

## Regresion y Prediccion

### Test de Lectura

Como discutiesemos, estos *tests de lectura* son una serie de preguntas simples que permiten monitorear si estan siguiendo efectivamente las clases. No seran corregidos ni evaluados, los administraran directamente ustedes. Cualquier duda que tengan sera discutida en nuestras sesiones sincronicas, la semana siguiente a publicado el material, luego de que hayan hecho un esfuerzo en trabajarlo, no antes. Es muy importante que trabajen este material y lo discutan entre ustedes.

1. Revisa los supuestos clasicos y el Teorema de Gauss Markov, de lo que hayas estudiado en una clase de econometria.
2. Verdadero o Falso (V o F): el estimador de MCO maximiza  $R^2$  porque la suma de cuadrados explicados es maxima.
3. V o F: un estimador de MCO que usa menos observaciones que las originales, es mas ineficiente que si las usa todas.
4. Cual es la principal diferencia entre estimar y predecir?
5. V o F: El predictor  $\hat{Y}_i$  discutido es eficiente porque es una combinacion lineal de un estimador lineal, insesgado y de varianza minima.
6. V o F: intuitivamente, el error cuadratico medio (ECM) de un estimador mide cuan lejos esta el estimador de su valor esperado.
7. V o F: si un estimador es insesgado, entonces el ECM es igual a la varianza.
8. Trabaja los detalles de la demostracion de la descomposicion sesgo/varianza del ECM.
9. Que diferencia hay entre el ECM y el error de pronostico?
10. Por que decimos que si  $Y = f(X) + u$ ,  $f(X)$  es el mejor predictor de  $Y$ ?
11. Revisa los detalles de la demostracion de que  $E(Y)$  es el mejor predictor de  $Y$ .
12. V o F: minimizar el error de pronostico equivale a minimizar el ECM
13. Revisa la demostracion de la descomposicion del error de pronostico en ECM mas  $\sigma^2$ . Ojo, aca hay un paso 'polemico'.

14. V o F: Si la prediccion se basa en un estimador insesgado de  $f$ , entonces minimizar el error de pronostico es identico a minimizar la varianza de  $f$ .
15. Por que en econometria no aparece el ECM?
16. Desde la perspectiva de un curso de econometria, por que no hay muestra de entrenamiento y de evaluacion?
17. Si se trata de minimizar el error de pronostico, porque es lo mismo mirar su promedio o su raiz cuadrada?
18. Googlea el desafio Netflix del millon de dolares, y lee el caso de como implementaron la muestra de entrenamiento y de evaluacion.
19. Este es un ejercicio interesante. Considera un modelo simple  $y_i = x_i\beta + u_i$ , es decir, un modelo en donde hay una sola variable explicativa (sin intercepto) y que ademas las variables han sido normalizadas (tienen promedio cero y varianza uno). Valen todos los supuestos clasicos, y  $x_i$  es una variable no aleatoria. Considera el siguiente estimador:

$$\hat{\beta}_r = \frac{\hat{\beta}}{1 + \lambda},$$

en donde  $\hat{\beta}$  es el estimador de MCO y  $\lambda$  es un escalar arbitrario mayor o igual a cero. Este estimador se llama ‘ridge’ y cumplira un rol crucial mas adelante.

- a) Obtene el sesgo y la varianza de este estimador, para un valor fijo y arbitrario de  $\lambda$ . Comparalos con el sesgo y la varianza de estimador MCO  $\hat{\beta}$ .
- b) Calcula el ECM para el estimador  $\hat{\beta}_r$  y para  $\hat{\beta}$ .
- c) Muestra que existe un valor de  $\lambda$  para el cual  $\hat{\beta}_r$  tiene menor ECM que  $\hat{\beta}$ .
- d) Que conclusion sacas de este ejercicio, en relacion a la practica estandar de la econometria de prefirir tercamente estimadores insesgados?