# Examen Estadística Aplicada III

Octubre 3, 2019.

**Instrucciones (leerlas con calma)**

En el archivo anexo en Excel y llamado “**BREXIT**” los datos relativos a los resultados de una serie de variables que se relacionan con tal ejercicio democrático en cada una de los *regiones electorales* del Reino Unido. En caso de que requieran más detalles pueden consultar la siguiente dirección: <https://www.bbc.com/news/uk-politics-32810887>

1. La primera actividad es la relativa a obtener una muestra de una proporción del 85% de toda la base usando como *semilla* su propio número de estudiante y guardarla en un documento (recomiendo que sea en Excel) y lo llamen “mi\_muestra\*\*\*\*\*\*\*” Los signos \*\*\*\*\*\*\* representan su número de estudiante. Este archivo es único y es a partir de los datos contenidos en esta matriz que deben responder las preguntas que se exponen más adelante.

Una vez que tienen su propia muestra, responder las preguntas enunciadas en la **siguiente sección**. Reportar tales respuestas en un documento aparte (en Word) que llamarán “RESPUESTAS\*\*\*\*\*\*\*”, en donde “\*\*\*\*\*\*\* “ es, una vez más, su número de estudiante en el ITAM. En este documento deberán incluir cada una de los argumentos para responder cada una de las preguntas. NO incluir gráficos que no se comenten o que no aporten elementos importantes. El documento conteniendo las respuestas deberán enviármelo a la siguiente dirección electrónica:

**hernande@itam.mx**

la hora límite para realizar el envío es 12h25 pues hay clase a las 11h30h. NO envíen la base de datos.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**PREGUNTAS DEL EXAMEN**

1. El problema principal es el tratar de explicar la decisión tomada en cada región electoral: ¿hay alguna relación entre tal variable y el resto de las variables? La respuesta debe incluir valores de algunos estadísticos así como gráficos que ustedes consideren pertinentes. Notar que en este caso se trata de un análisis exploratorio justificando sus respuestas. La primera actividad a realizar debe ser la creación de una nueva variable que se llame RESULTADO con dos valores posibles: 1 si la proporción de votos por la opción Remain es mayor que 50% y será igual a 0 si dicho porcentaje es menor o igual que dicho 50%. Realizar un análisis exploratorio en los límites que ustedes consideren pertinente para que con los datos de la base se pueda realizar otro tipo de análisis. Redactar un resumen de este análisis exploratorio de los datos de su muestra y teniendo en cuanto la pregunta ¿Por qué el resultado final no es siempre el mismo (en cada valor observado en la base de la muestra)? ¿Qué posibles factores (variables) podrían estar asociados a la respuesta (leave o remain).
2. A partir de la variable dicotómica REMAIN definida en el punto anterior, construir un modelo de regresión logística que permita tener el mejor porcentaje de aciertos con los datos de la muestra. Justificar la elección de las variables incluidas en el modelo. Por supuesto, pueden crear nuevas variables a partir de la información disponible en el archivo. Realizar una tabla de dos por dos en la que se puedan ver las proporciones de los pronósticos obtenidos con su modelo y usando los datos de la muestra.
3. Usar el criterio encontrado en el punto anterior y aplicarlo en el conjunto de los valores que no fueron usados para el modelo descrito en el punto 2.
4. Escribir un breve ensayo acerca de este problema real con sus principales resultados. Por ejemplo, ¿Es posible pensar que hubo regiones más favorables al REMAIN? ¿Cómo se caracterizan? Si se tuviese que realizar un BREXIT2 ¿Dónde se debería hacer más campaña en favor del BREXIT? (nombrar las regiones o las áreas).
5. Si así lo desean pueden agregar comentarios adicionales de este problema y el propuesto