

PAUTA Auxiliar 1 - Semestre Otoño 2016

15 de Marzo, 2016

Problema 1: Definiciones

1. Defina:

a. Estimador

R:

Es una función, (o regla, fórmula, método) la cual a partir de los datos obtenidos en la muestra, o parámetros muestrales, permite determinar un parámetro poblacional (también se puede entender como parámetro 'natural', 'fundamental', verdadero o por el estilo). Este valor es el que mejor caracteriza la muestra, como si condensara toda la probabilidad en un único punto. Es importante tener en cuenta que al ser función de la muestra, corresponde a una variable aleatoria, y por lo tanto posee las características correspondientes, como distribución, esperanza, varianza, etc.

b. Esperanza y Varianza (de un estimador)

R:

Entendiendo que un estimador es una variable aleatoria al estar en función de la muestra, la esperanza es entendida como el valor que se esperaría obtener una vez que se hace una realización de la variable aleatoria (es decir, que ocurra el fenómeno, que se saque un sujeto al azar, etc). La varianza se entiende como la desviación de los datos con respecto a la media. Es el valor esperado de las diferencias de todos los datos con el valor esperado de la variable aleatoria. En el caso del estimador, permite determinar qué tan distantes del valor poblacional a estimar se encuentran las estimaciones realizadas.

c. Sesgo y consistencia

R:

Ambas son características propias de los estimadores. El sesgo se define como la diferencia entre el valor del estimador y el parámetro poblacional: $S = \mathbb{E}(\hat{\beta}) - \beta$, por eso se define como insesgado un estimador donde $\mathbb{E}(\hat{\beta}) = \beta$. Por su parte, si ese mismo sesgo converge en plim (límite en probabilidad) al verdadero valor del parámetro. Técnicamente "Se dice que la sucesión de variables aleatorias X_1, X_2, \dots, X_n converge en probabilidad a la v.a. o constante X si:

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} Pr[|X_n - X| < \varepsilon] = 1$$

Conviene saber que: $plim\left(\frac{X}{Y}\right) = \frac{plim X}{plim Y}$ y que $plim(X \cdot Y) = plim(X) \cdot plim(Y)$.

- d. Diferencia entre u y \hat{u}

R:

El primero, u corresponde a un término aleatorio de la función de regresión poblacional que es desconocido y absoluto. Este término es imposible de medir, y es sólo una corrección a la ecuación (para evitar poner $Y > \beta X$ por ejemplo). Es lo que afecta a la variable dependiente pero no es observado ni medido. Y se considera aleatorio. Uno de los supuestos fundamentales consiste en suponer que dicho efecto es independiente de las variables explicativas. Por otra parte, el término \hat{u} corresponde al error muestral. Que se define entonces como $\hat{u} = Y - \hat{\beta}X = Y - \hat{Y}$, desde donde se puede ver que corresponde a la diferencia entre el valor poblacional absoluto y verdadero de Y y el valor estimado de la función de regresión muestral. Es entonces la distancia que existe entre la regresión promedio, la función, y los valores observados, los datos.

Problema 2: Características MCO

1. Identificando los supuestos de Gauss-Markov relevantes, muestre que el estimador MCO es insesgado y consistente, y calcule su varianza.

R:

Utilizando la definición del estimador obtenida:

$$\begin{aligned}\mathbb{E}(\beta|X_i) &= \mathbb{E}\left(\frac{\sum_i x_i y_i}{\sum_i x_i^2}\right) \\ &= \mathbb{E}\left(\frac{\sum_i x_i (\beta_1 x_i + u_i)}{\sum_i x_i^2}\right) \\ &= \mathbb{E}\left(\frac{\beta_1 \sum_i x_i^2 + \sum_i x_i u_i}{\sum_i x_i^2}\right) \\ &= \beta_1 + \frac{\sum_i \mathbb{E}(x_i u_i)^*}{\sum_i x_i^2} = \beta\end{aligned}$$

En * se cumple el supuesto de media de errores nula, u ortogonalidad de los errores con las variables explicativas, es decir: $\mathbb{E}(UX) = 0$. Dado que no hay sesgo, se concluye también la consistencia.

La varianza puede ser obtenida de la siguiente manera, teniendo en cuenta que se condiciona en X_i (es decir que se fija el X , trabajándose como constante):

$$\begin{aligned}
\mathbb{V}(\hat{\beta}|X_i) &= \frac{1}{(\sum_{i=1}^n X_i^2)^2} \mathbb{V}\left(\sum_{i=1}^n X_i Y_i\right) \\
&= \frac{1}{(\sum_{i=1}^n X_i^2)^2} \mathbb{V}\left(\sum_{i=1}^n \beta X_i^2 + X_i u_i\right) \\
&= \frac{1}{(\sum_{i=1}^n X_i^2)^2} \sum_{i=1}^n \mathbb{V}(X_i u_i) \\
&= \frac{\sigma^2}{(\sum_{i=1}^n X_i^2)^2} \sum_{i=1}^n X_i^2
\end{aligned}$$

***Los demás supuestos se pueden ver en la linealidad de los parámetros, la NO multicolinealidad exacta, la varianza constante de los errores, autocorrelación nula de los errores (estos dos resumidos en la matriz esférica de errores), además de la no-singularidad de la matriz y que el número de regresores sea menor que el número de datos!**

2. Calcule la esperanza y varianza de los siguientes estimadores y compare con MCO.

Suponga que se cumplen los supuestos para todos los estimadores.

$$\begin{aligned}
\text{I. } \hat{\beta}_1 &= \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i} \\
\text{II. } \hat{\beta}_2 &= \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i}
\end{aligned}$$

R:

Para el primer estimador, se tiene que su esperanza es

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_1|X) &= \frac{1}{\sum_{i=1}^n X_i} \mathbb{E}\left(\sum_{i=1}^n Y_i\right) \\
&= \frac{1}{\sum_{i=1}^n X_i} \left(\sum_{i=1}^n \mathbb{E}(\beta X_i + u_i)\right) \\
&= \beta
\end{aligned}$$

Por lo que $\hat{\beta}_1$ es un estimador insesgado de β . Por otro lado, su varianza esta dada por

$$\begin{aligned}
\mathbb{V}(\hat{\beta}_1|X) &= \frac{1}{(\sum_{i=1}^n X_i)^2} \mathbb{V}\left(\sum_{i=1}^n Y_i\right) \\
&= \frac{1}{(\sum_{i=1}^n X_i)^2} \mathbb{V}\left(\sum_{i=1}^n (\beta X_i + u_i)\right) \\
&= \frac{1}{(\sum_{i=1}^n X_i)^2} \left(\sum_{i=1}^n \mathbb{V}(u_i)\right) = \\
&= \frac{n\sigma^2}{(\sum_{i=1}^n X_i)^2}
\end{aligned}$$

Notar que la varianza en este caso depende de n , lo que quiere decir que a medida que aumentemos el número de datos, la varianza aumenta! ¿Qué tan bueno será este estimador?

Para el segundo estimador, se tiene que la esperanza esta dada por

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_2) &= \frac{1}{\sum_{i=1}^n X_i} \mathbb{E}\left(\sum_{i=1}^n X_i Y_i\right) \\
&= \frac{1}{\sum_{i=1}^n X_i} \mathbb{E}\left(\sum_{i=1}^n \beta X_i^2 + X_i u_i\right) \\
&= \frac{1}{\sum_{i=1}^n X_i} \mathbb{E}\left(\sum_{i=1}^n \beta X_i^2\right) \\
&= \beta \frac{\sum_{i=1}^n X_i^2}{\sum_{i=1}^n X_i}
\end{aligned}$$

Por lo que el estimador es sesgado, De igual forma, calculamos su varianza

$$\begin{aligned}
\mathbb{V}(\hat{\beta}_2) &= \frac{1}{(\sum_{i=1}^n X_i)^2} \mathbb{V}\left(\sum_{i=1}^n X_i Y_i\right) \\
&= \frac{1}{(\sum_{i=1}^n X_i)^2} \mathbb{V}\left(\sum_{i=1}^n X_i^2 + X_i u_i\right) \\
&= \frac{1}{(\sum_{i=1}^n X_i)^2} \sum_{i=1}^n \mathbb{V}(X_i u_i) \\
&= \frac{\sigma^2}{(\sum_{i=1}^n X_i)^2} \sum_{i=1}^n X_i^2
\end{aligned}$$

al compararse con la varianza del estimador MCO, esta última posee un término positivo que la hacer mayor. En ambos estimadores, la varianza sólo empeora, y se puede ver que entre los 3, el MCO es más eficiente.