IN4402 - Aplicaciones de Probabilidades y Estadística en Gestión

Ronald Leblebici - Angelo Muñoz

ronald.leblebici@gmail.com

gelox97@gmail.com

Facultad de Ciencias Físicas y Matemáticas
Universidad de Chile

24 de septiembre del 2018

¿Qué estudiaremos hoy?

P1. Teorema de Gauss-Markov

P2. Teorema de Frisch-Waugh / Multicolinealidad

P3. Bondad de ajuste y varianza de los estimadores

P1. Teorema de Gauss-Markov

[P4, Control 1, 2017] Comente la siguiente afirmación : "Si los supuestos del teorema de Gauss-Markov no se cumplen, entonces el modelo de regresión múltiple no es una buena aproximacion a la expectativa condicional de Y dado X, $\mathbb{E}(Y|X)$".

P1. Teorema de Gauss-Markov

[P4, Control 1, 2017] Comente la siguiente afirmación : "Si los supuestos del teorema de Gauss-Markov no se cumplen, entonces el modelo de regresión múltiple no es una buena aproximacion a la expectativa condicional de Y dado X, $\mathbb{E}(Y|X)$".

Respuesta:

Verdadero. Los supuestos del teorema de Gauss-Markov garantizan que $\hat{\beta}_{MCO}$ es el estimdor lineal insesgado de mínima varianza (menor incertidumbre). Por lo tanto, la estimación $\hat{Y}|X = X\beta_{MCO}$ representa la predicción más precisa de $E[Y|X]$. Si se viola(n) alguno(s) de los supuestos, se pierde esta propiedad, generando sesgo y/o ineficiencia.

P1. Teorema de Gauss-Markov

[P4, Control 1, 2014] Suponga el modelo de regresión lineal simple $Y_n = \beta_1 + \beta_2 X_n + U_n$, y considere el siguiente estimador para el parámetro β_2: $\dot{\beta}_2 = \sum_n Y_n / \sum_n X_n$.

P1. Teorema de Gauss-Markov

[P4, Control 1, 2014] Suponga el modelo de regresión lineal simple $Y_n = \beta_1 + \beta_2 X_n + U_n$, y considere el siguiente estimador para el parámetro β_2: $\dot{\beta}_2 = \sum_n Y_n / \sum_n X_n$.

a. Evalúe si el estimador $\dot{\beta}_2$ es lineal e insesgado. Explicite cualquier supuesto que deba realizar para demostrar o refutar esta propiedad.

P1. Teorema de Gauss-Markov

[P4, Control 1, 2014] Suponga el modelo de regresión lineal simple $Y_n = \beta_1 + \beta_2 X_n + U_n$, y considere el siguiente estimador para el parámetro β_2: $\dot{\beta}_2 = \sum_n Y_n / \sum_n X_n$.

a. Evalúe si el estimador $\dot{\beta}_2$ es lineal e insesgado. Explicite cualquier supuesto que deba realizar para demostrar o refutar esta propiedad.

b. Dados los supuestos realizados, ¿podría proponer usted un mejor estimador que $\dot{\beta}_2$?. ¿Qué estimador sería este? ¿En qué fundamenta esta propuesta?

P1. Teorema de Gauss-Markov

[P4, Control 1, 2014] Suponga el modelo de regresión lineal simple $Y_n = \beta_1 + \beta_2 X_n + U_n$, y considere el siguiente estimador para el parámetro β_2: $\dot{\beta}_2 = \sum_n Y_n / \sum_n X_n$.

a. Evalúe si el estimador $\dot{\beta}_2$ es lineal e insesgado. Explicite cualquier supuesto que deba realizar para demostrar o refutar esta propiedad.

b. Dados los supuestos realizados, ¿podría proponer usted un mejor estimador que $\dot{\beta}_2$?. ¿Qué estimador sería este? ¿En qué fundamenta esta propuesta?

c. Estime la varianza del estimador $\dot{\beta}_2$ y obtenga una expresión numérica para el nivel relativo de ineficiencia de dicho estimador respecto al MCO, es decir $\mathbb{V}(\dot{\beta}_2)/\mathbb{V}(\widehat{\beta}_2)$ ¿Hay alguna circunstancia en la que $\dot{\beta}_2$ sea tan bueno como $\widehat{\beta}_2$?

P2. Teorema de Frisch-Waugh / Multicolinealidad

[P5, Control 1, 2014] Un empresario productor de leche desea estudiar si dos tipos de forraje (alimento) X y Z logran elevar la producción de leche Y de 250 vacas que tiene en su predio.

P2.1. Teorema de Frisch-Waugh / Multicolinealidad

Suponga que a usted le pidieran diseñar el experimento para estudiar el efecto de los dos tipos de forraje X y Z sobre la producción de leche. Explique cómo asignaría cantidades de X y Z a las vacas para determinar los efectos de X y Z sobre la producción de leche con la mayor precisión.

P2.1. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Se tiene que la varianza del estimador asociado a un coeficiente en una regresión múltiple está dado por:

$$\widehat{\mathbb{V}}(\widehat{\beta}_A) = \frac{\mathbb{V}(Y)(1-R^2)}{(N-K)\mathbb{V}(X_A)} \cdot \frac{1}{1-R^2_{A,-A}}$$

La mejor estimación posible se obtiene al diseñar el experimento de modo tal que la varianza del estimador se reduzca lo más posible. Esto puede hacerse de dos formas:

P2.1. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Se tiene que la varianza del estimador asociado a un coeficiente en una regresión múltiple está dado por:

$$\widehat{\mathbb{V}}(\widehat{\beta}_A) = \frac{\mathbb{V}(Y)(1-R^2)}{(N-K)\mathbb{V}(X_A)} \cdot \frac{1}{1-R^2_{A,-A}}$$

La mejor estimación posible se obtiene al diseñar el experimento de modo tal que la varianza del estimador se reduzca lo más posible. Esto puede hacerse de dos formas: (1) asignando cantidades de X y Z a las vacas de forma tal que el factor inflador de varianza sea 1 (o correlación entre X y Z sea 0)

P2.1. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Se tiene que la varianza del estimador asociado a un coeficiente en una regresión múltiple está dado por:

$$\widehat{\mathbb{V}}(\widehat{\beta}_A) = \frac{\mathbb{V}(Y)(1-R^2)}{(N-K)\mathbb{V}(X_A)} \cdot \frac{1}{1-R^2_{A,-A}}$$

La mejor estimación posible se obtiene al diseñar el experimento de modo tal que la varianza del estimador se reduzca lo más posible. Esto puede hacerse de dos formas: (1) asignando cantidades de X y Z a las vacas de forma tal que el factor inflador de varianza sea 1 (o correlación entre X y Z sea 0) y (2) Elevando al máximo posible la variabilidad de X y Z.

P2.1. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Se tiene que la varianza del estimador asociado a un coeficiente en una regresión múltiple está dado por:

$$\widehat{\mathbb{V}}(\widehat{\beta}_A) = \frac{\mathbb{V}(Y)(1 - R^2)}{(N - K)\mathbb{V}(X_A)} \cdot \frac{1}{1 - R^2_{A,-A}}$$

La mejor estimación posible se obtiene al diseñar el experimento de modo tal que la varianza del estimador se reduzca lo más posible. Esto puede hacerse de dos formas: (1) asignando cantidades de X y Z a las vacas de forma tal que el factor inflador de varianza sea 1 (o correlación entre X y Z sea 0) y (2) Elevando al máximo posible la variabilidad de X y Z. (En la práctica esto último puede tener algunos inconvenientes no estipulados con claridad, ya que una variabilidad demasiado extrema podría dejar sin comer a las vacas o tener un costo financiero demasiado alto.)

P2.2. Teorema de Frisch-Waugh / Multicolinealidad

El empresario siguió su diseño experimental, pero encP2. Teorema de Frisch-Waugh / Multicolinealidadarga a un veterinario amigo que genere los resultados. El veterinario, solo sabe de regresion simple, por lo cual obtiene los siguientes resultados de dos regresiones donde las desviaciones estandar de los coeficientes están en el paréntisis bajo los respectivos coeficientes estimados. ¿Qué puede concluir sobre el verdadero impacto de X y Z sobre la producción de leche?

$\widehat{Y}_n = \underset{(0{,}21)}{5,61} + \underset{(0{,}03)}{0,51}X_n,\ R^2 = 0,15$	$\widehat{Y}_n = \underset{(0{,}18)}{5,55} + \underset{(0{,}12)}{0,22}Z_n,\ R^2 = 0,08$

P2.2. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Como mediante el diseño experimental anterior se logró que las variables no estuviesen correlacionadas, el resultado que se hubiese obtenido por regresión múltiple sobre cada una de las variables es el mismo que se obtuvo al hacer una regresión bivariada.

P2.2. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Como mediante el diseño experimental anterior se logró que las variables no estuviesen correlacionadas, el resultado que se hubiese obtenido por regresión múltiple sobre cada una de las variables es el mismo que se obtuvo al hacer una regresión bivariada. Por lo tanto, los efectos estimados de X y Z por el veterinario en dos regresiones bivariadas son efectivamente los estimadores MCO del modelo de regresión múltiple. El veterinario obtiene la respuesta correcta.

P2.3. Teorema de Frisch-Waugh / Multicolinealidad

Suponga que en lugar de hacerle caso a ud., el empresario sigue el consejo de su amigo veterinario para diseñar el experimento. El veterinario sugiere que se asignen cantidades de X y Z aleatoriamente a cada vaca, completando exactamente 20 kilos de forraje en total. (Sugerencia: escriba la matriz de regresores). ¿Sería posible determinar los efectos individuales de X y Z bajo este diseño experimental? ¿Cómo cambia su respuesta si la cantidad de X más Z suman 20 con error aleatorio?

P2.3. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Bajo este diseño experimental no es posible determinar los efectos individuales de X y Z, ya que X y Z tienen una correlación igual a -1, ya que $Z = 20 - X$

P2.3. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Bajo este diseño experimental no es posible determinar los efectos individuales de X y Z, ya que X y Z tienen una correlación igual a -1, ya que $Z = 20 - X$, con lo que su correlación es perfecta y la varianza asociada al coeficiente será infinita de acuerdo a la fórmula ya mostrada

P2.3. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Bajo este diseño experimental no es posible determinar los efectos individuales de X y Z, ya que X y Z tienen una correlación igual a -1, ya que $Z = 20 - X$, con lo que su correlación es perfecta y la varianza asociada al coeficiente será infinita de acuerdo a la fórmula ya mostrada , pero más aún el estimador MCO no está definido. No pudiendo verse así el efecto individual de cada variable.

P2.3. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Bajo este diseño experimental no es posible determinar los efectos individuales de X y Z, ya que X y Z tienen una correlación igual a -1, ya que $Z = 20 - X$, con lo que su correlación es perfecta y la varianza asociada al coeficiente será infinita de acuerdo a la fórmula ya mostrada , pero más aún el estimador MCO no está definido. No pudiendo verse así el efecto individual de cada variable. Mientras que si ambas suman 20 con un error aleatorio, será posible determinar los efectos de ambos, ya que habrá parte de la varianza sobre la producción de leche asociada tan sólo a cada variable (esto se deduce del teorema de Friesch-Waugh).

P2.3. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Bajo este diseño experimental no es posible determinar los efectos individuales de X y Z, ya que X y Z tienen una correlación igual a -1, ya que $Z = 20 - X$, con lo que su correlación es perfecta y la varianza asociada al coeficiente será infinita de acuerdo a la fórmula ya mostrada , pero más aún el estimador MCO no está definido. No pudiendo verse así el efecto individual de cada variable. Mientras que si ambas suman 20 con un error aleatorio, será posible determinar los efectos de ambos, ya que habrá parte de la varianza sobre la producción de leche asociada tan sólo a cada variable (esto se deduce del teorema de Friesch-Waugh). Sin embargo existirá una correlación negativa entre X y Z, por lo cual el diseño experimental no será el ideal con ${R^2_{-}}_X = 0$.

P2.3. Teorema de Frisch-Waugh / Multicolinealidad

Respuesta:

Bajo este diseño experimental no es posible determinar los efectos individuales de X y Z, ya que X y Z tienen una correlación igual a -1, ya que $Z = 20 - X$, con lo que su correlación es perfecta y la varianza asociada al coeficiente será infinita de acuerdo a la fórmula ya mostrada , pero más aún el estimador MCO no está definido. No pudiendo verse así el efecto individual de cada variable. Mientras que si ambas suman 20 con un error aleatorio, será posible determinar los efectos de ambos, ya que habrá parte de la varianza sobre la producción de leche asociada tan sólo a cada variable (esto se deduce del teorema de Friesch-Waugh). Sin embargo existirá una correlación negativa entre X y Z, por lo cual el diseño experimental no será el ideal con $R^2_{-X} = 0$. Habrá mayor varianza que en el diseño anterior, y no se encontrarán respuestas adecuadas con regresión bivariada.

P3. Bondad de ajuste y varianza de los estimadores

1. Ejecute ambas regresiones en *Stata* (con y sin el ingreso total del hogar) y analice el efecto de incluir esta variable sobre las desviaciones estándar de los estimadores y los coeficientes R^2 y R_a^2.

P3. Bondad de ajuste y varianza de los estimadores

1. Ejecute ambas regresiones en *Stata* (con y sin el ingreso total del hogar) y analice el efecto de incluir esta variable sobre las desviaciones estándar de los estimadores y los coeficientes R^2 y R_a^2.
2. Es su último día de práctica usted ha llegado a casa feliz pues le fue bastante bien. Además, como usted es un(a) alumno(a) responsable empezó a hacer su informe de práctica con anticipación, por lo que está casi listo(a) para salir de vacaciones. Solo le falta reportar un par de resultados que tiene impresos en su mochila. Sin pensarlo dos veces usted la abre y se da cuenta que se le reventó un lapiz cuya tinta no le permite ver algunos datos. Usted no tiene *Stata* instalado en su computador, sin embargo sabe utilizar las fórmulas necesarias y cuenta con *Matlab*. Obtenga los datos faltantes para poder salir de vacaciones.

¿Qué aprendimos hoy?

- Repaso de estimadores (linealidad, sesgo, consistencia, etc.).

¿Qué aprendimos hoy?

- Repaso de estimadores (linealidad, sesgo, consistencia, etc.).
- Aprendimos sobre el teorema de Gauss-Markov y algunas consecuencias del incumplimiento de los supuestos de MCO.

¿Qué aprendimos hoy?

- Repaso de estimadores (linealidad, sesgo, consistencia, etc.).
- Aprendimos sobre el teorema de Gauss-Markov y algunas consecuencias del incumplimiento de los supuestos de MCO.
- Cómo considerar la colinealidad al momento de diseñar experimentos y cuáles son los efectos de algunas decisiones que podemos tomar en este proceso de diseño.

¿Qué aprendimos hoy?

- Repaso de estimadores (linealidad, sesgo, consistencia, etc.).
- Aprendimos sobre el teorema de Gauss-Markov y algunas consecuencias del incumplimiento de los supuestos de MCO.
- Cómo considerar la colinealidad al momento de diseñar experimentos y cuáles son los efectos de algunas decisiones que podemos tomar en este proceso de diseño.
- Entendemos cómo se descompone el estimador de varianza de β_{MCO} y qué es el factor inflador de varianza (FIV).

¿Qué aprendimos hoy?

- Repaso de estimadores (linealidad, sesgo, consistencia, etc.).
- Aprendimos sobre el teorema de Gauss-Markov y algunas consecuencias del incumplimiento de los supuestos de MCO.
- Cómo considerar la colinealidad al momento de diseñar experimentos y cuáles son los efectos de algunas decisiones que podemos tomar en este proceso de diseño.
- Entendemos cómo se descompone el estimador de varianza de β_{MCO} y qué es el factor inflador de varianza (FIV).
- Nuevamente, somos capaces de calcular e interpretar valores que *Stata* nos entrega. En este caso vimos cómo obtener los estimadores de varianza de $\widehat{\beta}_{\text{MCO}}$ y los coeficientes de bondad de ajuste R^2 y R_a^2.

¿Qué aprendimos hoy?

- Repaso de estimadores (linealidad, sesgo, consistencia, etc.).
- Aprendimos sobre el teorema de Gauss-Markov y algunas consecuencias del incumplimiento de los supuestos de MCO.
- Cómo considerar la colinealidad al momento de diseñar experimentos y cuáles son los efectos de algunas decisiones que podemos tomar en este proceso de diseño.
- Entendemos cómo se descompone el estimador de varianza de β_{MCO} y qué es el factor inflador de varianza (FIV).
- Nuevamente, somos capaces de calcular e interpretar valores que *Stata* nos entrega. En este caso vimos cómo obtener los estimadores de varianza de $\widehat{\beta}_{\text{MCO}}$ y los coeficientes de bondad de ajuste R^2 y R_a^2.