Guía de Ejercicios de Econometría I

Gabriel Montes-Rojas
gabriel.montes@fce.uba.ar

Pregunta 1: Valores predichos y errores de la regresión

Considere un modelo de regresión simple, es decir, $y_i = \beta_0 + x_i\beta_1 + u_i$, donde x_i es un escalar y u_i satisface los supuestos de Gauss-Markov. $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ es el **valor de predicción** de y dado x_i, esto es, un estimador de $E(y|x_i)$. $\hat{u}_i = y_i - \hat{y}_i$ es el **residuo de la regresión** o **error de predicción** para la observación i, o sea un estimador de $y_i - \beta_0 - \beta_1 x_i$.

1. Demostrar que $\sum_{i=1}^{n} \hat{u}_i = 0$ y $\sum_{i=1}^{n} x_i\hat{u}_i = 0$.

2. Demostrar que $E[\hat{u}_i|x] = 0$ y $E[x_i\hat{u}_i|x] = 0$ $\forall i$.

3. Demostrar que $\bar{y} = \bar{\hat{y}} = \frac{1}{n}\sum_{i=1}^{n} \hat{y}_i$.

4. Demostrar que $E[\hat{y}_i|x] = y_i$ $\forall i$.

5. Simule una muestra de datos con el modelo $y_i = \beta_0 + x_i\beta_1 + u_i$ donde $\beta_0 = \beta_1 = 1$, $u_i \sim$ iid $N(0,1)$ y $x_i \sim$ iid $N(0,1)$. Compruebe si se cumplen las igualdades.

   Nota: A continuación se resuelve para la primera parte.

```
clear
set obs 1000
gl beta0=1
gl beta1=1
gen u=rnormal(0,1)
gen x=rnormal(0,1)
gen y=$beta0+$beta1*x+u
reg y x
predict yhat
summ yhat
predict uhat, resid
summ uhat
gen xuhat=x*uhat
summ xuhat
```

Pregunta 2: Modelo con ordenada al origen

Considere un modelo de regresión sin intercepto, es decir, $y_i = x_i\beta + u_i$, donde x_i es un escalar.

1. Graficar una muestra $\{y_i, x_i\}_{i=1}^n$ con estas condiciones. ¿Qué restricciones tiene este modelo con respecto al modelo general? Este modelo también se llama de ordenada al origen.

2. Plantear el estimador de MCO sin intercepto como una minimización y mostrar que
$$\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

3. Presente el Teorema de Gauss-Markov para este modelo. Detalle todos los supuestos.

4. Pruebe que el estimador de mínimos cuadrados ordinarios (MCO) es insesgado. Indique claramente donde usa cada uno de los supuestos del punto anterior.

5. Derive la varianza de este estimador bajo los supuestos de Gauss-Markov.

6. Pruebe que el estimador de MCO es el que tiene la menor varianza dentro de la clase de estimadores insesgados $\sum_i c_i y_i$, donde c_i son constantes (pueden depender de x_i).

7. Simule una muestra de datos con el modelo $y_i = \beta_0 + x_i\beta_1 + u_i$ donde $\beta_0 = 0$, $\beta_1 = 1$, $u_i \sim$ *iid* $N(0,1)$ y $x_i \sim$ *iid* $N(0,1)$. Compruebe si $\hat{\beta}$ estima bien. Simule ahora $y_i = \beta_0 + x_i\beta_1 + u_i$ donde $\beta_0 = \beta_1 = 1$, $u_i \sim$ *iid* $N(0,1)$ y $x_i \sim$ *iid* $N(1,1)$. ¿Qué cambia?

   Nota: A continuación se resuelve para la primera parte.

   ```
   clear
   set obs 1000
   gl beta0=0
   gl beta1=1
   gen u=rnormal(0,1)
   gen x=rnormal(0,1)
   gen y=$beta0+$beta1*x+u
   reg y x, nocons
   reg y x
   ```

Pregunta 3: Endogeneidad

Considere un modelo de regresión sin intercepto, es decir,

$$MODELO\ 1: \ y_i = x_i\beta + w_i\gamma + e_i,$$

donde x_i y w_i son escalares.

1. Derive las expresiones de MCO para $\hat{\beta}$ y $\hat{\gamma}$.

2. Presente el Teorema de Gauss-Markov para este modelo. Detalle todos los supuestos.

3. Pruebe que el estimador de MCO es insesgado. Indique claramente donde usa cada uno de los supuestos del punto anterior.

4. Derive la varianza de $\hat{\beta}$ bajo los supuestos de Gauss-Markov.

5. Pruebe que el estimador de MCO $\hat{\beta}$ es el que tiene la menor varianza dentro de la clase de estimadores insesgados $\sum_i c_i y_i$, donde c_i con constantes (que pueden depender de (x_i, w_i)).

6. Use el Teorema de Frisch-Waugh-Lovell para derivar $\hat{\beta}$. Chequear este resultado usando una simulación en STATA.

```
clear
set obs 1000
gen u=rnormal(0,1)
gen x=rnormal(0,1)
gen w=rnormal(0,1)
gen y=x+w+u
reg y x w, nocons
reg y w, nocons
predict y1, resid
reg x w, nocons
predict x1, resid
reg y1 x1
```

7. Suponga un modelo alternativo que usa

$$MODELO\ 2: \ y_i = x_i\alpha + u_i,$$

es decir que omite w_i. Demuestre qué forma tiene el sesgo para estimar β y bajo qué casos no habría sesgo por omitir variables. También para este modelo derive la varianza de α y compárela con la de $\hat{\beta}$ del modelo con dos variables.

8. Suponga ahora un instrumento para x_i que soluciona el problema de endogeneidad del punto anterior. Explicite qué supuestos debería satisfacer este instrumento. Derive el estimador de variables instrumentales.

9. Simule una muestra de datos del MODELO 1 usando donde $\beta = \gamma = 1$ y con $e_i \sim \ iid\ N(0,1)$, x_i, w_i normales con media 0, varianza 1 y correlación 0.5.

   Nota: A continuación se resuelve para la primera parte.

```
clear
set obs 1000
gl beta0=1
gl beta1=1
gen e=rnormal(0,1)
gen u1=rnormal(0,1)
gen u2=rnormal(0,1)
gen w=u1
gen x=(u1+u2)/2^.5
corr x w
gen y=x+w+e
reg y x w
reg y x
gen z=(u2+rnormal(0,1))/2^.5
reg x z
ivreg y (x=z)
```

Pregunta 4: Modelos no lineales

Considere un modelo de regresión $y_i = \beta_0 + x_i\beta_1 + x_i^2\beta_2 + e_i$, donde $E[e|x] = 0i$.

1. Derive las expresiones de MCO para los $\hat{\boldsymbol{\beta}}$.

2. Presente el Teorema de Gauss-Markov para este modelo. Detalle todos los supuestos.

3. Pruebe que el estimador de MCO es insesgado. Indique claramente donde usa cada uno de los supuestos del punto anterior.

4. Suponga que se estima el modelo $y_i = \gamma_0 + \gamma_1 x_i + u_i$. Determine si los estimadores de MCO (γ_0, γ_1) son estimadores insesgados de (β_0, β_1). ¿Bajo qué condiciones lo son? Exprese este sesgo como uno de variables omitidas.

5. Simule una muestra de datos con $e_i \sim \ iid\ N(0,1)$, $x_i \sim \ iid\ N(\mu, 1)$ para distintos valores de los parámetros $\boldsymbol{\beta}$ y μ. Estime los modelos cuadráticos y lineales y compare.

Pregunta 5: Heterocedasticidad

Considere un modelo de regresión simple, es decir, $y_i = \beta_0 + x_i\beta_1 + (1 + x_i\gamma)u_i$, donde x_i es un escalar y u_i es iid e independiente de x_i.

1. Pruebe que el estimador de MCO es insesgado. Indique claramente donde usa cada uno de los supuestos del Teorema de Gauss-markov.

2. Derive la varianza de $\hat{\beta}$ y compárela con la derivada bajo todos los supuestos de Gauss-Markov.

3. Simule una muestra de datos con los modelos anteriores usando donde $\beta_0 = \beta_1 = \gamma = 1$ y con $e_i \sim \ iid\ N(0, 1)$, $x_i \sim \ iid\ N(0, 1)$.