

## IN4402-1 Aplicaciones de Probabilidades y Estadística en Gestión

Profesor: Andrés Fernández Vergara

Equipo docente: Catalina Hurtado, Daniel Monsalve, Angelo Muñoz

Semestre: Primavera 2020

### TAREA 1

En esta tarea se le pide a cada grupo reportar un informe de investigación, lo que involucrará: revisar literatura, plantear una hipótesis, explorar los datos, analizar los datos, testear la hipótesis y concluir al respecto.

Entrega: **jueves 15 de octubre a las 18:00 hrs.**

(Entregas posteriores verán reducida su nota en 10 décimas por día de atraso)

Auxiliar encargado/a: **Catalina Hurtado.**

El consumo de carne a nivel global se ha vuelto un tema preocupante para muchos países, principalmente debido a las implicancias ambientales que su producción significa, y las condiciones sanitarias que produce en la gente. Para esta tarea se les pide estudiar algunos factores que afectan la producción de carne per cápita para diferentes países alrededor del mundo. Para ello, deberán recolectar datos<sup>1</sup> sobre las siguientes variables: Consumo de carne per cápita por país, Producto Interno Bruto (PIB) per cápita, Población, Índice de Desarrollo Humano (IDH), Edad de la población, Religión predominante o distribución de población según culto religioso, Tasa de urbanización, u otras que ustedes consideren relevantes.

Esta es una tarea “no estructurada”, por lo tanto no existe una única forma de llevarla a cabo. Se evaluará la coherencia del proceso y del análisis. Para llevar a cabo la investigación, Ud. deberá hacer lo siguiente:

1) **Elaborar una hipótesis** sobre la relación de un factor (por ejemplo, religión predominante, o el nivel de desarrollo humano) con el nivel de producción de carne. Ustedes deben justificar su hipótesis utilizando al menos dos referencias de estudios empíricos nacionales o internacionales que sean relevantes a su hipótesis. Para esto se recomienda utilizar el buscador Google Scholar o Web of Science. [10 puntos]

2) **Construir una única base de datos** que permita examinar la hipótesis planteada. Expliquen brevemente el proceso de construcción de la base de datos a utilizar, e identifiquen claramente la unidad de análisis. Noten que algunas variables ustedes deberán crearlas o transformarlas para poder estudiar la hipótesis planteada (por ejemplo, usar logaritmo, comparar proporciones, etc.). Además, noten que pueden existir algunos datos faltantes (no borren los datos, utilícenlos adecuadamente). Argumente las decisiones y supuestos realizados en este proceso. [20 puntos]

3) **Realizar un análisis exploratorio de datos (EDA)**, que incluya todos los puntos vistos en clases. Analicen gráficamente el comportamiento de la variable dependiente y del factor de interés escogido para su hipótesis. Estudie también esa relación diferenciando por algún otro factor que sea relevante para el análisis (por ejemplo, si esa relación cambia en países de menor o mayor PIB). Utilice los conceptos de visualización vistos en clases para apoyar el análisis. Este análisis descriptivo debe apoyar coherentemente los modelos de análisis que ustedes realizarán posteriormente para testear sus hipótesis. [25 puntos]

---

<sup>1</sup> Los sitios web donde pueden encontrarse estos datos se presentan al final de este enunciado.

4) **Utilizar regresiones lineales que permitan comprobar su hipótesis** sobre factores que explican la diferencia en la producción de carne entre países. Para ello, presenten distintos modelos agregando variables explicativas en forma ordenada y justificada<sup>2</sup>, explicando qué sucede con la relación estudiada a medida que agrega o quita cada variable. Incorporen también una variable de interacción que guarde relación con los gráficos presentados en la parte 3) sobre el EDA. Es importante ir entendiendo los resultados que va entregando cada modelo (esto es, no ir describiendo mecánicamente cada resultado, sino leerlos a la luz de la hipótesis planteada). Además, deben hacerse cargo de los supuestos estadísticos (por ejemplo, homoscedasticidad) para el modelo que ustedes consideren más relevante. [25 puntos]

5) **Discutir y concluir basándose en la evidencia encontrada**, incluyendo cualquier supuesto que ustedes hayan utilizado para llevar a cabo su análisis. La conclusión debe estar vinculada a su hipótesis inicial. En la discusión, brevemente discutan cada tipo de validez (interna, externa y de constructo) vista en el curso. [15 puntos]

Después de subir su tarea al módulo de tareas en U-Cursos, y a más tardar **el 16 de octubre de 2020**, cada estudiante debe subir a [este link](#) su principal conclusión respecto a la tarea [5 puntos].

#### **Instrucciones para el informe:**

Esta tarea se realizará en los grupos ya inscritos de 4 personas. Es muy importante que todas/os se involucren en el proceso de aprendizaje completo de la tarea. De haber problemas de contribución en un grupo por favor comuníquese con la auxiliar encargada de la tarea.

Formato: Máximo 6 páginas, excluyendo portada, referencias y anexos. Letra Times New Roman número 12, interlineado 1,15 o 1,5; márgenes normales de página y tamaño carta.

Se recomienda **resumir la información**: mostrar sólo gráficos relevantes y tablas resumen de los resultados. Los anexos no se leerán a menos que estén explícitamente referenciados dentro del informe (esto es, si una tabla es enviada a anexos, en el párrafo del informe debe decirse “Se realizó tal análisis, cuyo resultado es A. La tabla que muestra este resultado se puede ver en el Anexo X”, o algo por el estilo).

La tarea tiene 100 puntos (60 pts = 4.0).

#### **Algunos aspectos fundamentales a considerar:**

- Las hipótesis no son objetivos – se puede encontrar evidencia a favor o en contra de ésta. Procure que su hipótesis no predetermine su conclusión.
- Es mejor usar supuestos razonables y ser coherente con un análisis “enfocado” que explorar muchas cosas sin robustez (“quien mucho abarca poco aprieta”).
- Sea cuidadosa/o con temas de formato, incluyendo ortografía y gramática, lo cual también será considerado en la evaluación.
- Se aconseja fuertemente NO dejar las tareas para último minuto. Por favor NO solicitar extensión en los tiempos de entrega.

---

<sup>2</sup> Considere los ejemplos vistos en clases donde se analizan varios modelos; por ejemplo, en la Clase 4 (diapositiva 20) sobre la brecha de género en el salario; o el ejemplo de la Clase 5 (diapositiva 13) sobre el salario según escolaridad.

### Enlaces para la búsqueda de datos:

- En este [enlace](#) encontrarán información sobre la producción de carne en los países.
- En este [enlace](#) encontrarán información sobre el Producto Interno Bruto (PIB) de los países. También se recomienda leer (puede ser [aquí](#)) sobre la definición de PIB para entenderlo a cabalidad. Para obtener el PIB per cápita se puede dividir por el total de la población de cada país.
- En este [enlace](#) encontrarán información sobre la población de los países. Utilicen la información del año más reciente o más cercano al año en que esté la información de la variable dependiente.
- En este [enlace](#) encontrarán información sobre la edad de la población de cada país, particularmente la mediana.
- En este [enlace](#) encontrarán información sobre la religión profesada en cada país. Se recomienda dividir la participación según la población de cada país, utilizando una medida fácil de interpretar (por ejemplo, 1 cristiano por cada 100.000 personas en un país). Para los valores menores a 100.000 puede utilizar el mínimo 1.
- En este [enlace](#) encontrarán información sobre la tasa de urbanización de los países.