Gamma_p(\frac{n_f-2}{2})\right\}^{-1}$ is the normalizing constant. This is the prediction distribution of the FRSSM based on the future responses, $\boldsymbol{Y}_f$, conditional on the realized responses, $\boldsymbol{Y}$, from the multivariate simple regression model with normal errors. The density in (5.16) is a modified form of the generalized beta density. However, it can be shown that $\boldsymbol{S}^{-1}\boldsymbol{S}_f$ is a generalized beta variable with arguments $(n_f-2)/2$ and $(n-p-1)/2$. Obviously, for the existence of the above prediction distribution of $\boldsymbol{S}_f$ we must have $n_f > 2$ in addition to $n > p+1$. Khan (2001) and Khan (2004) obtained a similar prediction distribution of the FRSSM, conditional on the realized responses, for the multiple regression model with multivariate normal and matrix T errors respectively.

6 The Bayesian Approach

In this Section we consider the prediction distributions of the FRM and FRSSM under the Bayesian approach. Here we assume that the joint prior distribution of the regression matrix and inverse of the variance-covariance matrix is uniform. Such a prior is due to Jeffreys (1961) and many scholars have used this prior for numerous studies (see for example Bernardo and Rueda, 2002 and the references there in). Thus we adopt the following prior distribution

$$p\left(\boldsymbol{\beta}, \boldsymbol{\Sigma}^{-1}\right) \propto \left|\boldsymbol{\Sigma}^{-1}\right|^{\frac{p+1}{2}}. \tag{6.1}$$

Interested readers may refer to Press (1989, p.134) or Rowe (2003, p.41) for details. From (5.1) in the previous Section the joint density function of the error statistics $\boldsymbol{B}_E$, $\boldsymbol{S}_E$, $\boldsymbol{B}_{E_f}$ and $\boldsymbol{S}_{E_f}$ is

$$p\Big(\boldsymbol{B}_E, \boldsymbol{S}_E, \boldsymbol{B}_{E_f}, \boldsymbol{S}_{E_f}\Big) \propto |\boldsymbol{S}_E|^{\frac{n-p-2-1}{2}} \left|\boldsymbol{S}_{E_f}\right|^{\frac{n_f-p-2-1}{2}} \times e^{-\frac{1}{2}tr\left\{h_1(\boldsymbol{B}_E,\boldsymbol{Z})+h_2(\boldsymbol{B}_{E_f},\boldsymbol{Z}_f)\right\}}. \tag{6.2}$$

The inherent relation of the model in (3.9) yields the Jacobian of the transformation,

$$J\Big\{[\boldsymbol{B}_E, \boldsymbol{S}_E] \to [\boldsymbol{\beta}, \boldsymbol{\Sigma}^{-1}]\Big\} = \left|\boldsymbol{\Sigma}^{-1}\right|, \tag{6.3}$$

so the joint distribution of $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{B}_{E_f}$, and $\boldsymbol{S}_{E_f}$ becomes

$$p\Big(\boldsymbol{\beta}, \boldsymbol{\Sigma}^{-1}, \boldsymbol{B}_{E_f}, \boldsymbol{S}_{E_f}\Big) \propto |\boldsymbol{S}|^{\frac{n-2}{2}}|\boldsymbol{S}_{Y_f}|^{\frac{n_f-p-1}{2}}|\boldsymbol{\Sigma}^{-1}|^{\frac{n+n_f-2}{2}} e^{-\frac{1}{2}tr\boldsymbol{\Sigma}^{-1}\{\xi(\boldsymbol{B},\boldsymbol{\beta})+\boldsymbol{S}+\eta(\boldsymbol{B}_{E_f})+\boldsymbol{S}_{E_f}\}} \tag{6.4}$$

where $\xi(\boldsymbol{B},\boldsymbol{\beta}) = (\boldsymbol{B}-\boldsymbol{\beta})\boldsymbol{Z}\boldsymbol{Z}'(\boldsymbol{B}-\boldsymbol{\beta})'$; $\eta(\boldsymbol{B}_{E_f}) = \boldsymbol{B}_{E_f}\boldsymbol{Z}_f\boldsymbol{Z}_f'\boldsymbol{B}_{E_f}$; $\boldsymbol{B} = \boldsymbol{B}_Y$ and $\boldsymbol{S} = \boldsymbol{S}_Y$.

Similarly, for the future model the inherent relation in (5.4) yields the Jacobian of the transformations,

$$J\Big\{[\boldsymbol{B}_{E_f},\boldsymbol{S}_{E_f}]\rightarrow[\boldsymbol{\beta},\boldsymbol{\Sigma}^{-1}]\Big\}=\left|\boldsymbol{\Sigma}^{-1}\right|. \tag{6.5}$$

Therefore, the joint distribution of $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{B}_{Y_f}$, and $\boldsymbol{S}_{Y_f}$ is obtained as

$$\begin{aligned} p\Big(\boldsymbol{\beta},\boldsymbol{\Sigma}^{-1},\boldsymbol{B}_{Y_f},\boldsymbol{S}_{Y_f}\Big) \quad \propto \quad & \left|\boldsymbol{S}_{Y_f}\right|^{\frac{n_f-p-2-1}{2}}\left|\boldsymbol{\Sigma}^{-1}\right|^{\frac{n+n_f-2p-2}{2}} \\ & e^{-\frac{1}{2}tr\boldsymbol{\Sigma}^{-1}\{\xi(\boldsymbol{B},\boldsymbol{\beta})+\boldsymbol{S}+\eta(\boldsymbol{B}_{E_f},\boldsymbol{\beta})+\boldsymbol{S}_{E_f}\}}. \end{aligned} \tag{6.6}$$

Then incorporating the prior distribution in (6.1) for the parameters of the multivariate regression model the joint posterior distribution of $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}^{-1}$, $\boldsymbol{B}_{Y_f}$, and $\boldsymbol{S}_{Y_f}$ is obtained as follows

$$\begin{aligned} p\Big(\boldsymbol{\beta},\boldsymbol{\Sigma}^{-1},\boldsymbol{B}_{f},\boldsymbol{S}_{Y_f}\Big) \quad \propto \quad & \left|\boldsymbol{S}_{Y_f}\right|^{\frac{n_f-p-2-1}{2}}\left|\boldsymbol{\Sigma}^{-1}\right|^{\frac{n+n_f-p-1}{2}} \\ & e^{-\frac{1}{2}tr\boldsymbol{\Sigma}^{-1}\{\xi(\boldsymbol{B},\boldsymbol{\beta})+\boldsymbol{S}+\zeta(\boldsymbol{B}_f,\boldsymbol{\beta})+\boldsymbol{S}_{Y_f}\}} \end{aligned} \tag{6.7}$$

where $\zeta(\boldsymbol{B}_f,\boldsymbol{\beta})=(\boldsymbol{B}_f-\beta)\boldsymbol{Z}\boldsymbol{Z}'(\boldsymbol{B}_f-\beta)'$ in which $\boldsymbol{B}_f=\boldsymbol{B}_{Y_f}$ and $\boldsymbol{S}_f=\boldsymbol{S}_{Y_f}$.

Now, as in the previous Section, integration of the above density with respect to $\boldsymbol{\Sigma}^{-1}$, gives the joint distribution of β, $\boldsymbol{B}_f$ and $\boldsymbol{S}_f$. Then using the same representation of $\Big[\xi(\boldsymbol{B},\beta)+\zeta(\boldsymbol{B}_f,\boldsymbol{\beta})\Big]$ as for $\boldsymbol{Q}$ in (5.9) we integrate out $\boldsymbol{\beta}$ to obtain the joint density function of the FRM and FRSSM as follows

$$\begin{aligned} p\Big(\boldsymbol{B}_f,\boldsymbol{S}_f\Big) \quad = \quad & \Psi_{12}\times[\boldsymbol{S}_f]^{\frac{n_f-p-2-1}{2}} \\ & \times\left|\boldsymbol{S}+\boldsymbol{S}_f+(\boldsymbol{B}_f-\boldsymbol{B})H^{-1}(\boldsymbol{B}_f-\boldsymbol{B})'\right|^{-\frac{n+n_f-p-1}{2}} \end{aligned} \tag{6.8}$$

where $\Psi_{12}=\Big\{|\boldsymbol{H}|^{-\frac{1}{2}}\Gamma_p(\frac{n+n_f-p-1}{2})|\boldsymbol{S}|^{\frac{n-p-1}{2}}\Big\}\Big\{(\pi)^{\frac{2p}{2}}\Gamma_p(\frac{n-p-1}{2})\Gamma_p(\frac{n_f-2}{2})\Big\}^{-1}$ is the normalizing constant. This is the same distribution as obtained in (5.13) by the classical approach. Following the same procedures as in the previous Section we get the prediction distribution of FRM to be $p\times 2$-variate matrix T and that of the FRSSM to be scaled generalized beta distribution. Thus the Bayesian method with uniform prior produces the same prediction distributions for the FRM and FRSSM as produced by the classical method for the multivariate simple regression model with normal errors.

7 Concluding Remarks

The paper derives the prediction distribution of the FRV and FRSSM for the multivariate simple regression model with correlated normal responses. The distributions of the FRM and FRSSM obtained by both the classical and Bayesian methods are the same. The foregoing analyses reveal the fact that for the multivariate simple regression model with normal errors the distribution of the SRM is matrix variates normal, and it is independent of SRSSM. But the predictive distribution of the FRV and FRSSM are not independently distributed. Khan (2004) showed that the above statistics are not independently distributed for the multiple regression model with multivariate normal as well as multivariate Student-t errors. Furthermore, since the predictive distribution of the future regression matrix for the normal model is a matrix T distribution, each column of $\boldsymbol{B}_{Y_f}$ follows a multivariate Student-t distribution, and the components of the random vectors are dependent. Also, the shape parameter of the prediction distribution of the future regression matrix depends on the size of the realized sample as well as the number of regression parameter of the multivariate simple regression model.

The distribution of the SRSSM, $\boldsymbol{S}_Y$, is a Wishart distribution. But, the prediction distribution of the future residual sum of squares matrix (FRSSM) of the future regression model, conditional on the realized responses, follows a scaled generalized beta distribution.

Since the simple regression model is a special case of the multivariate simple regression model when $p = 1$, the results in this paper generalizes those for the simple regression model. In other words, setting $p = 1$ in this paper, the corresponding results for the simple regression model are observed. Thus for the simple regression model with normal error the prediction distribution of the future regression vector (intercept and slope) is a bi-variate Student-t distribution and that of the future residual sum of squares is a beta distribution.

Acknowledgements

The author is grateful to Professor A K Md Ehsanes Saleh, Carleton University, Canada for some valuable suggestions on an earlier version of the paper.

References

[1] J. Aitchison, (1964). *Bayesian tolerance regions*, J. of Royal Statistical Society, B, **127**, 161-175.

[2] J. Aitchison and I. R. Dunsmore, (1975). Statistical Prediction Analysis, Cambridge University Press, Cambridge.

[3] J. Aitchison and D. Sculthorpe, (1965). *Some problems of statistical prediction*, Biometrika, *55*, 469-483.

[4] J. Bernardo and R. Rueda, (2002). *Bayesian hypothesis testing: A reference approach*, International Statistical Review. Vol.70, 351-372.

[5] M. L. Eaton, (1983). Multivariate Statistics - Vector Space Approach, Wiley, New York.

[6] D. A. S. Fraser, (1968). The Structure of Inference, Wiley, New York.

[7] D. A. Fraser, (1979). Inference and Linear Models, McGraw-Hill, New York

[8] D. A. Fraser and M. S. Haq, (1969). *Structural probability and prediction for the multilinear model*, J. of Royal Statistical Society, B, **31**, 317-331.

[9] S. Geisser, (1993). Predictive Inference: An Introduction, Chapman & Hall, London.

[10] I. Guttman, (1970). Statistical Tolerance Regions, Griffin, London.

[11] M. S. Haq, (1982). *Structural relation and prediction for multivariate models*, Statistische Hefte, **23**, 218-228.

[12] M. S. Haq and S. Khan, (1990). *Prediction distribution for a linear regression model with multivariate Student-t error distribution*, Communications In Statistics- Theory and Methods, **19** (12), 4705-4712.

[13] M. S. Haq and S. Rinco, (1976). *β-expectation tolerance regions for a generalized multilinear model with normal error variables*, J. Multiv. Analysis, **6**, 414-21.

[14] H. Jeffreys, (1961). Theory of Probability, 3rd ed, Oxford, Clarendon Press

[15] S. Khan, (2005). *Estimation of parameters of the simple multivariate linear model with Student-t errors*, Journal of Statistical Research, **39**, 79-94.

[16] S. Khan, (2004). *Predictive distribution of regression vector and residual sum of squares for normal multiple regression model*, Communications In Statistics: Theory and Methods, Vol. 33(10), 2423-2443.

[17] S. Khan, (2002). *Distribution of sum of squares and products for the generalized multilinear matrix-T model*, J. of Multiv. Analysis, **83**, 124-140.

[18] S. Khan, (2001). *Predictive Distribution of Regression Vector and Residual Sum of Squares for Normal Multiple Regression Model*, Working Paper Series, No. SC-MC-0125, Faculty of Sciences, University of Southern Queensland, Australia.

[19] S. Khan and M. S. Haq, (1994). *Prediction inference for multilinear model with errors having multivariate Student-t distribution and first-order autocorrelation structure*, Sankhya, Part B: The Indian Journal of Statistics, **56**, 95-106.

[20] S. J. Press, (1989). Bayesian Statistics: Pronciples, Models, and Applications. Wiley, New York.

[21] C. Ren, D. Sun, and D. K. Dey, (2004). *Comparison of Bayesian and Frequentist estimation and prediction for normal population*, Sankhya: The Indian Journal of Statistics, **66(4)**, 678-706.

[22] D. B. Rowe, (2003). Multivariate Bayesian Statistics. Chapman and Hall/CRC, London.

[23] A. K. Md E. Saleh, (2006). Theory of Preliminary Test and Stein-type Estimators with Application. Wiley, Toronto.