Capítulo 7

El modelo de regresión lineal múltiple univariante. Estimación por máxima verosimilitud.

7.1. Introducción.

En el tema anterior hemos obtenido el estimador mínimo cuadrático ordinario de los parámetros del modelo así como ciertas propiedades del mismo. También se ha analizado el caso de funciones estimables, culminando con el teorema de Gauss-Markov. Para todo ello no ha sido necesario suponer ninguna forma concreta sobre la distribución de los términos de perturbación, bastando con asegurar que el proceso de los errores está formado por variables incorreladas de media cero y varianza finita. Sin embargo, aunque hemos demostrado propiedades como la insesgadez y eficiencia, no hemos podido determinar propiedades sobre sus distribuciones ni, por lo tanto, establecer contrastes de significación. Vamos a dar a continuación una hipótesis adicional sobre la distribución de los errores y a determinar qué consecuencias se producen en los estimadores.

La hipótesis sobre la que estamos hablando es la siguiente

$$\epsilon_i \rightsquigarrow \mathbf{N}_1\left[0;\sigma^2\right] \quad ; \quad i = 1,\ldots,N$$

siendo todas las variables ϵ_i independientes.

La introducción de esta hipótesis abre la posibilidad de emplear el método de máxima verosimilitud para estimar los parámetros del modelo.

7.2. Estimación por máxima verosimilitud.

Notemos que la hipótesis de normalidad puede ser expresada para las observaciones de la variable dependiente. En efecto, para cada conjunto fijo de valores $x_{i1},\ldots,x_{ik}$, la distribución de la variable y_i es normal con media $\beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik}$ y varianza σ^2. Además dichas distribuciones son también independientes, por lo que es inmediato expresar la función de verosimilitud asociada. Llamando $\beta = (\beta_0,\ldots,\beta_k)'$ e $\mathbf{y} = (y_1,\ldots,y_N)'$, tendremos

$$\mathbb{L}(\mathbf{y};\beta,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}\exp\left(-\frac{1}{2\sigma^2}(\mathbf{y}-\mathbf{X}\beta)'(\mathbf{y}-\mathbf{X}\beta)\right) =$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i-\beta_0-\beta_1 x_{i1}-\cdots-\beta_k x_{ik})^2\right) \ .$$

El logaritmo de la función de verosimilitud será

$$\log(\mathbb{L}(y_1, \ldots, y_N; \beta_0, \beta_1, \cdots, \beta_k, \sigma^2)) =$$

$$= -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\sigma^2) - -\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik})^2 \ .$$

Derivando respecto a los parámetros e igualando las parciales a cero obtendremos el mismo sistema de ecuaciones normales obtenido en la estimación por mínimos cuadrados ordinarios.

Al igual que ocurría en regresión simple, el estimador máximo verosímil de la varianza no coincide con el mínimo cuadrático

$$\widetilde{\sigma}^2 = \frac{\sum_{i=1}^{N} e_i^2}{N} = \frac{e'e}{N} \ .$$

No obstante, dicho estimador no es centrado, por lo que recurriremos en lo que sigue al estimador mínimo cuadrático

$$\widehat{\sigma}^2 = \frac{\sum_{i=1}^{N} e_i^2}{N-k-1} = \frac{e'e}{N-k-1} \ .$$

A continuación vamos a obtener la distribución de los estimadores obtenidos. Para ello nos es preciso conocer algunas cuestiones referentes a distribuciones multivariantes normales y formas cuadráticas de variables normales.

El hecho de introducir estas cuestiones se debe a simplificar, lo más posible, todo el cálculo y las expresiones que surgen, al mismo tiempo que motiva toda una filosofía concerniente a la inferencia en distribuciones multidimensionales.

7.3. Breves notas sobre la normal multivariante.

Un vector aleatorio p-dimensional $\mathbf{x} = (x_1, \cdots, x_p)'$ se dice que se distribuye atendiendo a una normal p-dimensional de parámetros $\mu_{p\times 1}$ y $\Sigma_{p\times p}$, donde Σ es una matriz real y definida positiva, si la densidad correspondiente es de la forma

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\right) \qquad \mathbf{x} = (x_1, \cdots, x_p)\,.$$

En particular, si $p = 1$ se tiene $\Sigma = \sigma^2$ y se recae en la definición de normal unidimensional

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)\,.$$

Existe toda una teoría sobre la normal multivariante, en la cual no vamos a entrar, pero a continuación citaremos algunas de sus propiedades más importantes puesto que después serán utilizadas.

Resultado 1. *Si Σ es diagonal, entonces las variables aleatorias x_i, $i = 1, \ldots, p$ son independientes y se distribuyen según una normal univariante $N_1\left[\mu_i; \sigma_i^2\right]$, $i = 1, \ldots, p$.* [1]

Resultado 2. *Sea $\mathbf{x}$ un vector p-dimensional distribuido según una normal $N_p\left[\mu; \Sigma\right]$. Sea $\mathbf{A}_{q\times p}$ una matriz con rango q $(q \leq p)$. Entonces $\mathbf{y} = \mathbf{A}\mathbf{x}$ se distribuye según una normal q-dimensional $N_q\left[\mathbf{A}\mu; \mathbf{A}\Sigma\mathbf{A}'\right]$.*

Resultado 3. *Toda combinación lineal de las componentes de un vector aleatorio normal p-dimensional es una variable normal unidimensional. En particular, las componentes del vector son normales unidimensionales.*

Resultado 4. *(Caracterización de la ley normal). La condición necesaria y suficiente para que un vector aleatorio se distribuya de forma normal multivariante es que cualquier combinación lineal de sus componentes sea una variable aleatoria unidimensional que se distribuya según una ley normal univariante.*

Resultado 5. *Cualquier vector $\mathbf{x}$ p-dimensional normal con matriz de varianzas-covarianzas Σ no singular puede convertirse, mediante una transformación lineal, en un vector $\mathbf{z}$ normal p-dimensional con vector de medias cero y matriz de varianzas-covarianzas igual a la identidad (normal estándar o tipificada).*

[1] Observemos que si introdujéramos las hipótesis sobre el vector de perturbación aleatoria en la forma $\varepsilon \rightsquigarrow \mathrm{N}_N\left[\mathbf{0}; \sigma^2\mathbf{I}_N\right]$ o sobre el vector $\mathbf{y}$ como $\mathbf{y} \rightsquigarrow \mathrm{N}_N\left[\mathbf{X}\beta; \sigma^2\mathbf{I}_N\right]$ no haría falta imponer la independencia puesto que el resultado 1 nos proporcionaría dicha condición.

$$f(\mathbf{z}) = \frac{1}{(2\pi)^{\frac{p}{2}}} \exp\left(-\frac{1}{2}\mathbf{z}'\mathbf{z}\right) = \prod_{i=1}^{p} \frac{1}{(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}z_i^2\right) .$$

En efecto, basta tomar $\mathbf{z} = \mathbf{A}^{-1}(\mathbf{x} - \mu)$, donde $\mathbf{A}$ procede de la factorización $\Sigma = \mathbf{A}\mathbf{A}'$ (factorización de Cholesky), y aplicar el resultado 2.

Volviendo por un momento a la función de verosimilitud considerada anteriormente

$$\mathbb{L}(\mathbf{y}; \beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta)\right)$$

vemos que representa la densidad de un vector normal N-dimensional de media $\mathbf{X}\beta$ y matriz de varianzas-covarianzas $\sigma^2\mathbf{I}_N$. Por lo tanto $\mathbf{y} \rightsquigarrow \mathrm{N}_N\left[\mathbf{X}\beta; \sigma^2\mathbf{I}_N\right]$. Asimismo, dado que $\mathbf{y} = \mathbf{X}\beta + \varepsilon$, entonces es inmediato que $\epsilon \rightsquigarrow \mathrm{N}_N\left[\mathbf{0}; \sigma^2\mathbf{I}_N\right]$.

7.4. Distribución de $\widehat{\beta}$.

Ya conocemos que el estimador máximo verosímil de β viene dado por

$$\widehat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} .$$

Ahora bien, dado que $\mathbf{y} \rightsquigarrow \mathrm{N}_N\left[\mathbf{X}\beta;\sigma^2\mathbf{I}_N\right]$ y aplicando el resultado 2 de las propiedades de la normal multivariante se tendrá que el vector $\widehat{\beta}$ se distribuirá de forma normal con media β y matriz de varianzas-covarianzas

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}_N\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

como ya sabíamos del tema anterior. Así pues

$$\widehat{\beta} \rightsquigarrow \mathrm{N}_{k+1}\left[\beta;\sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right] \ .$$

Por lo tanto cualquier combinación lineal de dicho vector se distribuirá de forma normal. En particular sus componentes serán normales

$$\widehat{\beta}_i \rightsquigarrow \mathrm{N}_1\left[\beta_i;\sigma^2 q_{ii}\right]$$

donde q_{ii} es el i-ésimo elemento de la diagonal de la matriz $(\mathbf{X}'\mathbf{X})^{-1}$.

7.5. Sobre la distribución de formas cuadráticas normales. Distribución de $\widehat{\sigma}^2$. Algunas distribuciones de interés.

En principio podemos basarnos en la distribución de $\widehat{\beta}$ para hacer contrastes de hipótesis sobre los elementos de β. Sin embargo no podremos realizar esos contrastes salvo que conozcamos el valor de σ^2. Para realizar contrastes cuando σ^2 es desconocido y determinar sus propiedades necesitamos conocer la distribución de $\widehat{\sigma}^2$ y, por lo tanto, la de $\mathbf{e}'\mathbf{e}$. Observemos que como $\mathbf{e} = \left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]\mathbf{y}$, entonces $\mathbf{e}'\mathbf{e} = \mathbf{y}'\left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]\mathbf{y}$. Esta última expresión es una forma cuadrática con variables normales, por lo que habrá que introducir algunos conceptos sobre distribución de formas cuadráticas normales.

Resultado 6. *Si $\mathbf{x}$ es un vector aleatorio N-dimensional con distribución normal de media cero y matriz de varianzas-covarianzas $\mathbf{I}_N$, entonces*

$$\mathbf{x}'\mathbf{x} \rightsquigarrow \chi^2_N \ .$$

Resultado 7. *Si $\mathbf{x}$ es un vector aleatorio N-dimensional con distribución normal de media μ y matriz de varianzas-covarianzas Σ, entonces*

$$(\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu) \rightsquigarrow \chi^2_N \ .$$

Resultado 8. *Sea $\mathbf{x}$ un vector N-dimensional normal $N_N[0; \mathbf{I}_N]$. Si $\mathbf{A}$ es una matriz simétrica y constante de rango r, entonces*

$$\mathbf{x}'\mathbf{A}\mathbf{x} \rightsquigarrow \chi^2_r \Leftrightarrow \quad \mathbf{A} \textit{ es idempotente} \ .$$

Resultado 9. *Sea $\mathbf{x}$ un vector N-dimensional normal $N_N[\mu; \mathbf{I}_N]$. Si $\mathbf{A}$ es una matriz simétrica y constante de rango r, entonces*

$$\mathbf{x}'\mathbf{A}\mathbf{x} \rightsquigarrow \chi^2_r(\delta) \Leftrightarrow \quad \mathbf{A} \textit{ es idempotente} \quad \textit{siendo} \quad \delta = \mu' A\mu \ .$$

Resultado 10. *Si* $\mathbf{x}$ *es un vector normal* $N_N[\mu;\Sigma]$ *y si* $\mathbf{A}$ *y* $\mathbf{B}$ *son matrices de orden* N *no aleatorias, entonces*

$$\mathbf{BA} = 0 \Rightarrow \mathbf{Bx} \ \text{ y } \ \mathbf{x}'\mathbf{Ax} \quad \text{son independientes} \ .$$

Resultado 11. *Si* $\mathbf{x}$ *es un vector normal* $N_N[\mu;\mathbf{I}_N]$ *y si* $\mathbf{A}$ *y* $\mathbf{B}$ *son matrices de orden* N *no aleatorias, entonces*

$$\mathbf{AB} = 0 \Leftrightarrow \mathbf{x}'\mathbf{Ax} \ \text{ y } \ \mathbf{x}'\mathbf{Bx} \quad \text{son independientes}$$

Por último vamos a obtener la distribución de ciertas funciones de los estimadores mínimo cuadráticos. Estas distribuciones van a ser muy importantes en el próximo tema a la hora de establecer contrastes sobre los parámetros del modelo. Asimismo probaremos la independencia de los estimadores mínimo cuadráticos $\widehat{\beta}$ y $\widehat{\sigma}^2$.

Proposición 1. *Los estimadores* $\widehat{\beta}$ *y* $\widehat{\sigma}^2$ *son independientes.*

<u>Demostración</u>

Comprobaremos la independencia de $\widehat{\beta}$ y de $\mathbf{e}'\mathbf{e}$. Para ello basta con observar que $\widehat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ y que $\mathbf{e}'\mathbf{e} = \mathbf{y}'\left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]\mathbf{y}$. Puesto que $\mathbf{X}\left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] = \mathbf{0}$, el resultado 10 nos proporciona la independencia.

Proposición 2. *Sea* $\mathbf{A}$ *una matriz no aleatoria de dimensión* $s \times (k+1)$ $(s \leq k+1)$ *y de rango* s. *Son ciertas las siguientes afirmaciones:*

1. *Si β^0 es un vector de constantes, entonces*

$$\Phi_1 = \frac{1}{\sigma^2}(\widehat{\beta} - \beta^0)'\mathbf{A}'\left[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\right]^{-1}\mathbf{A}(\widehat{\beta} - \beta^0) \rightsquigarrow \chi^2_s(\delta)$$

siendo

$$\delta = \frac{(\beta - \beta^0)'\mathbf{A}'\left[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\right]^{-1}\mathbf{A}(\beta - \beta^0)}{\sigma^2} .$$

2. $\Phi_2 = \frac{1}{\sigma^2}\mathbf{e}'\mathbf{e} = \frac{(N-k-1)\widehat{\sigma}^2}{\sigma^2} \rightsquigarrow \chi^2_{N-k-1}$.

<u>Demostración</u>

1. Puesto que $\widehat{\beta} \rightsquigarrow \mathrm{N}_{k+1}\left[\beta; \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right]$, entonces

$$\widehat{\beta} - \beta^0 \rightsquigarrow \mathrm{N}_{k+1}\left[\beta - \beta^0; \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right] .$$

Con ello

$$\frac{\mathbf{A}(\widehat{\beta} - \beta^0)}{\sigma} \rightsquigarrow \mathrm{N}_s\left[\frac{\mathbf{A}(\beta - \beta^0)}{\sigma}; \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\right] .$$

Ahora bien, puesto que $\mathbf{A}$ es de rango s también lo será $\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'$ por lo que es definida positiva. Sea $\mathbf{C}$ la matriz tal que (descomposición de Cholesky) $\mathbf{C}\mathbf{C}' = \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'$. Con ello

$$v = \frac{\mathbf{C}^{-1}\mathbf{A}(\widehat{\beta} - \beta^0)}{\sigma} \rightsquigarrow \mathbf{N}_s \left[\frac{\mathbf{C}^{-1}\mathbf{A}(\beta - \beta^0)}{\sigma}; \mathbf{I}_s \right] .$$

Aplicando el resultado 9 sobre formas cuadráticas se tiene que $\Phi_1 = v'v \rightsquigarrow \chi^2_s(\delta)$, como se quería.

Observemos que si $\beta^0 = \beta$ entonces la distribución es centrada.

2. Como $\varepsilon \rightsquigarrow \mathbf{N}_N \left[0; \sigma^2 \mathbf{I}_N\right]$ entonces $\dfrac{\varepsilon}{\sigma} \rightsquigarrow \mathbf{N}_N \left[0; \mathbf{I}_N\right]$ y puesto que

$$\mathbf{e} = \left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] \mathbf{y} = \left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] \varepsilon$$

se tiene

$$\Phi_2 = \frac{\mathbf{e}'\mathbf{e}}{\sigma^2} = \frac{\varepsilon'}{\sigma} \left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] \frac{\varepsilon}{\sigma} .$$

Ahora bien, la matriz $\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ es idempotente por lo que su rango es igual a su traza. Puesto que $\mathrm{tr}\left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] = N - k - 1$, basta aplicar el resultado 8 sobre formas cuadráticas para concluir. [2]

[2]Notemos que esta distribución es centrada, tal y como se comentó en el apartado de regresión simple. Por lo tanto su comportamiento en contrastes de hipótesis no estará marcado por las hipótesis nulas que se establezcan.

Notemos la importancia que tiene Φ_1 en tanto en cuanto nos valdrá para realizar contrastes sobre subconjuntos de parámetros dentro del vector paramétrico β. Para ello bastará con hacer uso de una matriz $\mathbf{A}$ adecuada al caso que nos interese.

Sin embargo podemos avanzar algo más sobre estas cuestiones. En efecto, sería deseable que ambas distribuciones, la de Φ_1 y la de Φ_2, fueran independientes y así tener un estadístico válido para futuros contrastes. Este hecho es cierto sin más que observar que Φ_1 depende sólo de $\widehat{\beta}$ y que Φ_2 depende sólo de $\widehat{\sigma}^2$ y aplicar el resultado 9. No obstante, para hacer uso de los resultados sobre formas cuadráticas normales volvamos a confirmar este hecho como consecuencia de la siguiente proposición

Proposición 3. *Bajo las condiciones de la proposición 2 se tiene:*

1. Φ_1 *y* Φ_2 *son independientes.*
2. $\dfrac{\Phi_1/s}{\Phi_2/N-k-1} \rightsquigarrow F_{s,N-k-1}(\delta)$ *siendo*

$$\delta = \frac{(\beta-\beta^0)'\mathbf{A}'\left[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\right]^{-1}\mathbf{A}(\beta-\beta^0)}{\sigma^2}.$$

Demostración

1. Perseguimos ponernos en las condiciones del resultado 11 sobre las distribuciones de formas cuadráticas normales, que es el que tenía en su tesis la independencia. Para ello es preciso expresar Φ_1 en función de ε.

Como $\widehat{\beta} = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon$ se tiene, por un lado,

$$\widehat{\beta} - \beta^0 = \beta - \beta^0 + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\left[\varepsilon + \mathbf{X}(\beta - \beta^0)\right]$$

con lo cual

$$\Phi_1 = \frac{1}{\sigma^2}(\widehat{\beta} - \beta^0)'\mathbf{A}'\left[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\right]^{-1}\mathbf{A}(\widehat{\beta} - \beta^0) =$$

$$= \frac{(\varepsilon + \mathbf{X}(\beta - \beta^0))'}{\sigma}\left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\left[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'\right]^{-1}\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]\frac{(\varepsilon + \mathbf{X}(\beta - \beta^0))}{\sigma} =$$

$$= \frac{(\varepsilon + \mathbf{X}(\beta - \beta^0))'}{\sigma}\mathbf{A}_1\frac{(\varepsilon + \mathbf{X}(\beta - \beta^0))}{\sigma}$$

y, por otro lado

$$\mathbf{e} = \left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]\varepsilon = \left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]\left[\varepsilon + \mathbf{X}(\beta - \beta^0)\right]$$

con lo cual

$$\Phi_2 = \frac{(\varepsilon + \mathbf{X}(\beta - \beta^0))'}{\sigma}\left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right]\frac{(\varepsilon + \mathbf{X}(\beta - \beta^0))}{\sigma} =$$

$$= \frac{(\varepsilon + \mathbf{X}(\beta - \beta^0))'}{\sigma} \mathbf{A}_2 \frac{(\varepsilon + \mathbf{X}(\beta - \beta^0))}{\sigma} \; .$$

De esta forma ambas formas cuadráticas están expresadas en función del mismo vector

$$\frac{\varepsilon + \mathbf{X}(\beta - \beta^0)}{\sigma}$$

que se distribuye de forma normal

$$\frac{\varepsilon + \mathbf{X}(\beta - \beta^0)}{\sigma} \rightsquigarrow \mathrm{N}_N \left[\frac{\mathbf{X}(\beta - \beta^0)}{\sigma}; \mathbf{I}_N \right] \; .$$

Resta por comprobar que $\mathbf{A}_1\mathbf{A}_2 = \mathbf{0}$. En efecto:

$$\left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}' \left[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}' \right]^{-1} \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] \left[\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right] =$$

$$= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}' \left[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}' \right]^{-1} \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' -$$

$$- \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}' \left[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}' \right]^{-1} \mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{0}$$

por lo que Φ_1 y Φ_2 son independientes.

2. Resulta inmediato puesto que $\Phi_1 \rightsquigarrow \chi_s^2(\delta)$ y $\Phi_2 \rightsquigarrow \chi_{N-k-1}^2$ y ambas son independientes. Obviamente, si $\beta^0 = \beta$ la distribución es centrada.