Søren Johansen

Asymptotic Inference in Random Coefficient Regression Models

Preprint
January
1981

1

Institute of Mathematical Statistics
University of Copenhagen

Søren Johansen

ASYMPTOTIC INFERENCE IN RANDOM COEFFICIENT REGRESSION MODELS

Preprint 1981 No. 1

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

JANUARY 1981

Abstract

In a random coefficient linear regression model which is not balanced it is shown how one can make asymptotic inference on the parameters when the number of observations tends to infinity in such a way that the variance due to the design stays bounded.

1. Introduction and main result

Consider the model for a random coefficient regression with normal errors

$$Y_i = X_i B_i + U_i, \quad i = 1, \ldots, n$$

where $(B_1, \ldots, B_n, U_1, \ldots, U_n)$ are independent normally distributed in such a way that

$$B_i \sim N_m(\beta, \Sigma) \text{ and } U_i \sim N_{p_i}(0, \sigma_i^2 I) .$$

We shall assume that the design matrix X_i has rank m for all i, and that $p_i > m$, $i = 1, \ldots, n$.

It follows that

$$Y_i \sim N_{p_i}(X_i \beta,\ X_i \Sigma X_i' + \sigma_i^2 I), \ i = 1, \ldots, n ,$$

are independent.

We want to estimate β, Σ and $\sigma_1^2, \ldots, \sigma_n^2$, and find the asymptotic distribution of the estimate for β.

Let P_i denote the projection of R^{p_i} onto the subspace spanned by X_i, then for $i = 1, \ldots, n$,

$$(I - P_i) Y_i \sim N_{p_i}(0, \sigma_i^2 (I - P_i))$$

and

$$P_i Y_i \sim N_{p_i}(X_i \beta,\ X_i \Sigma X_i' + \sigma_i^2 P_i)$$

are independent.

The variable $P_i Y_i$ is in one-to-one correspondence with $\hat{\beta}_i = (X_i' X_i)^{-1} X_i' Y_i$, the estimate of the "individual coefficients".

The distribution of $\hat{\beta}_i$ is given by $N_m(\beta, \Sigma + \sigma_i^2(X_i'X_i)^{-1})$.

Thus the likelihood splits into a product of two factors, one containing σ_i^2 and one derived from $\hat{\beta}_i$ which contains Σ, β and σ_i^2.

If σ_i^2 and Σ are known then it is easily seen that the maximum likelihood for β is given by

$$\hat{\beta}(\Sigma, \sigma^2) = (\Sigma_1^n w_i)^{-1} \Sigma_1^n w_i \hat{\beta}_i \ ,$$

where

$$w_i = (\Sigma + \sigma_i^2(X_i'X_i)^{-1})^{-1}, \ i = 1, \ldots, n \ .$$

Swamy (1971) suggested using the estimates

$$\sigma_{i*}^2 = |(I - P_i)Y_i|^2/(p_i - m), \ i = 1, \ldots, n$$

and

$$\Sigma_* = SSD(\hat{\beta})/(n-1) - \Sigma_1^n \sigma_{i*}^2 (X_i'X_i)^{-1}/n$$

which are easily seen to be unbiased for σ_i^2 and Σ respectively.

Thus we use the residual sum of squares from each regression to estimate the variance within each experiment. We then find the empirical variance matrix of $\hat{\beta}_i$ and correct it to be unbiased for Σ. Note that we take the $SSD(\hat{\beta})$ around $\bar{\beta} = \frac{1}{n} \Sigma_1^n \hat{\beta}_i$.

Swamy now defines

$$\beta_* = \hat{\beta}(\Sigma_*, \sigma_*^2)$$

and shows consistency and asymptotic normality under the assumption

A. $$\sqrt{n} \sup_{1 \le i \le n} \sigma_i^2 \operatorname{tr}(X_i'X_i)^{-1} \to 0, n \to \infty \ .$$

Now $\hat{\beta}_i$ has a variance composed of two components, the population variance Σ and the design variance $\sigma_i^2(X_i'X_i)^{-1}$. The condition A ensures that the design variance disappears in the limit, and hence that the estimator β becomes asymptotically equivalent to $\bar{\beta}$, but then there may not be so much reason for weighting after all. Often $X_i'X_i$ will be of the order of p_i and hence the condition is that $\sqrt{n} \sup_{1 \leq i \leq n} \sigma_i^2 p_i^{-1} \to 0$, that is, p_i should tend to infinity much faster than $\sqrt{n}$.

For the applications it is certainly of interest to derive a limit result without this severe restriction. One would often have many experimental units, (n large) each giving rise to an unbalanced design (X_i) which is not necessarily very large (p_i bounded). Thus a reasonable condition is that

B. $$\sup_{1 \leq i \leq n} \sigma_i^2 \operatorname{tr}(X_i'X_i)^{-1} \leq C < \infty .$$

We can now prove the following

<u>Theorem 1</u> The asymptotic distribution of β_* when $n \to \infty$, such that B is satisfied is given by

$$N_m(\beta, (\Sigma_1^n E\tilde{w}_i)^{-1}(\Sigma_1^n E\tilde{w}_i w_i^{-1}\tilde{w}_i)(\Sigma_1^n E\tilde{w}_i)^{-1})$$

with

$$\tilde{w}_i^{-1} = \Sigma + \sigma_{i*}^2(X_i'X_i)^{-1} .$$

For this to be useful one must have a consistent estimate of the asymptotic variance of β_*.

<u>Theorem 2</u> A consistent estimate of the asymptotic variance of β_* is given by

$$(\Sigma_1^n w_i^*)^{-1}(\Sigma_1^n w_i^{**})(\Sigma_1^n w_i^*)^{-1}$$

where

$$w_i^{**} = w_i^* \Sigma_* w_i^* +$$

$$\int_0^1 (\Sigma_* + z\sigma_{i*}^2(X_i'X_i)^{-1})^{-1} \sigma_{i*}^2(X_i'X_i)^{-1}(\Sigma_* + z\sigma_{i*}^2(X_i'X_i)^{-1})^{-1} \lambda_i z^{\lambda_i - 1} dz$$

where

$$\lambda_i = (p_i - m)/2 \text{ and } w_i^{*-1} = \Sigma_* + \sigma_{i*}^2(X_i'X_i)^{-1} .$$

Note that if $p_i \to \infty$ sufficiently fast then $\sigma_{i*}^2 \overset{P}{\to} \sigma_i^2$ and the asymptotic variance will be equivalent to $(\Sigma_1^n w_i)^{-1}$ which again is equivalent to $n^{-1}\Sigma$.

If $\sigma_i^2 = \sigma^2$ for $i = 1,\ldots,n$, then σ^2 can be estimated consistently and $\Sigma_1^n(\Sigma_* + \sigma_*^2(X_i'X_i)^{-1})^{-1}$ will estimate the asymptotic variance consistently.

The estimator in Theorem 2 for the asymptotic variance is not easy to calculate. Some different expressions are given in connection with the proof. A different estimator is given in

<u>Theorem 3</u> If $p_i > m + 2$, $i = 1,\ldots,n$, then a consistent estimator of the asymptotic variance of β_* is given by

$$(\Sigma_1^n w_i^*)^{-1} \Sigma_1^n w_i^{***} (\Sigma_1^n w_i^*)^{-1}$$

where

$$w_i^{***} = w_i^* \Sigma_* w_i^* + w_i^+ \sigma_{i+}^2 (X_i'X_i)^{-1} w_i^+ (1 - \lambda_i^{-1})$$

and

$$\sigma_{i+}^2 \sim \sigma_i^2 \chi^2/f, \quad f = p_i - m - 2 = 2(\lambda_i - 1)$$

and

$$w_i^+ = (\Sigma_* + \sigma_{i+}^2 (X_i' X_i)^{-1} (1 - \lambda_i^{-1}))^{-1} \quad .$$

Thus by discarding two degrees of freedom from each of the variances estimates σ_{i*}^2, one can produce another simpler estimate.

2. Proof of the results

We shall repeatedly use the ordering of positive definite symmetric matrices given by the usual definition: If ϕ and Σ are positive definite symmetric $m \times m$ matrices then $\phi > \Sigma$ (or $\Sigma < \phi$) if $\phi - \Sigma$ is positive definite.

We shall need the following results:

$$\phi < I \operatorname{tr} \phi$$

$$\phi \to \phi^{-1} \text{ is convex and decreasing}$$

$$\lambda \to (\Sigma + \lambda\phi)^{-1}(\Sigma + \lambda_0\phi)(\Sigma + \lambda\phi)^{-1}$$
$$\text{is convex and decreasing}$$

The convexity results are easily proved by diagonalizing Σ and ϕ simultaneously by some non-singular transformation.

We shall also employ the L_2 norm of an $m \times m$ matrix A given by

$$|A|^2 = \sup_x \frac{|Ax|^2}{|x|^2} = \sup_x \frac{x'A'Ax}{x'x} = \lambda_{max}(A'A),$$

the maximal eigenvalue. Note that $|\phi| = \lambda_{max}(\phi)$ for a positive definite symmetric ϕ.

We apply these notions to find that under condition B we have

$$\Sigma < w_i^{-1} = \Sigma + \sigma_i^2(X_i'X_i)^{-1} < \Sigma + CI$$

and hence $\Sigma^{-1} > w_i > (\Sigma + CI)^{-1}$.

Thus for a fixed Σ, the weights as well as the variances lie in a compact set.

We need the following simple result: for $\bar{\beta} = \frac{1}{n}\Sigma_1^n \hat{\beta}_i$ and

$SS(\hat{\beta}-\beta) = \Sigma_1^n(\hat{\beta}_i-\beta)(\hat{\beta}_i-\beta)'$ we have

$$E|\bar{\beta}-\beta| \in O(1/\sqrt{n})$$

$$E|SS(\hat{\beta}-\beta) - E\ SS(\hat{\beta}-\beta)| \in O(\sqrt{n}) \quad .$$

This follows easily by calculating the variances

$$V(\bar{\beta}-\beta) = \frac{1}{n^2}\ \Sigma_1^n\ w_i^{-1}$$

and

$$V(SS(\hat{\beta}-\beta)) = \frac{1}{n^2}\ \Sigma_1^n\ w_i^{-1} \otimes w_i^{-1}$$

both of which are $O(1/n)$ by the remarks above.

We shall use the following notation: for a sequence $\{X_n\}$ of random variables we write $X_n \in O_P(1)$ if the sequence is tight, that is if for all $\varepsilon > 0$ there exists a constant C such that $P\{|X_n| \geq C\} \leq \varepsilon$ for all n.

If $\{b_n\}$ is a sequence of real numbers we write $X_n \in O_P(b_n)$ if $b_n^{-1}\ X_n \in O_P(1)$. It is easily seen that if $E|X_n|$ or $V(X_n)$ is bounded then it follows from Chebychev's inequality that $X_n \in O_P(1)$. Similarly if $X_n \overset{P}{\to} a$, then also $X_n \in O_P(1)$.

<u>Lemma 2.1</u> Under condition B, the estimator Σ_* is consistent for Σ, in fact $\Sigma_* - \Sigma \in O_P(1/\sqrt{n})$.

<u>Proof</u> It is seen that

$$\begin{aligned}\sqrt{n}(\Sigma_* - \Sigma) &= \frac{\sqrt{n}}{n-1}\ (SS(\hat{\beta}-\beta) - E\ SS(\hat{\beta}-\beta)) \\ &- \frac{n\sqrt{n}}{n-1}\ (\bar{\beta}-\beta)(\bar{\beta}-\beta)' \\ &- \frac{\sqrt{n}}{n}\ \Sigma_1^n(\sigma_{i*}^2 - \sigma_i^2)(X_i'X_i')^{-1} \quad .\end{aligned}$$

Now the first and second terms are $\in O_P(1)$ and the third is evaluated by its variance

$$V\left(\frac{\sqrt{n}}{n}\Sigma_1^n(\sigma_{i*}^2-\sigma_i^2)(X_i'X_i)^{-1}\right)=\frac{2}{n}\Sigma_1^n\frac{\sigma_i^4}{p_i-m}(X_i'X_i)^{-1}\otimes(X_i'X_i)^{-1} .$$

This is contained in $O(1)$ under condition B which completes the proof.

<u>Lemma 2.2</u> Under condition B we have

$$\sup_i |w_i^* - \tilde{w}_i| \in O_P(1/\sqrt{n})$$

and

$$\left(\frac{1}{n}\Sigma_1^n w_i^*\right)^{-1} - \left(\frac{1}{n}\Sigma_1^n E\tilde{w}_i\right)^{-1} \in O_P(1/\sqrt{n}) .$$

<u>Proof</u> We have

$$w_i^* - \tilde{w}_i = w_i^*(\Sigma - \Sigma_*)\tilde{w}_i .$$

For the first factor we have the inequality

$$w_i^* < \Sigma_*^{-1}$$

and for the third

$$\tilde{w}_i < \Sigma^{-1}$$

which shows that they are both in $O_P(1)$ uniformly in i. The second factor is $O_P(1/\sqrt{n})$ which proves the first result.

The second result is proved as follows:

$$\Delta_1 = \left(\frac{1}{n}\Sigma_1^n w_i^*\right)^{-1} - \left(\frac{1}{n}\Sigma_1^n\tilde{w}_i\right)^{-1}$$

$$= \left(\frac{1}{n}\Sigma_1^n w_i^*\right)^{-1}\left(\frac{1}{n}\Sigma_1^n(\tilde{w}_i - w_i^*)\right)\left(\frac{1}{n}\Sigma_1^n\tilde{w}_i\right)^{-1} .$$

Now the second factor is, by the above, $\in O_P(1/\sqrt{n})$. We have to show that the first and third factors are bounded.

From the convexity of $\phi \to \phi^{-1}$ we have

$$(\frac{1}{n}\ \Sigma_1^n w_i^*)^{-1} < \frac{1}{n}\ \Sigma_1^n w_i^{*-1} < \Sigma_* + \frac{1}{n}\ \Sigma_1^n \frac{\sigma_{i*}^2}{\sigma_i^2}\ CI$$

and

$$(\frac{1}{n}\ \Sigma_1^n \tilde{w}_i)^{-1} < \frac{1}{n}\ \Sigma_1^n \tilde{w}_i^{-1} < \Sigma + \frac{1}{n}\ \Sigma_1^n \frac{\sigma_{i*}^2}{\sigma_i^2}\ CI$$

which shows that both factors are $\in O_P(1)$, since $E\ \frac{1}{n}\ \Sigma_1^n \frac{\sigma_{i*}^2}{\sigma_i^2} = 1$, hence $\Delta_1 \in O_P(1/\sqrt{n})$.

Next consider

$$\begin{aligned}\Delta_2 &= (\frac{1}{n}\ \Sigma_1^n \tilde{w}_i)^{-1} - (\frac{1}{n}\ \Sigma_1^n E\,\tilde{w}_i)^{-1} \\ &= -\ (\frac{1}{n}\ \Sigma_1^n \tilde{w}_i)^{-1}(\frac{1}{n}\ \Sigma_1^n (\tilde{w}_i - E\,\tilde{w}_i))(\frac{1}{n}\ \Sigma_1^n E\,\tilde{w}_i)^{-1}\ .\end{aligned}$$

Again the first factor is bounded in probability.

The third is evaluated as follows. From the convexity of $\phi \to \phi^{-1}$ it follows that

$$\frac{1}{n}\ \Sigma_1^n \tilde{w}_i > (\Sigma + \frac{1}{n}\ \Sigma_1^n \sigma_{i*}^2\,(X_i'X_i)^{-1})^{-1}$$

and again

$$\begin{aligned}E\ \frac{1}{n}\ \Sigma_1^n \tilde{w}_i &> (\Sigma + \frac{1}{n}\ \Sigma_1^n \sigma_i^2 (X_i'X_i)^{-1})^{-1} \\ &> (\Sigma + CI)\end{aligned}$$

which shows that the third factor is bounded.

Finally the second factor is evaluated by its variance

$$\frac{1}{n^2}\ \Sigma_1^n\ V(\tilde{w}_i) \in O(1/n)$$

which shows that $\Delta_2 \in O(1/\sqrt{n})$. Thus $\Delta_1 + \Delta_2 \in O(1/\sqrt{n})$ which proves the last assertion of Lemma 2.2.

Lemma 2.3 Under condition B we have that

$$\Sigma_1^n \tilde{w}_i (\hat{\beta}_i - \beta)$$

is asymptotically normally distributed with parameters 0 and $\Sigma_1^n E\tilde{w}_i w_i^{-1} \tilde{w}_i$.

Proof We are considering a sum of independent random variables $Z_i = \tilde{w}_i (\hat{\beta}_i - \beta)$.

We find

$$E\ Z_i = E\ \tilde{w}_i\ E\ (\hat{\beta}_i - \beta) = 0$$

$$V\ Z_i = E\ \tilde{w}_i\ V(\hat{\beta}_i)\ \tilde{w}_i = E\ \tilde{w}_i w_i^{-1} \tilde{w}_i$$

$$E|Z_i|^3 \leq |\Sigma^{-1}|^3\ E|\hat{\beta}_i - \beta|^3$$

Hence

$$\Sigma_1^n E|Z_i|^3 \leq |\Sigma^{-1}|^3\ \Sigma_1^n E|\hat{\beta}_i - \beta|^3 \in O(n)$$

and

$$\Sigma_1^n\ V(Z_i) = \Sigma_1^n E\tilde{w}_i w_i^{-1} \tilde{w}_i\ .$$

By the convexity of the function

$$s^2 \to (\Sigma + s^2 (X'X)^{-1})^{-1} (\Sigma + \sigma^2 (X'X)^{-1}) (\Sigma + s^2 (X'X)^{-1})^{-1}$$

it follows that

$$E\ \tilde{w}_i w_i^{-1} \tilde{w}_i > w_i > (\Sigma + CI)^{-1}$$

and hence

$$\Sigma_1^n V(Z_i) > n(\Sigma + CI)^{-1} .$$

Thus we can check Ljapunov's condition

$$\frac{(\Sigma_1^n E|Z_i|^3)^2}{|\Sigma_1^n V(Z_i)|^3} \leq \frac{an^2}{n^3|(\Sigma + CI)^{-1}|} \to 0, \quad n \to \infty$$

which proves Lemma 2.3.

We can now combine the results and prove Theorem 1.

Let $\Sigma_{**} = \frac{1}{n} SS(\hat{\beta} - \beta) - \frac{1}{n} \Sigma_1^n \sigma_{i*}^2 (X_i'X_i)^{-1}$ and note that $\Sigma_{**} - \Sigma \in O_P(1/\sqrt{n})$ and that $\Sigma_{**} - \Sigma_* =$
$\frac{1}{n-1} (\frac{1}{n} SS(\hat{\beta} - \beta) - n(\overline{\beta} - \beta)(\overline{\beta} - \beta)') \in O_P(1/n)$.

Now

$$\Sigma_1^n w_i^* (\hat{\beta}_i - \beta) = \Sigma_1^n \tilde{w}_i (\hat{\beta}_i - \beta) - \Sigma_1^n (\tilde{w}_i - w_i^{**})(\hat{\beta}_i - \beta)$$
$$- \Sigma_1^n (w_i^{**} - w_i^*)(\hat{\beta}_i - \beta) .$$

The first term is by Lemma 2.3 asymptotically distributed $N(0, \Sigma_1^n E \tilde{w}_i w_i^{-1} \tilde{w})$ and hence $\in O_P(\sqrt{n})$. We want to prove that the other terms are $O_P(1)$.

The last term is evaluated as follows

$$|\Sigma_1^n (w_i^* - w_i^{**})(\hat{\beta}_i - \beta)|$$
$$\leq |\Sigma_1^n w_i^* (\Sigma_* - \Sigma_{**}) w_i^{**} (\hat{\beta}_i - \beta)|$$
$$\leq \sup_i |w_i^*| \sup_i |w_i^{**}| \sup_i |\hat{\beta}_i - \beta| n |\Sigma_* - \Sigma_{**}|$$
$$\in O_P(1)$$

The second term which contains the essence of all the difficulties is evaluated by its variance. We easily find

$$E\{(\hat{\beta}_i - \beta) \mid SS(\hat{\beta} - \beta)\} = 0$$

$$V\{(\hat{\beta}_i - \beta, \hat{\beta}_j - \beta) \mid SS(\hat{\beta} - \beta)\} = 0, \; i \neq j$$

and

$$|V\{(\hat{\beta}_i - \beta) \mid SS(\hat{\beta} - \beta)\}| \leq E\{|\hat{\beta}_i - \beta|^2 \mid SS(\hat{\beta} - \beta)\} \; .$$

Hence for $R = \Sigma_1^n (w_i^{**} - \tilde{w}_i)(\hat{\beta}_i - \beta)$ we get $ER = 0$ and

$$V(R) = E\{V(R) \mid SS(\hat{\beta} - \beta), \sigma_*^2)\}$$

$$= E \; \Sigma_1^n (w_i^{**} - \tilde{w}_i) V\{(\hat{\beta}_i - \beta) \mid SS(\hat{\beta} - \beta)\} (w_i^{**} - \tilde{w}_i)$$

which shows that

$$|V(R)| \leq E \sup_i |w_i^{**} - \tilde{w}_i|^2 \; tr(SS(\hat{\beta} - \beta)) \in O_P(1) \; .$$

Thus we find that

$$\Sigma_1^n w_i^* (\hat{\beta}_i - \beta) = \Sigma_1^n \tilde{w}_i (\hat{\beta}_i - \beta) + \varepsilon$$

with $\varepsilon \in O_P(1)$.

Finally we write

$$\beta_* - \beta = (\frac{1}{n} \; \Sigma_1^n w_i^*)^{-1} \; \frac{1}{n} \; \Sigma_1^n \; w_i^* (\hat{\beta}_i - \beta)$$

$$= ((\frac{1}{n} \; \Sigma_1^n \, E\tilde{w}_i)^{-1} + \varepsilon_1) \; \frac{1}{n} (\Sigma_1^n \tilde{w}_i (\hat{\beta}_i - \beta) + \varepsilon)$$

with $\varepsilon_1 \in O_P(1/\sqrt{n})$ from Lemma 2.2. Hence

$$\beta_* - \beta = (\Sigma_1^n \, E\tilde{w}_i)^{-1} \; \Sigma_1^n \tilde{w}_i (\hat{\beta}_i - \beta) \; + \; \varepsilon_2 \; ,$$

where $\varepsilon_2 \in O_P(1/n)$ which is the desired conclusion and the proof of Theorem 1 is completed.

Next we prove Theorem 2. From the above results we have

$$(\frac{1}{n}\,\Sigma_1^n w_i^*)^{-1} - (\frac{1}{n}\,\Sigma_1^n E\,\tilde{w}_i)^{-1} \overset{P}{\to} 0$$

and similarly

$$\frac{1}{n}\,\Sigma_1^n w_i^*\;\Sigma_* w_i^* - \frac{1}{n}\,\Sigma_1^n E\,\tilde{w}_i\,\Sigma\tilde{w}_i \quad \overset{P}{\to} 0$$

we see that it is enough to prove that

$$E\int_0^1 (\Sigma + z\sigma_{i*}^2(X_i'X_i)^{-1})^{-1}\sigma_{i*}^2(X_i'X_i)^{-1}(\Sigma + z\sigma_{i*}^2(X_i'X_i)^{-1})^{-1}\lambda_i z^{\lambda_i - 1}dz$$

$$= E\;\tilde{w}_i\;\sigma_i^2(X_i'X_i)^{-1}\tilde{w}_i\;,$$

since then $\frac{1}{n}\,\Sigma_1^n\;w_i^{**} - \frac{1}{n}\,\Sigma_1^n E\,\tilde{w}_i\;w_i^{-1}\tilde{w}_i \overset{P}{\to} 0.$

The expression in the integral contains the two matrices Σ and $(X_i'X_i)^{-1}$. We diagonalize them simultaneously such that Σ becomes the identity and $(X_i'X_i)^{-1}$ the diagonal matrix $\mathrm{diag}(d_1,\ldots,d_m)$.

We then have to verify that

$$E\int_0^1 (1+zV)^{-2}\;V\lambda z^{\lambda-1}dz = E\;(1+V)^{-2}\;\lambda\beta$$

where $V = d_k\sigma_{i*}^2$ is Gamma distributed with parameters $\lambda = \lambda_i$ and $\beta = \sigma_i^2\;d_k/\lambda_i$ here $i = 1,\ldots,n$ and $k = 1,\ldots,m$.

To prove this relation we interchange the integration as follows:

$$E\int_0^1 (1+zV)^{-2}\;V\lambda z^{\lambda-1}\;dz$$

$$= \int_0^\infty\int_0^1 (1+zv)^{-2}\;v\lambda z^{\lambda-1}\;v^{\lambda-1}e^{-v/\beta}dv/(\Gamma(\lambda)\,\beta^\lambda)$$

$$= \int_0^\infty\int_0^v (1+u)^{-2}\;u^{\lambda-1}\;du\,\lambda\,e^{-v/\beta}dv/(\Gamma(\lambda)\,\beta^\lambda)$$

$$= \int_0^\infty (1+u)^{-2}\;\lambda\beta\,e^{-u/\beta}\,u^{\lambda-1}du/(\Gamma(\lambda)\,\beta^\lambda)$$

$$= E(1+V)^{-2}\;\lambda\beta\;.$$

This completes the proof of Theorem 2. In order that the result be useful one has to calculate the integral $\int_0^1 (1+zx)^{-2} z^{\lambda-1} dz$. We have found the following two expressions

$$\int_0^1 (1+zx)^{-2} z^{\lambda-1} dz = (1+x)^{-1} - (\lambda-1) \sum_{\mu=0}^{\infty} \frac{(-x)^{\mu}}{\mu+\lambda}$$

$$= (1+x)^{-1} + (\lambda-1)(-x)^{-\lambda} \{\ln(1+x) + \sum_{\nu=1}^{\lambda-1} (-x)^{\nu}/\nu\}.$$

Finally we want to prove Theorem 3.

Just as with Theorem 2 we only have to prove that

$$(1-\lambda_i^{-1})\, E\, w_i^{+} \sigma_{i+}^{2} (X_i' X_i)^{-1} w_i^{+} = E\, \tilde{w}_i \sigma_i^2 (X_i' X_i)^{-1} \tilde{w}_i \quad .$$

Diagonalizing as before this result reduces to

$$(1-\lambda^{-1})\, E\{U(1+U(1-\lambda^{-1}))^{-2}\} = \lambda\beta\, E\,(1+V)^{-2}$$

where V, λ and β has the same meaning as before, and U is distributed as $\Gamma(\lambda-1, \beta\lambda/(\lambda-1))$. This relation follows from

$$\lambda\beta \int_0^{\infty} (1+v)^{-2}\, v^{\lambda-1}\, e^{-v/\beta} dv$$

$$= \lambda\beta \int_0^{\infty} v(1+v)^{-2}\, v^{\lambda-2}\, e^{-v/\beta} dv$$

$$= \lambda\beta \int_0^{\infty} u(1+u(1-\lambda^{-1}))^{-2}\, u^{\lambda-2}\, e^{-\frac{u(1-\lambda^{-1})}{\beta}} du\,(1-\lambda^{-1})^{\lambda}$$

by dividing by $\Gamma(\lambda)\beta^{\lambda}$.

Acknowledgement

It is a pleasure to thank Steffen Lauritzen for a helpful comment and discussions that resulted in considerable simplification of the arguments.

References

Swamy, P.A.V.B.: (1971) Statistical Inference in Random Coefficient Regression Models. Lecture Notes in Oper. Research and Math. Syst., Springer.

PREPRINTS 1980

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Olkin, Ingram and Væth, Michael: Maximum Likelihood Estimation in a Two-way Analysis of Variance with Correlated Errors in One Classification.

No. 2 Skovgaard, Ib M.: Transformation of an Edgeworth Expansion by a Sequence of Smooth Functions.

No. 3 Asmussen, Søren: Equilibrium Properties of the M/G/1 Queue.

No. 4 Johansen, Søren and Keiding, Susanne: A Family of Models for the Elimination of Substrate in the Liver.

No. 5 Skovgaard, Ib M.: Edgeworth Expansions of the Distributions of Maximum Likelihood Estimators.

No. 6 Tjur, Tue: A Connection Between Rasch's Item Analysis Model and a Multiplicative Poisson Model.

No. 7 Lauritzen, Steffen L. : Examples of Extreme Point Models in Statistice.

PREPRINTS 1981

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Johansen, Søren: Asymptotic Inference in Random Coefficient Regression Models.