Material para las clases teóricas virtuales

Clase 6: Teorema de Gauss-Markov e Inferencia en el modelo lineal general. (Link a lista de reproducción) (Link de la lista de reproducción)

1. Presentación en pdf para imprimir disponible en este link

2. Videos (son 4 videos en total):
   - **Video 1 – Repaso**
     Se repasa el modelo con K variables y los estimadores MCO usando notación matricial. Se discute la propiedad de semidefinición positiva de la matriz de varianzas y covarianzas.

     Preguntas guía para el video 1:
     a) En la página 7 de las filminas, el supuesto 2 dice que $VarCov(u) = \sigma^2 I_n$. ¿Por qué decimos que esto implica suponer homocedasticidad y también ausencia de correlación serial?
     b) En la página 8 de las filminas se introduce la siguiente notación: A_{kl} es el elemento en la fila *k* y columna *l* de la matriz $(\mathrm{X'X})^{-1}$. Con esa notación, ¿cuál es la expresión para la varianza de $\hat{\beta}_2$? ¿y cuál es la expresión para la covarianza entre $\hat{\beta}_2$ y $\hat{\beta}_5$? ¿y cuáles son las expresiones para los estimadores de esa varianza y de esa covarianza?
     c) En la página 10 de las filminas se presenta un ejemplo para K=2. Se pide desarrollar un ejemplo similar, pero para K=3: (i) Desarrollar la expresión $c'\hat{\beta}$; notar que se trata de una variable aleatoria unidimensional. (ii) obtener la expresión para la varianza de $c'\hat{\beta}$ mediante el desarrollo de la forma cuadrática de $c'VarCov(\hat{\beta})c$. (iii) Reescribir esa expresión usando la notación del punto (b). (iv) Obtener la expresión para la varianza estimada de $c'\hat{\beta}$ si los elementos del vector c son los siguientes: $c_1 = 0,\ c_2 = 1,$ $c_3 = 120$.

   - **Video 2 – Teorema de Gauss-Markov**
     Se discute la intuición y uso práctico del Teorema de Gauss-Markov en el modelo lineal general. Se presenta el enunciado formal del teorema y su demostración. La demostración propiamente dicha está disponible aquí.

     Preguntas guía para el video 2:
     a) Para K=2, mostrar que la definición semipositiva de $VarCov(\tilde{\beta}) - VarCov(\hat{\beta})$ implica que la varianza de cualquiera de los $\hat{\beta}_k$ con *k=1, 2* (o de cualquier combinación lineal de los $\hat{\beta}_k$) es menor o igual que la varianza de cualquiera de los $\tilde{\beta}_k$ (o de la correspondiente combinación lineal).
     b) ¿Por qué en (a) decimos menor o igual? ¿En qué caso serían iguales?
     c) En la página 14 de las filminas, aparece la matriz COV($\hat{\beta}$, $\hat{\gamma}$). ¿Qué dimensión tiene esa matriz y cuáles son los elementos que la componen? Tal vez ayude si empezás pensando en el caso de K=1, luego K=2 y recién después generalizás para cualquier K.

d) Al final de la página 14 dice que si probamos que COV($\hat{\beta}$, $\hat{\gamma}$)=0 probaremos el teorema, ¿por qué? Asegurate también de entender los porqués de la página 15.

e) A partir del minuto 10:50 del video, la profesora dice "sabiendo que uno de los supuestos no se cumple, sabemos que entonces va a haber otro estimador lineal e insesgado que no es el mínimo cuadrático y que va a ser más eficiente que el mínimo cuadrático". Esto no está mal, pero es incompleto. Déjenme aclararlo. Como los supuestos son necesarios, cuando uno de los supuestos no se cumple, sabemos que no se cumple el resultado del Teorema, es decir, el estimador mínimo cuadrático dejará de ser MELI. Ser MELI implica simultáneamente 3 cosas: ser Lineal, ser Insesgado y ser el Mejor (en el sentido de menor varianza). Como por construcción los estimadores MCO siempre son lineales, que algún supuesto no se cumpla implica que (i) los estimadores mínimo-cuadráticos dejen de ser insesgados o (ii) que dejen de ser los más eficientes entre los estimadores lineales e insesgados (o sea, que habrá otro estimador lineal e insesgado con menor varianza que el de mínimos cuadrados).

f) Si en el modelo lineal general se cumplen todos los supuestos clásicos y tenemos, por un lado, a $\hat{\beta}$ que es el vector de estimadores MCO. También tenemos otro vector de estimadores insesgados, al que llamamos $\breve{\beta}$. ¿Cuál tiene menor varianza?

- **Video 3 – Tests de hipótesis simples**

Se presenta la formal general para expresar los tests de hipótesis simples usando notación matricial. Se ven casos particulares que incluyen tests sobre parámetros individuales y tests sobre combinaciones lineales de parámetros.

Preguntas guía para el video 3:

a) Explicar por qué el supuesto de normalidad del término aleatorio del modelo implica que el vector de estimadores MCO $\hat{\beta}$ tiene distribución normal multivariada.

b) En la página 19 de las filminas se presentan varios casos particulares de hipótesis nulas simples. ¿Cuál es el vector *c* y el número *r* que corresponden a cada caso?

c) De la página 20: prestar mucha atención en que $\hat{\beta}$ tiene distribución normal multivariada mientras que c'$\hat{\beta}$ tiene distribución normal univariada. Asegurarse de entender la diferencia y el por qué.

d) Con relación a lo que se explica en la página 22 de las filminas, asegurarse de entender por qué decimos que el vector *c* "selecciona" algunas filas y columnas de la matriz $(\mathrm{X'X})^{-1}$. Para eso, puede ayudar concentrarse otra vez en un caso simplificado, por ejemplo, para K=3. Escribir la matriz $(\mathrm{X'X})^{-1}$ usando la notación que vimos en la página 8, es decir, llamar A_{kl} al elemento en la fila *k* y columna *l* de la matriz $(\mathrm{X'X})^{-1}$. Luego desarrollar el producto $c'(\mathrm{X'X})^{-1}$c. Notar que si $c_1 = 1\ y\ \ c_2 = c_3 = 0$ ese producto da por resultado A_{11}, es decir, el elemento de la fila 1 y columna 1 de $(\mathrm{X'X})^{-1}$. Probar con $c_2 = 1\ y\ \ c_1 = c_3 = 0,\ \ y\ con$ si $c_3 = 1\ y\ \ c_1 = c_2 =$

0 . Después probar con vectores que "seleccionan" más de una fila/columna. Por ejemplo, $c_1 = 1 \ y \ c_2 = -1 \ y \ c_3 = 0$.

- **Video 4 – Tests de hipótesis compuestas**

  Se presentan los contrastes de hipótesis para evaluar hipótesis nulas compuestas. En particular, se ven los tests de significatividad global y de significatividad conjunta de un subconjunto de variables. Por último, se presenta un ejemplo empírico donde se aplican los contenidos de la clase.

  Preguntas guía para el video 4:

  a) Respecto del test de significatividad global (página 24 de las filminas). ¿Qué decimos acerca del modelo propuesto si no podemos rechazar la hipótesis nula? ¿Y si podemos rechazarla?
  b) Intuitivamente, ¿por qué un valor más alto del estadístico F de la página 24 nos hace sospechar que la hipótesis nula es falsa?
  c) Misma pregunta anterior, pero con relación al estadístico F de la página 26.
  d) En la página 31 del ejemplo empírico: ¿cuál es el mínimo nivel de significatividad que permite rechazar la hipótesis nula de que las horas al cuadrado no son significativas?
  e) Responder las últimas 2 preguntas de la página 36.
  f) Decir qué test aplicarías para evaluar si la siguiente afirmación es verdadera o falsa: "lo único que puede explicar el número de materias aprobadas son las horas de estudio".