Modelos Lineales

Inferencia

Fernando Massa; Bruno Bellagamba

martes 9 de abril 2024

FCEA Facultad de Ciencias Económicas y de Administración
IESTA Instituto de Estadística
Universidad de la República Uruguay

Temario

1. Distribución
2. Estimación
3. Distribución de los estimadores
4. Hipótesis lineales
5. Próxima clase

- Veremos las coincidencias y diferencias entre los estimadores *MCO* y *ML*.
- Plantearemos la distribución de los estimadores.
- Veremos la forma general del estadístico F.
- Cerraremos con un ejemplo de aplicación de este estadístico.

Distribución de ε

Hasta ahora hemos obtenido resultados algebraicos y probabilísticos bajo un enfoque de mínimos cuadrados. No obstante, gran parte del trabajo con modelos lineales consiste en su uso para:

- poner a prueba hipótesis sobre el poder explicativo de las variables contenidas en X;
- realizar predicciones.

Distribución de ε

Hasta ahora hemos obtenido resultados algebraicos y probabilísticos bajo un enfoque de mínimos cuadrados. No obstante, gran parte del trabajo con modelos lineales consiste en su uso para:

- poner a prueba hipótesis sobre el poder explicativo de las variables contenidas en X;
- realizar predicciones.

En ambos casos nos interesará evaluar la incertudimbre de los valores estimados o predichos con intervalos de confianza o realizar pruebas de hipótesis que contesten ciertas preguntas de interés. Es decir, necesitaremos realizar uno u otro proceso **inferencial**.

Distribución de ε

Hasta ahora hemos obtenido resultados algebraicos y probabilísticos bajo un enfoque de mínimos cuadrados. No obstante, gran parte del trabajo con modelos lineales consiste en su uso para:

- poner a prueba hipótesis sobre el poder explicativo de las variables contenidas en X;
- realizar predicciones.

En ambos casos nos interesará evaluar la incertudimbre de los valores estimados o predichos con intervalos de confianza o realizar pruebas de hipótesis que contesten ciertas preguntas de interés. Es decir, necesitaremos realizar uno u otro proceso **inferencial**.

Sin embargo, para esto necesitaremos realizar algún supuesto sobre la distribución de los errores del modelo.

Distribución de ε

En este curso, el enfoque *oficial* es de carácter paramétrico, por lo que supondremos una distribución de probabilidad en concreto para el vector de errores.

Distribución de ε

En este curso, el enfoque *oficial* es de carácter paramétrico, por lo que supondremos una distribución de probabilidad en concreto para el vector de errores.

$$\varepsilon \sim N_n(\mathbf{0}, \sigma^2 I_n)$$

De esta manera, suponemos que la distribucicón de los errores es normal multivariada e incorporamos los supuestos de media nula, varianza constante y covarianzas nulas.

Distribución de ε

En este curso, el enfoque *oficial* es de carácter paramétrico, por lo que supondremos una distribución de probabilidad en concreto para el vector de errores.

$$\varepsilon \sim N_n(\mathbf{0}, \sigma^2 I_n)$$

De esta manera, suponemos que la distribucicón de los errores es normal multivariada e incorporamos los supuestos de media nula, varianza constante y covarianzas nulas.

Si retomamos el modelo lineal en su expresión matricial:

$$Y = X\beta + \varepsilon$$

Podemos obtener la distribución del la variable de respuesta Y:

$$Y \sim N_n(X\beta, \sigma^2 I_n)$$

Debido a que las covarianzas de los elementos de Y son cero, es posible afirmar que las observaciones son independientes. A partir de esta observación existen 2 formas de plantear la función de verosimilitud.

Debido a que las covarianzas de los elementos de Y son cero, es posible afirmar que las observaciones son independientes. A partir de esta observación existen 2 formas de plantear la función de verosimilitud.

1. En forma escalar

$$\mathscr{L}(Y|X,\beta,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{\left(y_i - x_i'\beta\right)^2}{\sigma^2}}$$

2. En forma vectorial

$$\mathscr{L}(Y|X,\beta,\sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta)}$$

Log-verosimilitud

La obtención de las estimaciones surge de la maximización de esta función. En la práctica, se suele trabajar con la *log-verosimilitud* debido a que es más cómodo trabajar con ella que con la verosimilitud y a que se obtienen resultados idénticos.

La obtención de las estimaciones surge de la maximización de esta función. En la práctica, se suele trabajar con la *log-verosimilitud* debido a que es más cómodo trabajar con ella que con la verosimilitud y a que se obtienen resultados idénticos.

De aquí en más continuaremos con la forma vectorial.

$$\ell(Y|X,\beta,\sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(Y - X\beta)^{'}(Y - X\beta)$$

La obtención de las estimaciones surge de la maximización de esta función. En la práctica, se suele trabajar con la *log-verosimilitud* debido a que es más cómodo trabajar con ella que con la verosimilitud y a que se obtienen resultados idénticos.

De aquí en más continuaremos con la forma vectorial.

$$\ell(Y|X,\beta,\sigma^2) = -\frac{n}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left(Y - X\beta\right)'\left(Y - X\beta\right)$$

A continuación, se obtienen los estimadores, derivando e igualando a cero.

$$\begin{aligned} \tfrac{\partial}{\partial\beta}\ell(Y|X,\beta,\sigma^2) &= 0 \\ \tfrac{\partial}{\partial\sigma^2}\ell(Y|X,\beta,\sigma^2) &= 0 \end{aligned}$$

Estimación MV de β

$$\begin{aligned}
\frac{\partial}{\partial\beta}\ell(Y|X,\beta,\sigma^2) &= \frac{\partial}{\partial\beta}\left[-\frac{n}{2}\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta)\right]=0\\
&= -\frac{\partial}{\partial\beta}\frac{1}{2\sigma^2}(Y-X\beta)'(Y-X\beta)=0\\
&= \frac{\partial}{\partial\beta}(Y-X\beta)'(Y-X\beta)=0
\end{aligned}$$

Estimación MV de β

$$\begin{aligned}
\frac{\partial}{\partial\beta}\ell(Y|X,\beta,\sigma^2) &= \frac{\partial}{\partial\beta}\left[-\frac{n}{2}\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}(Y-X\beta)^{'}(Y-X\beta)\right]=0\\
&= -\frac{\partial}{\partial\beta}\frac{1}{2\sigma^2}(Y-X\beta)^{'}(Y-X\beta)=0\\
&= \frac{\partial}{\partial\beta}(Y-X\beta)^{'}(Y-X\beta)=0
\end{aligned}$$

En este punto podemos observar que estamos ante el mismo problema de optimización que teníamos cuando obtuvimos los estimadores MCO, por lo que:

$$\hat{\beta}_{MV}=\left(X^{'}X\right)^{-1}X^{'}Y=\hat{\beta}_{MCO}$$

Y se mantienen todas las propiedades que estudiamos anteriormente.

Estimación *MV* de σ^2

$$
\begin{aligned}
\frac{\partial}{\partial \sigma^2}\ell(Y|X,\beta,\sigma^2) &= \frac{\partial}{\partial \sigma^2}\left[-\frac{n}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\right]=0\\
&= \frac{\partial}{\partial \sigma^2}\left[-\frac{n}{2}\log\left(\sigma^2\right)-\frac{1}{2\sigma^2}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\right]=0\\
&= -\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)=0
\end{aligned}
$$

Estimación MV de σ^2

$$
\begin{aligned}
\frac{\partial}{\partial\sigma^2}\ell(Y|X,\beta,\sigma^2) &= \frac{\partial}{\partial\sigma^2}\left[-\frac{n}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\right]=0\\
&= \frac{\partial}{\partial\sigma^2}\left[-\frac{n}{2}\log\left(\sigma^2\right)-\frac{1}{2\sigma^2}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\right]=0\\
&= -\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)=0
\end{aligned}
$$

Despejando de esta última ecuación surge que:

$$
\begin{aligned}
-\frac{n}{2\sigma^2} &= -\frac{1}{2\sigma^4}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\\
\frac{n}{\sigma^2} &= \frac{1}{\sigma^4}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\\
n &= \frac{1}{\sigma^2}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\\
n\sigma^2 &= \left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\\
\hat{\sigma}^2 &= \frac{\left(Y-X\hat{\beta}\right)^{'}\left(Y-X\hat{\beta}\right)}{n}
\end{aligned}
$$

Estimación MV de σ^2

$$
\begin{aligned}
\frac{\partial}{\partial\sigma^2}\ell(Y|X,\beta,\sigma^2) &= \frac{\partial}{\partial\sigma^2}\left[-\frac{n}{2}\log\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\right]=0\\
&= \frac{\partial}{\partial\sigma^2}\left[-\frac{n}{2}\log\left(\sigma^2\right)-\frac{1}{2\sigma^2}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\right]=0\\
&= -\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)=0
\end{aligned}
$$

Despejando de esta última ecuación surge que:

$$
\begin{aligned}
-\frac{n}{2\sigma^2} &= -\frac{1}{2\sigma^4}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\\
\frac{n}{\sigma^2} &= \frac{1}{\sigma^4}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\\
n &= \frac{1}{\sigma^2}\left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\\
n\sigma^2 &= \left(Y-X\beta\right)^{'}\left(Y-X\beta\right)\\
\hat{\sigma}^2 &= \frac{\left(Y-X\hat{\beta}\right)^{'}\left(Y-X\hat{\beta}\right)}{n}
\end{aligned}
$$

Vemos como (del mismo modo que ocurría en el curso de *Inferencia I*) el estimador *MV* de σ^2 está **sesgado**. Por este motivo, trabajaremos con el estimador insesgado que obtuvimos anteriormente.

Distribución de $\hat{\beta}$

Debido a que el estimador MV coincide con el obtenido mediante MCO se sigue cumpliendo que: $\mathbb{E}(\hat{\beta}_{MV}) = \beta$ y que $\mathbb{V}ar(\hat{\beta}_{MV}) = \sigma^2 \left(X'X\right)^{-1}$

Distribución de $\hat{\beta}$

Debido a que el estimador *MV* coincide con el obtenido mediante *MCO* se sigue cumpliendo que: $\mathbb{E}(\hat{\beta}_{MV}) = \beta$ y que $\mathbb{V}ar(\hat{\beta}_{MV}) = \sigma^2 \left(X'X\right)^{-1}$

No obstante, habiendo agregado el supuesto de la distribución de los errores (y por ende, de Y), es posible obtener la distribución de $\hat{\beta}_{MV}$.

Distribución de $\hat{\beta}$

Debido a que el estimador *MV* coincide con el obtenido mediante *MCO* se sigue cumpliendo que: $\mathbb{E}(\hat{\beta}_{MV}) = \beta$ y que $\mathbb{V}ar(\hat{\beta}_{MV}) = \sigma^2 \left(X'X\right)^{-1}$

No obstante, habiendo agregado el supuesto de la distribución de los errores (y por ende, de Y), es posible obtener la distribución de $\hat{\beta}_{MV}$.

Pero antes repasemos una propiedad de la distribución normal multivariada.

Normal multivariada

Sea Y un vector aleatorio en $\mathbb{R}^p$ con distribución $N_p(\mu, \Sigma)$, y sea Z un nuevo vector aleatorio en $\mathbb{R}^q$ el resultado de la transformación lineal $Z = a + BY$. Siendo $a \in \mathbb{R}^q$ y $B \in \mathbb{R}^{q \times p}$, entonces $Z \sim N_q\left(a + B\mu, B\Sigma B'\right)$

Distribución de $\hat{\beta}$

A paritir de lo anterior, es fácil observar que, como $\hat{\beta}_{MV}$ es una transformación lineal de Y, su distribución también es normal multivariada.

$$\hat{\beta}_{MV} \sim N_{k+1}\left(\beta, \sigma^2 \left(X^{'} X\right)^{-1}\right)$$

Distribución de $\hat{\beta}$

A paritir de lo anterior, es fácil observar que, como $\hat{\beta}_{MV}$ es una transformación lineal de Y, su distribución también es normal multivariada.

$$\hat{\beta}_{MV} \sim N_{k+1}\left(\beta, \sigma^2 \left(X'X\right)^{-1}\right)$$

Este resultado nos servirá para comenzar a plantear intervalos de confianza y pruebas de hipótesis en el contexto del modelo de RLM. No obstante nótese como aún permanece el parámetro (desconocido) σ^2 en la matriz de varianzas y covarianzas.

Distribución de $\hat{\beta}$

A paritir de lo anterior, es fácil observar que, como $\hat{\beta}_{MV}$ es una transformación lineal de Y, su distribución también es normal multivariada.

$$\hat{\beta}_{MV} \sim N_{k+1}\left(\beta, \sigma^2 \left(X^{'}X\right)^{-1}\right)$$

Este resultado nos servirá para comenzar a plantear intervalos de confianza y pruebas de hipótesis en el contexto del modelo de RLM. No obstante nótese como aún permanece el parámetro (desconocido) σ^2 en la matriz de varianzas y covarianzas.

A continuación veremos cual es la distribución de $\hat{\sigma}^2$ y su relación con $\hat{\beta}_{MV}$

Distribución de $\hat{\sigma}^2$

Para obtener la distribución del estimador (insesgado) de σ^2 nos valdremos del siguiente resultado.

Para obtener la distribución del estimador (insesgado) de σ^2 nos valdremos del siguiente resultado.

Distribución de formas cuadráticas

Sea Y un vector aleatorio en $\mathbb{R}^p$ con distribución $N_p(\mu, \Sigma)$, y sea A una matriz simétrica de $p \times p$ y de rango r.

$$Y^{'}AY \sim \chi^2_r(\lambda) \Leftrightarrow \text{A}\Sigma \text{ es idempotente.}$$

Nota: el parámetro $\lambda = \frac{1}{2}\mu^{'}A\mu$ es llamado parámetro de no centralidad.

DISTRIBUCIÓN DE $\hat{\sigma}^2$

Para obtener la distribución del estimador (insesgado) de σ^2 nos valdremos del siguiente resultado.

DISTRIBUCIÓN DE FORMAS CUADRÁTICAS

Sea Y un vector aleatorio en $\mathbb{R}^p$ con distribución $N_p(\mu, \Sigma)$, y sea A una matriz simétrica de $p \times p$ y de rango r.

$$Y^{'}AY \sim \chi^2_r(\lambda) \Leftrightarrow \text{A}\Sigma \text{ es idempotente.}$$

Nota: el parámetro $\lambda = \frac{1}{2}\mu^{'}A\mu$ es llamado parámetro de no centralidad.

Recordemos el estimador de σ^2 y su expresión como forma cuadrática de Y.

$$\begin{aligned} \hat{\sigma}^2 &= \frac{\left(Y - X\hat{\beta}\right)^{'}\left(Y - X\hat{\beta}\right)}{n-k-1} \\ &= \frac{Y^{'}(I_n - H_X)Y}{n-k-1} \end{aligned}$$

Distribución de $\hat{\sigma}^2$

El estadístico cuya distribución queremos conocer es:

$$\frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} = \frac{Y'\left(I_n - H_X\right)Y}{\sigma^2}$$

DISTRIBUCIÓN DE $\hat{\sigma}^2$

El estadístico cuya distribución queremos conocer es:

$$\frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} = \frac{Y^{'}(I_n - H_X)Y}{\sigma^2}$$

Para aplicar el teorema de la diapositiva anterior recordaremos que $Y \sim N_n(X\beta, \sigma^2 I_n)$ y el papel de la matriz A lo ocupa la matriz de proyección sobre el complemento ortogonal de $\mathbb{C}(X)$ dividio σ^2.

Distribución de $\hat{\sigma}^2$

El estadístico cuya distribución queremos conocer es:

$$\frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} = \frac{Y^{'}\left(I_n - H_X\right)Y}{\sigma^2}$$

Para aplicar el teorema de la diapositiva anterior recordaremos que $Y \sim N_n(X\beta, \sigma^2 I_n)$ y el papel de la matriz A lo ocupa la matriz de proyección sobre el complemento ortogonal de $\mathbb{C}(X)$ dividio σ^2.

Para corroborar la dondición que exige el teorema, planteamos el producto $A\Sigma$.

$$A\Sigma = A\mathbb{V}ar(Y) = \frac{I_n - H_X}{\sigma^2} \times \sigma^2 I_n = I_n - H_X$$

Distribución de $\hat{\sigma}^2$

El estadístico cuya distribución queremos conocer es:

$$\frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} = \frac{Y^{'}\,(I_n - H_X)\,Y}{\sigma^2}$$

Para aplicar el teorema de la diapositiva anterior recordaremos que $Y \sim N_n(X\beta, \sigma^2 I_n)$ y el papel de la matriz A lo ocupa la matriz de proyección sobre el complemento ortogonal de $\mathbb{C}(X)$ dividio σ^2.

Para corroborar la dondición que exige el teorema, planteamos el producto $A\Sigma$.

$$A\Sigma = A\mathbb{V}ar(Y) = \frac{I_n - H_X}{\sigma^2} \times \sigma^2 I_n = I_n - H_X$$

La cual es una matriz idempotente, por lo que el estadístico tiene distribución χ^2 con $n-k-1$ grados de libertad.

Finalmente, resta por determinar el valor del parámetro de no centralidad.

$$\frac{1}{2}\mu^{'}A\mu = \frac{1}{2}\mathbb{E}(Y)^{'}A\mathbb{E}(Y) = \frac{1}{2}\beta^{'}X^{'}(I_n - H_X)X\beta = 0$$

Donde se empleó la condición de que $(I_n - H_X)X = 0$.

Distribución de $\hat{\sigma}^2$

Finalmente, resta por determinar el valor del parámetro de no centralidad.

$$\frac{1}{2}\mu^{'}A\mu = \frac{1}{2}\mathbb{E}(Y)^{'}A\mathbb{E}(Y) = \frac{1}{2}\beta^{'}X^{'}(I_n - H_X)X\beta = 0$$

Donde se empleó la condición de que $(I_n - H_X)X = 0$.

Por lo que:

$$\frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-k-1}$$

Antes de construir los estadísticos de prueba, es necesario obtener un último resultado que relaciona $\hat{\beta}_{MV}$ con $\hat{\sigma}^2$.

INDEPENDENCIA ENTRE UNA FORMA LINEAL Y UNA FORMA CUADRÁTICA

Sea Y un vector aleatorio en $\mathbb{R}^p$ con distribución $N_p(\mu, \Sigma)$, sea A una matriz simétrica de $p \times p$ y B una matriz de $m \times p$.

Las transformaciones BY y $Y^{'}AY$ son independientes $\Leftrightarrow \quad B\Sigma A = 0$

Antes de construir los estadísticos de prueba, es necesario obtener un último resultado que relaciona $\hat{\beta}_{MV}$ con $\hat{\sigma}^2$.

Independencia entre una forma lineal y una forma cuadrática

Sea Y un vector aleatorio en $\mathbb{R}^p$ con distribución $N_p(\mu, \Sigma)$, sea A una matriz simétrica de $p \times p$ y B una matriz de $m \times p$.

Las transformaciones BY y $Y^{'}AY$ son independientes $\Leftrightarrow \quad B\Sigma A = 0$

Independencia entre dos formas cuadráticas

Sea Y un vector aleatorio en $\mathbb{R}^p$ con distribución $N_p(\mu, \Sigma)$, sean A y B matrices simétricas de $p \times p$.

Las transformaciones $Y^{'}BY$ y $Y^{'}AY$ son independientes $\Leftrightarrow \quad B\Sigma A = 0$

Independencia de los estimadores

El motivo por el que se mencionan los teoremas anteriores es basicamente para determinar la independencia entre $\hat{\beta}_{MV}$ y $\hat{\sigma}^2$. Para ello recordemos la forma de cada uno.

El motivo por el que se mencionan los teoremas anteriores es basicamente para determinar la independencia entre $\hat{\beta}_{MV}$ y $\hat{\sigma}^2$. Para ello recordemos la forma de cada uno.

$$\begin{aligned} \hat{\beta}_{MV} &= \left(X^{'}X\right)^{-1} X^{'}Y = BY \\ \frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} &= \frac{Y^{'}(I_n - H_X)Y}{\sigma^2} = Y^{'}AY \end{aligned}$$

El motivo por el que se mencionan los teoremas anteriores es basicamente para determinar la independencia entre $\hat{\beta}_{MV}$ y $\hat{\sigma}^2$. Para ello recordemos la forma de cada uno.

$$\begin{aligned} \hat{\beta}_{MV} &= \left(X^{'}X\right)^{-1}X^{'}Y = BY \\ \frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} &= \frac{Y^{'}(I_n - H_X)Y}{\sigma^2} = Y^{'}AY \end{aligned}$$

Por lo visto anteriormente, para determinar la independencia entre ellos, es necesario plantear el producto $B\Sigma A$. Si tenemos en cuenta que en este caso $B = \left(X^{'}X\right)^{-1}X^{'}$, $A = \frac{1}{\sigma^2}(I_n - H_X)$ y que $\Sigma = \sigma^2 I_n$:

$$\begin{aligned} B\Sigma A &= \left(X^{'}X\right)^{-1}X^{'} \times \sigma^2 I_n \times \frac{1}{\sigma^2}(I_n - H_X) \\ &= \left(X^{'}X\right)^{-1}X^{'} \times I_n \times (I_n - H_X) \\ &= \left(X^{'}X\right)^{-1}X^{'} \times (I_n - H_X) \\ &= 0 \end{aligned}$$

Quedando demostrada la independencia entre ambos estadísticos.

A partir de ahora contamos con todas las herramientas para realizar diversos procesos inferenciales:

- Pruebas de hipótesis.
- Intervalos de confianza

A partir de ahora contamos con todas las herramientas para realizar diversos procesos inferenciales:

- Pruebas de hipótesis.
- Intervalos de confianza

Comenzaremos realizando pruebas de hipótesis dentro de un marco de trabajo general a partir del cual se podrán obtener varios casos especiales mediante distintas simplificaciones. Este marco es el de las **Pruebas de hipótesis lineales** que comprende:

A partir de ahora contamos con todas las herramientas para realizar diversos procesos inferenciales:

- Pruebas de hipótesis.
- Intervalos de confianza

Comenzaremos realizando pruebas de hipótesis dentro de un marco de trabajo general a partir del cual se podrán obtener varios casos especiales mediante distintas simplificaciones. Este marco es el de las **Pruebas de hipótesis lineales** que comprende:

- Significación global del modelo.

Hipótesis lineales

A partir de ahora contamos con todas las herramientas para realizar diversos procesos inferenciales:

- Pruebas de hipótesis.
- Intervalos de confianza

Comenzaremos realizando pruebas de hipótesis dentro de un marco de trabajo general a partir del cual se podrán obtener varios casos especiales mediante distintas simplificaciones. Este marco es el de las **Pruebas de hipótesis lineales** que comprende:

- Significación global del modelo.
- Significación de una variable.

Hipótesis lineales

A partir de ahora contamos con todas las herramientas para realizar diversos procesos inferenciales:

- Pruebas de hipótesis.
- Intervalos de confianza

Comenzaremos realizando pruebas de hipótesis dentro de un marco de trabajo general a partir del cual se podrán obtener varios casos especiales mediante distintas simplificaciones. Este marco es el de las **Pruebas de hipótesis lineales** que comprende:

- Significación global del modelo.
- Significación de una variable.
- Significación de un subconjunto de variables.

Hipótesis lineales

A partir de ahora contamos con todas las herramientas para realizar diversos procesos inferenciales:

- Pruebas de hipótesis.
- Intervalos de confianza

Comenzaremos realizando pruebas de hipótesis dentro de un marco de trabajo general a partir del cual se podrán obtener varios casos especiales mediante distintas simplificaciones. Este marco es el de las **Pruebas de hipótesis lineales** que comprende:

- Significación global del modelo.
- Significación de una variable.
- Significación de un subconjunto de variables.
- Significación de una combinación lineal.

Hipótesis lineales

A partir de ahora contamos con todas las herramientas para realizar diversos procesos inferenciales:

- Pruebas de hipótesis.
- Intervalos de confianza

Comenzaremos realizando pruebas de hipótesis dentro de un marco de trabajo general a partir del cual se podrán obtener varios casos especiales mediante distintas simplificaciones. Este marco es el de las **Pruebas de hipótesis lineales** que comprende:

- Significación global del modelo.
- Significación de una variable.
- Significación de un subconjunto de variables.
- Significación de una combinación lineal.
- Significación de varias combinaciones lineales.

Estadístico F

Anteriormente dijimos que:

$$
\begin{aligned}
\hat{\beta}_{MV} &\sim N_{k+1}\left(\beta, \sigma^2\left(X'X\right)^{-1}\right) \\
\frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} &\sim \chi^2_{n-k-1}
\end{aligned}
$$

Estadístico F

Anteriormente dijimos que:

$$\begin{aligned}\hat{\beta}_{MV} &\sim N_{k+1}\left(\beta, \sigma^2\left(X'X\right)^{-1}\right)\\ \frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} &\sim \chi^2_{n-k-1}\end{aligned}$$

Y también vimos que ambos estadísticos son independientes.

Anteriormente dijimos que:

$$\begin{aligned} \hat{\beta}_{MV} &\sim N_{k+1}\left(\beta, \sigma^2 \left(X'X\right)^{-1}\right) \\ \frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} &\sim \chi^2_{n-k-1} \end{aligned}$$

Y también vimos que ambos estadísticos son independientes.

La prueba de hipótesis que nos interesará será la siguiente:

$$\begin{aligned} H_0) &\quad C\beta = b \\ H_1) &\quad C\beta \neq b \end{aligned}$$

Siendo C una matriz de q filas y $k+1$ columnas y b un vector de dimensión q.

La construcción del estadístico de prueba gira sobre la distribución de cierta *norma* del vector $C\hat{\beta}_{MV} - b$.

ESTADÍSTICO F

En base a la normalidad de $\hat{\beta}_{MV}$, se puede afirmar que:

$$C\hat{\beta}_{MV} - b \sim N_q\left(C\beta - b, \sigma^2 C\left(X^{'}X\right)^{-1} C^{'}\right)$$

En base a la normalidad de $\hat{\beta}_{MV}$, se puede afirmar que:

$$C\hat{\beta}_{MV} - b \sim N_q\left(C\beta - b, \sigma^2 C\left(X^{'}X\right)^{-1}C^{'}\right)$$

Bajo el cumplimiento de la hipótesis nula ($C\beta = b$):

$$C\hat{\beta}_{MV} - b \sim N_q\left(0, \sigma^2 C\left(X^{'}X\right)^{-1}C^{'}\right)$$

ESTADÍSTICO F

En base a la normalidad de $\hat{\beta}_{MV}$, se puede afirmar que:

$$C\hat{\beta}_{MV} - b \sim N_q\left(C\beta - b, \sigma^2 C\left(X^{'}X\right)^{-1}C^{'}\right)$$

Bajo el cumplimiento de la hipótesis nula ($C\beta = b$):

$$C\hat{\beta}_{MV} - b \sim N_q\left(0, \sigma^2 C\left(X^{'}X\right)^{-1}C^{'}\right)$$

Luego, si se premultiplica por la *"raíz cuadrada"* de la inversa de la matriz de covarianzas:

$$\left[\sigma^2 C\left(X^{'}X\right)^{-1}C^{'}\right]^{-\frac{1}{2}}\left[C\hat{\beta}_{MV} - b\right] \sim N_q\left(0, I_q\right)$$

Estadístico F

Finalmente, se aplica el siguiente resultado:

Distribución χ^2

Sean $z_1, z_2, \ldots, z_r$ variables aleatorias *i.i.d.* con distribución $N(0,1)$, entonces $\sum_i z_i^2 \sim \chi_r^2$

Lo cual equivale a decir que $Z^{'}Z \sim \chi_r^2$ siendo r la dimensión de $Z \sim N_r(0, I_r)$.

Estadístico F

Finalmente, se aplica el siguiente resultado:

Distribución χ^2

Sean $z_1, z_2, \ldots, z_r$ variables aleatorias *i.i.d.* con distribución $N(0,1)$, entonces $\sum_i z_i^2 \sim \chi_r^2$

Lo cual equivale a decir que $Z^{'}Z \sim \chi_r^2$ siendo r la dimensión de $Z \sim N_r(0, I_r)$.

Retomando la diapositiva anterior:

$$\begin{aligned} \left(C\hat{\beta}_{MV} - b\right)^{'} \left[\sigma^2 C \left(X^{'}X\right)^{-1} C^{'}\right]^{-\frac{1}{2}} \left[\sigma^2 C \left(X^{'}X\right)^{-1} C^{'}\right]^{-\frac{1}{2}} \left(C\hat{\beta}_{MV} - b\right) &\sim \chi_q^2 \\ \left(C\hat{\beta}_{MV} - b\right)^{'} \left[\sigma^2 C \left(X^{'}X\right)^{-1} C^{'}\right]^{-1} \left(C\hat{\beta}_{MV} - b\right) &\sim \chi_q^2 \end{aligned}$$

Estadístico F

Finalmente, se aplica el siguiente resultado:

Distribución χ^2

Sean $z_1, z_2, \ldots, z_r$ variables aleatorias $i.i.d.$ con distribución $N(0,1)$, entonces $\sum_i z_i^2 \sim \chi_r^2$

Lo cual equivale a decir que $Z^{'}Z \sim \chi_r^2$ siendo r la dimensión de $Z \sim N_r(0, I_r)$.

Retomando la diapositiva anterior:

$$\begin{aligned}
\left(C\hat{\beta}_{MV} - b\right)^{'} \left[\sigma^2 C \left(X^{'}X\right)^{-1} C^{'}\right]^{-\frac{1}{2}} \left[\sigma^2 C \left(X^{'}X\right)^{-1} C^{'}\right]^{-\frac{1}{2}} \left(C\hat{\beta}_{MV} - b\right) &\sim \chi_q^2 \\
\left(C\hat{\beta}_{MV} - b\right)^{'} \left[\sigma^2 C \left(X^{'}X\right)^{-1} C^{'}\right]^{-1} \left(C\hat{\beta}_{MV} - b\right) &\sim \chi_q^2
\end{aligned}$$

El problema es que este estadístico incluye al parámetro σ^2...

ESTADÍSTICO F

La solución natural a este problema sería sustituir σ^2 por su estimador $\hat{\sigma}^2$, de forma que el estadístico sería:

$$\frac{\left(C\hat{\beta}_{MV}-b\right)^{'}\left[C\left(X^{'}X\right)^{-1}C^{'}\right]^{-1}\left(C\hat{\beta}_{MV}-b\right)}{\hat{\sigma}^2}$$

Estadístico F

La solución natural a este problema sería sustituir σ^2 por su estimador $\hat{\sigma}^2$, de forma que el estadístico sería:

$$\frac{\left(C\hat{\beta}_{MV} - b\right)^{'} \left[C\left(X^{'}X\right)^{-1}C^{'}\right]^{-1}\left(C\hat{\beta}_{MV} - b\right)}{\hat{\sigma}^2}$$

No obstante, planteado de esa manera, si bien podemos afirmar que las cantidades en numerador y denominador son independientes, no conocemos sus distribuciones.

Podemos solucionar esto, dividiendo en ambas partes entre σ^2.

$$\frac{\left(C\hat{\beta}_{MV} - b\right)^{'} \left[C\left(X^{'}X\right)^{-1}C^{'}\right]^{-1}\left(C\hat{\beta}_{MV} - b\right)/\sigma^2}{\hat{\sigma}^2/\sigma^2}$$

La cantidad en el denominador se parece, a la cantidad pivotal sobre la que trabajamos anteriormente, para completarla la multiplicamos por $(n - k - 1)$.

Estadístico F

La expresión final de la diapositiva anterior era:

$$\frac{\left(C\hat{\beta}_{MV}-b\right)^{'}\left[C\left(X^{'}X\right)^{-1}C^{'}\right]^{-1}\left(C\hat{\beta}_{MV}-b\right)/\sigma^2}{\frac{(n-k-1)SCRes}{\sigma^2}}$$

Donde se cumple que:

- Numerador y denominador son independientes.
- La cantidad del denominador corresponde a una χ^2 dividida sus grados de libertad.
- La cantidad del numerador corresponde a otra distribución χ^2

Esto se parece mucho a una distribución del final del curso de *Probabilidad I*.

Estadístico F

Antes de culminar, debemos recordar el siguiente resultado:

Distribución F

Sean W y Z dos variables aleatorias independientes con distribución χ^2 con ν_W y ν_Z grados de libertad respectivamente, entonces $\frac{W/\nu_W}{Z/\nu_Z} \sim F_{\nu_W,\nu_Z}$

Estadístico F

Antes de culminar, debemos recordar el siguiente resultado:

Distribución F

Sean W y Z dos variables aleatorias independientes con distribución χ^2 con ν_W y ν_Z grados de libertad respectivamente, entonces $\frac{W/\nu_W}{Z/\nu_Z} \sim F_{\nu_W,\nu_Z}$

Todo lo que falta hacer es dividir el numerador entre q y el resultado es el estadístico F.

$$F = \frac{\left(C\hat{\beta}_{MV} - b\right)^{'} \left[C\left(X^{'}X\right)^{-1}C^{'}\right]^{-1}\left(C\hat{\beta}_{MV} - b\right)/q}{Y^{'}(I_n - H_X)Y/(n-k-1)}$$

Estadístico F

Antes de culminar, debemos recordar el siguiente resultado:

Distribución F

Sean W y Z dos variables aleatorias independientes con distribución χ^2 con ν_W y ν_Z grados de libertad respectivamente, entonces $\frac{W/\nu_W}{Z/\nu_Z} \sim F_{\nu_W,\nu_Z}$

Todo lo que falta hacer es dividir el numerador entre q y el resultado es el estadístico F.

$$F = \frac{\left(C\hat{\beta}_{MV} - b\right)' \left[C\left(X'X\right)^{-1} C'\right]^{-1} \left(C\hat{\beta}_{MV} - b\right)/q}{Y'(I_n - H_X)Y/(n-k-1)}$$

En el caso de que la hipótesis nula sea cierta, este estadístico se distribuye $F_{q,n-k-1}$ y gracias a esta distribución podemos obtener valores críticos o p-valores que nos permitan tomar decisiones respecto de H_0.

Significación global del modelo

El caso particular más simple corresponde a la llamada *significación global* del modelo. La idea es que una vez que estimamos los parámetros del modelo lineal que nos interesa, la primer pregunta a realizar es:

El caso particular más simple corresponde a la llamada *significación global* del modelo. La idea es que una vez que estimamos los parámetros del modelo lineal que nos interesa, la primer pregunta a realizar es:

¿Alguna de las variables explicativas incluidas en el modelo contribuye a explicar a Y?

El caso particular más simple corresponde a la llamada *significación global* del modelo. La idea es que una vez que estimamos los parámetros del modelo lineal que nos interesa, la primer pregunta a realizar es:

¿Alguna de las variables explicativas incluidas en el modelo contribuye a explicar a Y?

Planteamos las siguientes hipótesis.

$$H_0)\beta_1 = \beta_2 = \ldots = \beta_k = 0$$
$$H_1)\text{Algún } \beta_j \neq 0$$

El siguiente paso es plantear estas hipótesis en el contexto $C\beta = b$

La manera de plantear esta hipótesis como una hipótesis lineal es con las siguientes elecciones de C y b.

$$C = (\mathbf{0} \quad I_k) \qquad\qquad b^{'} = (0, 0, \ldots, 0)$$

En forma matricial, las hipótesis que se están testeando **simultáneamente** son:

$$\begin{aligned} \beta_1 &= 0 \\ \beta_2 &= 0 \\ &\cdots \\ \beta_k &= 0 \end{aligned}$$

Observe como este enfoque tiene la ventaja de que el error de tipo 1 se mantiene en su valor nominal de 5 % pese a estar realizando k pruebas de hipótesis.

Hace un par de clases trabajamos sobre el ejemplo de las ventas de una empresa en función de su presupuesto de publicidad en distintos medios.

Ejemplo

Hace un par de clases trabajamos sobre el ejemplo de las ventas de una empresa en función de su presupuesto de publicidad en distintos medios.

En ese momento hicimos hincapié en la estimación de los coeficientes y su interpretación.

Ejemplo

Hace un par de clases trabajamos sobre el ejemplo de las ventas de una empresa en función de su presupuesto de publicidad en distintos medios.

En ese momento hicimos hincapié en la estimación de los coeficientes y su interpretación.

Hoy daremos un paso más y estudiaremos la significación global del modelo.

Ejemplo

Hace un par de clases trabajamos sobre el ejemplo de las ventas de una empresa en función de su presupuesto de publicidad en distintos medios.

En ese momento hicimos hincapié en la estimación de los coeficientes y su interpretación.

Hoy daremos un paso más y estudiaremos la significación global del modelo.

Vayamos al R.

La próxima hablaremos de:

- Otros casos especiales del estadístico F.
- Significación de cada variable.
- Restricciones lineales.
- Intervalos de confianza.

- Carmona, Francesc (2003). *Modelos Lineales (notas de curso).* Departament d'Estadística.
- Farraway, Julian (2014). *Linear Models with R, second edition.* Chapman Hall/CRC.
- Rencher, Alvin y Bruce Schaalje (2008). *Linear Models in Statistics, second edition.* John Wiley Sons, Inc.

¿Preguntas?

Muchas Gracias