

Examen #2

Métodos Cuantitativos Aplicados, CIDE A.C.

Instrucciones

El examen contiene una sección con 11 preguntas sobre regresión lineal (75 puntos en total). Tienes 90 minutos para contestarlo. Te recomiendo leer las preguntas con cuidado y tratar de no dejar respuestas en blanco. Si no tienes la respuesta final, siempre es mejor mostrar el procedimiento que tu crees que te llevaría a dicha respuesta.

Al terminar, deberás enviar el archivo con tus respuestas a mi dirección de correo electrónico (sergio.bejar@cide.edu). ¡Buena suerte!

Sección 1: Actitudes sobre Gasto de Gobierno en Bienes Públicos

Las preguntas en esta sección serán respondidas con la base de datos `gss.csv` disponible en la carpeta “Files” de la página Github de nuestra clase. Los datos vienen de la primera ola de la encuesta General Social Survey en los Estados Unidos. En ella, se tratan de investigar las actitudes de los ciudadanos con respecto al gasto nacional en bienes públicos.

Las variables de nuestro interés son las siguientes:

educ: educación, mide el número de años de escuela completados (min: 0; max: 20).

race: raza del individuo (1 = blanco, 2 = negro, 3 = otra).

age: edad (min: 18, max: 89).

polviews: ideología (min: 1-extremadamente liberal-, max: 7-extremadamente conservador-).

sumnatsoc: suma de opiniones respecto a gasto en varios bienes públicos (i.e. `natenvir`, `natheal`, `natdrug`, `nateduc`, `natrace`, `natfare`, `natroad`, `natmass`, `natpark`, `natsoc`, `natchld`). Es una variable de intervalo con una distribución casi normal. Valores más altos indican que el individuo tiene opiniones más favorables respecto a incrementar el gasto de gobierno en bienes públicos.

1. Recodifica `race` para tener una dummy que tome el valor de 1 para individuos de raza negra y 0 para otras razas. Denota esta variable como `raza_d`. (3 pts.)
2. Estima un modelo OLS que tenga la siguiente especificación: (5 pts.)

$$sumnatsoc = f(educacin, raza_d, edad, ideologia)$$

3. Presenta los resultados en una tabla bien legible (i.e. no usando `summary()` sino un paquete que nos permita presentar mejor los resultados. (5 pts.)
4. Explica los resultados poniendo énfasis especial en los coeficientes de las variables de educación, raza (la dummy que creaste), e ideología. (10 pts.)
5. Estima las pruebas de Durbin-Watson y Breush-Godfrey y explica detalladamente por que sí o por que no existe un problema de autocorrelación en los datos. (10 pts.)
6. Checa la linealidad del modelo que estimaste en (2) y determina hay problemas de linealidad. Explica detalladamente tu respuesta. (7 pts.)
7. Elabora dos pruebas de diagnóstico diferentes para determinar si existe problema de heteroskedasticidad en la regresión que estimaste. Explica cuidadosamente tu respuesta. (10 pts.)
8. Estima de nuevo el modelo pero ahora con errores estándar robustos (función `coeftest` de paquete “sandwich”). ¿Qué diferencias encuentras en los resultados? (10 pts.)
9. El profe piensa que existe la posibilidad de que el efecto de ideología política sea condicional dependiendo de la raza de los individuos. Estima un modelo que evalúe la hipótesis del profe. (5 pts.)
10. Elabora un gráfico que muestre el efecto condicional de ideología política en la variable dependiente. (5 pts.)
11. Explica detalladamente los resultados del gráfico que elaboraste en el inciso anterior. (5 pts.)