

Control 2
Aplicaciones de Estadística para Economía y Gestión
IN4402

Segundo Semestre 2017
Profesor: Benjamín Villena R.
Auxiliares: Ronald Leblebici

Fecha: 20 de noviembre de 2017

Reglas generales:

- Cualquier falta a la honestidad será sancionada con Nota 1.0, de acuerdo a los reglamentos y acuerdos éticos vigentes.
- Está prohibido el uso de celulares, computadores, tablets, u otros artefactos electrónicos.
- Si llega a un resultado que requiriera calcular una magnitud numérica, déjelo expresado, indicando con precisión los elementos necesarios.
- Todas las preguntas se ponderan de igual forma y se califican en escala de 0 % a 100 %.
- Una respuesta en blanco recibirá una calificación de 5 %, pero una respuesta totalmente incorrecta tendrá 0 %. Una respuesta correcta, pero que NO tenga relación alguna con la pregunta tendrá también recibirá 0 %. Si usted no sabe algo, tiene algún valor reconocer la falta de conocimiento; eso es preferible a jugar al azar o tener una concepción totalmente errónea.

1. ¿Para qué sirve LASSO? ¿Es útil LASSO para seleccionar una forma funcional de la expectativa condicional?

R: El método LASSO sirve para seleccionar variables o forma funcional correcta dentro de un modelo, incluso si hay un número mayor de variables que observaciones.

Específicamente, sirve para seleccionar un número pequeño de variables dentro de un grupo de K regresores donde K puede ser mayor al tamaño muestral N .

En clases vimos que cualquier forma funcional puede ser aproximada por polinomios y productos de polinomios de las variables del modelo, por lo cual es posible aproximar la forma funcional también por LASSO, al incluir polinomios y producto de polinomios en la lista de K regresores entre los cuales el método escogería aquellos que minimicen su función objetivo.

2. Como gerente de recursos humanos de una empresa, usted desea experimentar si el pago de una comisión a sus vendedores tendrá un impacto positivo sobre las ventas de cada vendedor. Para esto, debe tomar las siguientes decisiones de cómo implementar este estudio con la finalidad de obtener una estimación causal del efecto del tratamiento en los tratados.

Opciones de asignación: (a) Por orden de antigüedad en la empresa ; (b) Por participación voluntaria ; o (c) Por dígito verificador del RUT (por ejemplo 0-5 sin tratamiento y 6-K con tratamiento)

Opciones de publicidad: (a) Hacer estudio públicamente conocido en la empresa; (b) Hacer estudio conocido sólo para los tratados.

Opciones de extensión: (a) Hacer estudio para un grupo menor al 10 % de los vendedores de la empresa; (b) Hacer estudio para todos los vendedores de la empresa.

Explique qué opciones tomaría, sopesando costos y beneficios de cada opción tomada.

R: En un diseño experimental se asigna aleatoriamente la condición de grupo tratado o no tratado (también llamado grupo de control). De las opciones de asignación, el dígito verificador del RUT es la única que no nos da motivos para creer que exista diferencias sistemáticas entre ambos grupos en términos de ventas (no habría sesgo de selección), por lo que sería la opción más recomendable.

Suponiendo que hacer el estudio públicamente conocido en la empresa no implica que los tratados y no-tratados sepan en qué grupo están, se prefiere esta opción para prevenir el efecto Hawthorne (el cual se daría si el estudio fuese conocido solo para los tratados).

Sería costoso aplicar el experimento en toda la empresa (pues tiene que ver con el pago de una comisión). Por esto mismo, se preferiría aplicarla en el 10 % de los trabajadores (siempre que este número no sea muy pequeño y que la muestra sea representativa de la población).

3. Continuando con la pregunta anterior, un colega A suyo sugiere que se debe controlar las ventas observadas por trabajador en función de variables X tales como género, edad, nivel educacional, y antigüedad en la empresa para estimar el efecto de las comisiones en las ventas individuales. El colega B le dice que no es buena idea controlar por X porque el tratamiento ya ha sido asignado aleatoriamente, sin que imputar las características individuales. ¿Qué consejo seguiría y por qué?

R: En clases vimos que incluir observables relevantes reduce la varianza que debe ser explicada por el tratamiento, lo cual es bueno, pues se podría estimar este efecto con mayor certeza. Sin embargo, dado que la asignación del tratamiento se hizo de manera aleatoria, el efecto del tratamiento τ_{ATET} se podría estimar correctamente sin necesidad de incluir X . Por lo tanto, hace sentido seguir el consejo del colega B.

4. Continuando con lo anterior, otro colega C plantea que sería importante controlar por algunos factores psicológicos S tales como motivación, autoestima, empatía, etc. Si las variables S no son observables, ¿qué método podría recomendar para estimar el efecto del tratamiento en los tratados? ¿qué supuestos necesitaría para que este método tenga sentido?

Al estimar el efecto del tratamiento, no habría problema al separar la contribución de este y de S si ambas fueran observables, sin embargo S no lo es, por lo que inevitablemente tendríamos una variable omitida.

En clases hemos visto que la omisión de variables relevantes es una fuente de endogeneidad, problema que se puede solucionar a través del método de variables instrumentales. Para implementar dicho método, necesitamos encontrar un instrumento Z que cumpla los supuestos de relevancia ($Cov(Z, D) \neq 0$) y exclusión ($Cov(Z, U) = 0$), donde D sería la variable de tratamiento y U el error del modelo.

5. Respecto al estudio de Angrist y Krueger (1991) sobre el retorno marginal de la educación en los salarios:

- a) Explique por qué Angrist y Krueger (1991) consideran que el trimestre de nacimiento de un niño es un instrumento válido para la educación adquirida por el alumno.
- b) Explique la crítica de Bound et al (1995) a la condición de debilidad del instrumento, la validez del instrumento, y cómo ambos aspectos interactúan entre ellos.

Según el estudio de Angrist y Krueger (1991), el trimestre de nacimiento está correlacionado con el nivel de escolaridad debido a la legislación que en EEUU requiere que los alumnos estudien hasta una cierta edad (15, 16 o 17) como mínimo (condición de relevancia o no-debilidad). Por otro lado, el trimestre de nacimiento no estaría correlacionado con otras variables no observables como inteligencia o habilidad (condición de exclusión o validez).

Sin embargo Bound et al (1995) señala que la correlación entre educación alcanzada y trimestre de nacimiento es débil. Es así como a pesar de que se cumpla que $Cov(U, Z)$ sea pequeña, si $Cov(X, Z) \approx 0$ se puede dar que el sesgo del estimador de variables instrumentales sea muy alto ($\beta_{VI} = \beta + \frac{Cov(U, Z)}{Cov(X, Z)}$).

6. Respecto al estudio de regresión discontinua de Lee (2008), responda

- a)* ¿Cuál es el supuesto de identificación del efecto de reelección de incumbente respecto al margen de victoria en la elección anterior?
- b)* ¿Cuál es la lógica de mostrar que el número de victorias previas a la elección actual no tiene un “salto” en el margen de victoria de 0 % en la elección anterior?

El supuesto consiste en que la regla de asignación (elección del candidato en la primera instancia) es arbitraria, es decir que los políticos que fueron y no fueron elector por márgenes bajos son comparables. Si los políticos que se encuentran a ambos lados de la discontinuidad se encuentran ahí solo por “suerte” (razones estocásticas no correlacionadas sistemáticamente con la variable X), es posible identificar el efecto del tratamiento τ_{LATE} .

7. Como analista en el mercado nacional de azúcar, y basándose en sus conocimientos de microeconomía, usted considera que un modelo razonable para este mercado está dado por las siguientes ecuaciones de oferta y demanda, respectivamente

$$\log Q_t^o = \alpha_0 + \alpha_1 \log P_t + \alpha_2 \log L_t + \sum_{j=1}^{11} \gamma_j M_{j,t} + U_t$$

$$\log Q_t^d = \beta_0 + \beta_1 \log P_t + \beta_2 \log I_t + \beta_3 \log S_t + \sum_{j=1}^{11} \delta_j M_{j,t} + V_t$$

donde

Q_t es la cantidad ofrecida(o) o demandada/consumida(d) de azúcar en el mes t ;

P_t es el precio de mercado de la azúcar en el mes t ;

L_t es la cantidad de lluvia caída en la zona de producción de remolacha azucarera en el mes t ;

$M_{j,t}$ son variables dummy estacionales que toman valor 1 si el mes es j ; y 0 en otro caso;

I_t es el ingreso nominal en el mes t ;

S_t es el precio de sustitutos del azúcar como fructosa, aspartame, sucralosa, etc.

Utilizando la condición de orden, explique si las ecuaciones de este modelo está subidentificadas, exactamente identificadas, o sobreidentificadas.

R: Igualando oferta y demanda se puede obtener lo siguiente:

$$\log P_t = \frac{\beta_0 - \alpha_0 + \beta_2 \log I_t + \beta_3 \log S_t - \alpha_2 \log L_t + \sum_{j=1}^{11} (\delta_j - \gamma_j) M_{j,t} + V_t - U_t}{\alpha_1 - \beta_1}$$

Luego, se pueden obtener las siguientes covarianzas...

$$\begin{aligned} Cov(\log P_t, U_t) &= \frac{-\sigma_U^2}{\alpha_1 - \beta_1} \neq 0 \\ Cov(\log P_t, V_t) &= \frac{\sigma_V^2}{\alpha_1 - \beta_1} \neq 0 \\ Cov(\log Q_t^o, U_t) &= \frac{-\alpha_1 \sigma_U^2}{\alpha_1 - \beta_1} + \sigma_U^2 \neq 0 \\ Cov(\log Q_t^d, V_t) &= \frac{\beta_1 \sigma_V^2}{\alpha_1 - \beta_1} + \sigma_V^2 \neq 0 \end{aligned}$$

Esto permite concluir que $\log P_t$, $\log Q_t^o$ y $\log Q_t^d$ son endógenas ($g = 3$).

$$k_{\text{oferta}} = g_{\text{-oferta}} + e_{\text{-oferta}} = 1 + 2 = 3 > g - 1 = 2$$

$$k_{\text{demanda}} = g_{\text{-demanda}} + e_{\text{-demanda}} = 1 + 1 = 2 = g - 1$$

Se concluye que la ecuación de oferta está sobre-identificada, mientras que la de demanda está exactamente identificada, por lo que ambas se pueden estimar.

8. Respecto al punto anterior, explique cómo estimaría los parámetros estructurales de la ecuación de demanda. Sea explícito respecto a los supuestos requeridos.

Se puede aplicar el Método de Mínimos Cuadrados en 2 Etapas de la siguiente manera:

- **Etapas 1:** regresionar P_t en función de las variables exógenas del modelo: L_t , I_t , S_t y $\{M_{jt}\}_{j=1}^{11}$ y guardar $\hat{P}_t$.
- **Etapas 2:** regresionar $\log Q_t^d = \beta'_0 + \beta'_1 \log \hat{P}_t + \beta'_2 \log I_t + \beta'_3 \log S_t + \sum_{j=1}^{11} \delta'_j M_{j,t} + V'_t$.

Los supuestos requeridos son los de relevancia y exclusión. Esto quiere decir que las variables exógenas Z_t que hemos escogido como instrumento deben cumplir $Cov(Z, P) \neq 0$ y $Cov(Z, V) = 0$

9. Suponga que la oferta de azúcar no es sensible al factor de lluvias debido a la gran cantidad de inventarios disponibles. ¿Podría estimar las ecuaciones de oferta y demanda en este contexto?

En ese caso, se puede omitir la variable L_t de la ecuación de oferta, lo cual haría que esta misma tenga una variable exógena menos. En este caso, se tiene:

$$k_{\text{oferta}} = g_{\text{-oferta}} + e_{\text{-oferta}} = 1 + 2 = 3 > g - 1 = 2$$

$$k_{\text{demanda}} = g_{\text{-demanda}} + e_{\text{-demanda}} = 1 + 0 = 1 < g - 1 = 2$$

Por lo tanto, se podría estimar la ecuación de oferta (sobre-identificada), pero no la de demanda (sub-identificada).

10. Considere el modelo AR(1) $Y_t = \alpha + \phi Y_{t-1} + U_t$.

- a) Dado que el estimador MCO de ϕ es sesgado, ¿sería posible utilizar la distribución t-Student para hacer inferencia estadística?
- b) ¿Cómo testearía la hipótesis nula de $\phi = 1$?

R: No es posible utilizar la distribución t-Student en un modelo como este. Si bien el estadístico de prueba es análogo al estadístico t , la distribución que se utiliza para hacer inferencia en estos casos es la distribución ADF (Augmented Dickey-Fuller).

- $H_0: \phi = 1$ o equivalentemente $\delta = 0$, donde $\Delta Y_t = \alpha + \underbrace{(\phi - 1)}_{\delta} Y_{t-1} + U_t$
- Estadístico de prueba $t = \frac{\hat{\delta}}{\overline{DS(\hat{\delta})}}$
- $RR = \{t | t < ADF(5\%)\}$

Si $t \in RR$ se rechaza la hipótesis nula y se concluye que no hay raíz unitaria. En caso contrario, no se puede descartar la presencia de raíz unitaria.

11. Suponga un modelo de rezagos distribuidos para explicar la evolución de la tasa de interés nominal de corto plazo en Chile, I_t en función de la inflación π_t del siguiente tipo

$$I_t = \alpha_0 + \alpha_1 I_{t-1} + \alpha_2 \pi_t + U_t$$

. Respecto a este modelo,

- a) Calcule cuál sería el impacto de largo plazo en la tasa de interés de un aumento permanente de la inflación en un punto porcentua.
- b) Especifique cómo podría estimar este impacto.

$$I_t = \alpha_0 + \alpha_1 I_{t-1} + \alpha_2 \pi_t + U_t$$

$$\Rightarrow I_t = \frac{\alpha_0 + \alpha_2 \pi_t + U_t}{1 - \alpha_1 L}$$

$$\Rightarrow \mathbb{E}(I_t | \pi_t, I_{t-1}, \dots) = \frac{\alpha_0 + \alpha_2 \pi_t}{1 - \alpha_1}$$

$$\Rightarrow \frac{\partial \mathbb{E}(I_t | \pi_t, I_{t-1}, \dots)}{\partial \pi_t} = \frac{\alpha_2}{1 - \alpha_1}$$

$$\Rightarrow \sum_{j=0}^{\infty} \frac{\partial \mathbb{E}(I_t | \pi_t, I_{t-1}, \dots)}{\partial \pi_t} = \alpha_2 \sum_{j=0}^{\infty} \left(\frac{1}{1 - \alpha_1} \right)^j$$

Este impacto se podría estimar llevando a cabo la regresión $I_t = \alpha_0 + \alpha_1 I_{t-1} + \alpha_2 \pi_t + U_t$, obteniendo así los estimadores $\hat{\alpha}_1$ y $\hat{\alpha}_2$, los que pueden ser reemplazados en la expresión anterior.

12. Si usted realiza un test de Breusch-Godfrey y rechaza que los errores U_t no siguen un proceso AR(1), explique

- a) Las consecuencias en la estimación del impacto de largo plazo de π sobre I .
- b) Las posibles soluciones que propone al eventual problema que esta situación genera.

La existencia de autocorrelación de los errores pueden ser vistos como especificación dinámica incompleta del modelo en un contexto de datos con persistencia. En efecto:

$$\begin{aligned}
 I_t &= \alpha_0 + \alpha_1 I_{t-1} + \alpha_2 \pi_t + U_t \\
 \Rightarrow I_t &= \alpha_0 + \alpha_1 I_{t-1} + \alpha_2 \pi_t + \rho_1 U_{t-1} + \varepsilon_t \\
 \Rightarrow I_t &= \alpha_0 + \alpha_1 I_{t-1} + \alpha_2 \pi_t + \rho_1 (I_{t-1} - \alpha_0 - \alpha_1 I_{t-2} - \alpha_2 \pi_{t-1}) + \varepsilon_t \\
 \Rightarrow I_t &= \alpha_0 (1 - \rho_1) + (\alpha_1 + \rho_1) I_{t-1} - \alpha_1 \rho_1 I_{t-2} + \alpha_2 \pi_t - \alpha_2 \rho_1 \pi_{t-1} + \varepsilon_t
 \end{aligned}$$

Lo anterior genera sesgo e inconsistencia en los estimadores obtenidos.

Una posible solución a este problema es la incorporación de aquellas variables omitidas: π_{t-1} e I_{t-2}

Formulario

- Propiedades de esperanza, varianzas, y covarianzas Si X e Y son variables aleatoria, y a, b, c, d son escalares reales,
 1. $E[aX + bY] = aE[X] + bE[Y]$
 2. $Var[aX + bY] = a^2Var[X] + b^2Var[Y] + 2abCov(X, Y)$
 3. $Cov(aX + c, bY + d) = abCov(X, Y)$
- Ley de expectativas iteradas: $E[Y] = E[E[Y|X]]$
- Linealidad de expectativa condicional: $E[YG(X)|X] = G(X)E[Y|X]$
- Estimador MCO bivariado: Dado $Y = \hat{\beta}_0 + \hat{\beta}_1 X + \hat{U}$, $\hat{\beta}_1 = \frac{\sum_n (X_n - \bar{X})(Y_n - \bar{Y})}{\sum_n (X_n - \bar{X})^2}$ y $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$.
- Estimador MCO multivariado: Se minimiza $\Phi(\hat{\beta}) = \sum_n (Y_n - \hat{\beta}_0 - \hat{\beta}_1 X_{n1} - \dots - \hat{\beta}_K X_{nK})^2$. El resultado obtenido es $\hat{\beta} = (X'X)^{-1}X'Y = (\sum_n X'_n X_n)^{-1}(\sum_n X'_n Y_n)$.
- MCO modelo en desvíos respecto a la media: $\hat{\beta}^* = (x'x)^{-1}x'y$ con $y = Y - \bar{Y}$ y $x_k = X_k - \bar{X}_k$.
- Varianza MCO: $Var(\hat{\beta}) = E[(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))'] = \sigma^2(X'X)^{-1}$. Modelo en desvíos respecto a la media $Var(\hat{\beta}^*) = \sigma^2(x'x)^{-1}$.
- Estimador varianza de errores: $\hat{\sigma}^2 = \frac{\sum_n \hat{U}_n^2}{N-K-1}$.
- Descomposición de varianza: $Y'Y - N\bar{Y}^2 = \hat{\beta}'X'X\hat{\beta} - N\bar{Y}^2 + \hat{U}'\hat{U}$ o $SCT = SCE + SCR$.
- Coeficiente de determinación $R^2 = 1 - SCR/SCT$, $R_a^2 = 1 - \frac{SCR/(N-K-1)}{SCT/(N-1)}$.
- Teorema de Frisch-Waugh: $\tilde{\beta}_k = \frac{Cov(\hat{U}_{Y,-k}, \hat{U}_{X_k,-k})}{Var(\hat{U}_{X_k,-k})}$ donde $\hat{U}_{Y,-k}$ es el residuo de la regresión entre Y y todos los X , salvo X_k , y $\hat{U}_{X_k,-k}$ es el residuo de la regresión entre X_k y todos los otros X .
- Matriz aniquiladora: $M_X = I_N - X(X'X)^{-1}X'$.
- Varianza estimador individual en modelo multivariado: $Var(\hat{\beta}_k) = \frac{\sigma^2}{\sum_n x_{nk}^2 (1-R_{k,-k}^2)}$.
- Varianza estimada del estimador individual en modelo multivariado: $\widehat{Var}(\hat{\beta}_k) = \frac{Var(Y)(1-R^2)}{(N-K)Var(X_k)(1-R_{k,-k}^2)}$.
- Ley de Grandes Números: Para una secuencia de variables aleatorias $Q_1, Q_2, \dots, Q_N$ con esperanza $E[Q]$, tenemos que $N^{-1} \sum_n Q_n \xrightarrow{P} E[Q]$
- Convergencia en probabilidad: Si $a, b \in \mathbb{R}$, $Q_n \xrightarrow{P} Q$ y $S_n \xrightarrow{P} S$, entonces
 1. $aQ_n + bS_n \xrightarrow{P} aQ + bS$
 2. $aQ_n S_n \xrightarrow{P} aQS$
 3. si $G(\cdot)$ es una función continua, entonces $G(Q_n) \xrightarrow{P} G(Q)$.

- Teorema Central del Límite: Si $Z_1, Z_2, \dots, Z_N$ son variables aleatorias con media $E(Z) < \infty$ y varianza $Var(Z) < \infty$, entonces $\sqrt{N}(\bar{Z} - E(Z)) \xrightarrow{d} \mathcal{N}(0, Var(Z))$.

- Teorema de Slutsky: Si $A_n \xrightarrow{p} A$ y $Z_n \xrightarrow{d} Z$, entonces

1. $A_n + Z_n \xrightarrow{d} A + Z$
2. $A_n Z_n \xrightarrow{d} AZ$

- Normalidad asintótica de MCO:

$$\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} \mathcal{N}(0, \sigma^2 E[X'X]^{-1})$$

- Distribución del estimador MCO, con errores poblacionales normales:

$$\hat{\beta}|X \sim \mathcal{N}(\beta, \sigma^2(X'X)^{-1})$$

- Distribución de estimador varianza de los errores: $\frac{(N-K)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(N-K)$.
- Distribución estimador individual en modelo multivariado: $\frac{\hat{\beta}_k - \beta_k}{\widehat{DS}(\hat{\beta}_k)}|X \sim \mathcal{T}_{N-K}$ con $\widehat{DS}(\hat{\beta}_k) = \hat{\sigma}\sqrt{\omega_{kk}}$ donde ω_{kk} es el elemento correspondiente en $(X'X)^{-1}$.
- Potencia de test de hipótesis con $H_0 : \beta_k = b_0$ vs $H_1 : \beta_k = b_1$.

$$\text{Potencia} = \beta(b_0, b_1) = F_T \left(\frac{b_0 - b_1}{\widehat{DS}(\hat{\beta}_k)} + \mathcal{T}_{1-\alpha/2, N-K} \right) - F_T \left(\frac{b_0 - b_1}{\widehat{DS}(\hat{\beta}_k)} - \mathcal{T}_{1-\alpha/2, N-K} \right)$$

- Test de restricciones lineales múltiples

$$F = \frac{1}{R\hat{\sigma}^2} (C\hat{\beta} - b)' (C(X'X)^{-1}C')^{-1} (C\hat{\beta} - b)$$

- Test de modelo restringido vs libre

$$F = \frac{(SCR_R - SCR_L)/R}{SCR_L/(N-K-1)} = \frac{(R_L^2 - R_R^2)/R}{(1 - R_L^2)/(N-K-1)}$$

- Test significancia global

$$\begin{aligned} F &= \frac{\hat{\beta}^{*'} x' x \hat{\beta}^*}{K\hat{\sigma}^2} \\ &= \frac{SCE/K}{SCR/(N-K)} = \frac{R^2}{1 - R^2} \left(\frac{N-K}{K} \right) \end{aligned}$$

- Matriz Varianza Robusta a Heterocedasticidad (White) $\widehat{Var}(\hat{\beta}|X) = (X'X)^{-1} X' \hat{\Omega} X (X'X)^{-1}$ con

$$\widehat{E}[UU'|X] = \hat{\Omega} = \begin{cases} \omega_{ij} = \hat{U}_n^2, & \text{if } i = j \quad \forall i = 1, 2, \dots, N \\ \omega_{ij} = 0, & \text{if } i \neq j \quad \forall i = 1, 2, \dots, N \end{cases}$$

- Mínimos Cuadrados Generalizados

$$\hat{\beta}_{MCG} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1}Y \quad \text{con } E[UU'|X] = \Omega$$

- Expectativa estimador MCO con variables omitidas (con X_{-k} como matriz de regresores donde la variable X_k está omitida)

$$\begin{aligned} E[\tilde{\gamma}|X] &= (X'_{-k}X_{-k})^{-1}X'_{-k}[X_{-k}|X_k]\beta \\ &= \beta + (X'_{-k}X_{-k})^{-1}X'_{-k}X_k\beta \end{aligned}$$

- Media y varianza incondicional de AR(1): $Y_t = \alpha + \rho_1 Y_{t-1} + U_t$.

$$\begin{aligned} E[Y_t] &= \alpha/(1 - \rho_1) \\ Var[Y_t] &= \frac{\sigma^2}{1 - \rho_1^2} \end{aligned}$$

- Forma matricial de un AR(p)

$$\begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix}_{p \times 1} = \begin{pmatrix} \rho_1 & \rho_2 & \dots & \rho_{p-1} & \rho_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}_{p \times p} \begin{pmatrix} Y_{t-1} \\ Y_{t-2} \\ \vdots \\ Y_{t-p} \end{pmatrix}_{p \times 1} + \begin{pmatrix} U_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{p \times 1}$$

- Descomposición espectral de matrices cuadradas $P = V\Lambda V^{-1}$. V es vector propio, y Λ es matriz diagonal con valores propios en ella.
- Estacionariedad de Y_t ssi $f(Y_t, Y_{t-1}, \dots, Y_{t-k}) = f(Y_{t-j}, Y_{t-1-j}, \dots, Y_{t-k-j})$ para todo $k, j \in \mathbb{Z}$.
- Estacionariedad débil de Y_t : $E[Y_t|t]$, $Var[Y_t|t]$ y $Cov[Y_t, Y_{t-k}|t]$ NO son función de t para cualquier t y k .
- Operadores de rezago : L es un operador tal que $L^j Y_t \equiv Y_{t-j}$.
- Secuencia geométrica

$$\theta(L)^{-1} = \frac{1}{1 - \rho L} = 1 + \rho L + \rho^2 L^2 + \dots$$

- Efectos acumulativos hasta un horizonte determinado en modelos de rezagos distribuidos.

$$\sum_{j=0}^J \frac{\partial E[Y_t|X_t, X_{t-1}, \dots]}{\partial X_{t-j}}$$

- Test Breusch-Godfrey

$$\hat{U}_t = \rho_1 \hat{U}_{t-1} + \rho_2 \hat{U}_{t-2} + \dots + \rho_P \hat{U}_{t-P} + V_t$$

y luego se construye el estadístico $BG = TR^2 \sim \chi_P^2$.

- Modelo con tendencia lineal (determinística)

$$Y_t = \alpha + \beta t + U_t \text{ con } E[U_t|t] = 0, Var(U_t) = \sigma^2, Cov(U_t, U_{t-j}) = 0 \forall j$$

- Camino aleatorio (tendencia estocástica)

$$Y_t = Y_{t-1} + U_t \text{ con } E[U_t|t] = 0, Var(U_t) = \sigma^2, Cov(U_t, U_{t-j}) = 0 \forall j$$

- Tests Dickey-Fuler Aumentado (ADF)

$$\Delta Y_t = \delta Y_{t-1} + U_t \text{ (simple)}$$

$$\Delta Y_t = \lambda + \delta Y_{t-1} + U_t \text{ (constante)}$$

$$\Delta Y_t = \pi t + \delta Y_{t-1} + U_t \text{ (tendencia lineal)}$$

$$\Delta Y_t = \lambda + \pi t + \delta Y_{t-1} + \sum_{k=1}^K \eta_k \Delta Y_{t-k} + U_t \text{ (dinámica de ajuste)}$$

- Intervalo confianza predicción fuera de muestra estática

$$IC(\hat{Y}_0) = \hat{Y}_0 \pm \mathcal{T}_{N-K-1, 1-\alpha/2} \times \hat{\sigma} \sqrt{1 + 1/T + (X_0 - \bar{X})(x'x)^{-1}(X_0 - \bar{X})'}$$

- Distribución predicción dinámica para un AR(1) (asumiendo parámetros conocidos)

$$Y_{t+k}|Y_t \sim \mathcal{N}\left(\frac{\alpha}{1-\rho} + \rho^k \left(Y_t - \frac{\alpha}{1-\rho}\right), \frac{\sigma^2(1-\rho^{2k})}{1-\rho^2}\right)$$

- Descomposición de Error Cuadrático Medio

$$ECM = \underbrace{(\bar{Y} - \widehat{\bar{Y}})^2}_{\text{Sesgo}} + \underbrace{(DS(Y) - DS(\widehat{Y}))^2}_{\text{Varianza}} + \underbrace{2(1 - Corr(Y, \widehat{Y}))DS(Y)DS(\widehat{Y})}_{\text{Covarianza}}$$

- Condición de orden para identificación de sistemas de ecuaciones: Sea g_i número de variables endógenas en i y e_{-i} número de variables exógenas que no están presentes en la ecuación i .

- Si $g_i - 1 = e_{-i}$, la ecuación está exactamente identificada.
- Si $g_i - 1 < e_{-i}$, la ecuación está sobreidentificada.
- Si $g_i - 1 > e_{-i}$, la ecuación está subidentificada (o no identificada).

- Estimador de variables instrumentales $\widehat{\beta}_{VI} = (Z'X)^{-1}Z'Y$.

- Supuesto de estimador de variables instrumentales para modelo $Y = X\beta + U$.

- Validez o Exclusión: $E[Z'U] = 0$.
- Relevancia o No-debilidad: $E[Z'X] \neq 0$.

- Estimador de varianza de VI: $\widehat{Var}(\widehat{\beta}_{VI}) = \widehat{\sigma}^2(Z'X)^{-1}Z'Z(Z'X)^{-1}$
con $\widehat{\sigma}^2 = \frac{1}{T-K-1}(Y - X\widehat{\beta}_{VI})'(Y - X\widehat{\beta}_{VI})$

- Estimador de Mínimos Cuadrados en Dos Etapas (MC2E) para modelo $Y = X\beta + U$ con matriz de instrumentos Z : $\widehat{\beta}_{MC2E} = (\widehat{X}'X)^{-1}\widehat{X}'Y = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'Y$.

- Estimador de varianza de MC2E: $\widehat{Var}(\widehat{\beta}_{MC2E}) = \widehat{\sigma}^2 X'Z(Z'Z)^{-1}Z'X$
con $\widehat{\sigma}^2 = \frac{1}{T-K-1}(Y - X\widehat{\beta}_{MC2E})'(Y - X\widehat{\beta}_{MC2E})$

- Error de medición clásico con modelo $Y = X^*\beta + U$ y medición imperfecta $X = X^* + \epsilon$.

- Sesgo de atenuación con errores de medición, modelo con un regresor: $\widehat{\beta} \xrightarrow{p} \beta \frac{Var(x^*)}{Var(x^*) + Var(\epsilon)}$.

- Efecto promedio de tratamiento $\tau_{ATE} = E[\tau_n] = E[Y_n(1) - Y_n(0)]$.
- Efecto promedio de tratamiento en los tratados $\tau_{ATT} = E[\tau_n|D_n = 1] = E[Y_n(1)|D_n = 1] - E[Y_n(0)|D_n = 1]$.
- Efecto promedio de tratamiento en los no tratados $\tau_{ATN} = E[\tau_n|D_n = 0] = E[Y_n(1)|D_n = 0] - E[Y_n(0)|D_n = 0]$.
- Modelo efectos fijos para datos de panel $Y_{nt} = \sum_n \alpha_n D_n + \delta_t + X_{nt}\beta + U_{nt} \quad \forall n = 1, \dots, N; t = 1, \dots, T$.
- Diferencias en diferencias: $(E[Y_2|D = 1] - E[Y_2|D = 0]) - (E[Y_1|D = 1] - E[Y_1|D = 0])$
- Regresión discontinua nítida: $Y_n = \beta I(X_n \geq \bar{X}) + \text{polinomio de } X_n + \text{controles} + U_n$. ($I(X_n \geq \bar{X})$ es la regla arbitraria de asignación al tratamiento).
- Regresión discontinua difusa: Etapa 1: $D_n = \alpha I(X_n \geq x_0) + \text{polinomio de } X_n + \text{controles} + U_n$ (D_n es tratamiento). Etapa 2: $Y_n = \beta \hat{D}_n + \text{polinomio de } X_n + \text{controles} + U_n$.