

Especificaciones del proyecto del curso MA-3405

Estadística

I semestre del 2021

Profesores: Giovanni Sanabria Brenes, Bryan Rodríguez Castro, Juan Pablo Prendas Rojas, Emanuelle Soto Cascante, Esteban Ballesterio Alfaro.

Estimado estudiante, el presente documento es una guía para la elaboración del proyecto del curso que tiene **un valor del 50%** en la nota final. La guía incluye una descripción de cada una de las partes a realizar y la fecha de entrega. El trabajo se realizará en grupos **máximo de 4 personas**.

I parte: Planteamiento del problema

Se debe definir el problema a tratar, los parámetros, variables y los objetivos buscados. Debe hacer una breve introducción que justifique la elección del problema.

Deben estar involucrados dos poblaciones independientes A y B. Se analizarán los parámetros: media, proporción y variancia y su relación de estos parámetros entre las poblaciones.

Debe haber al menos 2 variables cuantitativas y 2 cualitativas ligadas a una misma población

Este planteamiento debe contar con el aval del profesor para continuar. Tome en cuenta que, en la siguiente parte del proyecto, se solicitará obtener datos de al menos 1000 individuos ligados a esas poblaciones.

Ejemplo:

- 1) **Problema:** Resultados de próximas elecciones (AQUÍ SE DEBE EXPLICAR A GRANDES RAZGOS EL PROBLEMA Y DEFINIRLO EN TIEMPO Y ESPACIO, ADEMÁS DEBE INDICAR LOS OBJETIVOS DEL ESTUDIO)

Parámetros:

	Población 1: ciudadanos de Cartago	Población 2: ciudadanos de San José
Edad promedio de los que no votarán	μ_1	μ_2
% que votarán por el candidato X	p_1	p_2
Variancia de la edad de los que no votarán	σ_1	σ_2

Variables cuantitativas:

- Edad de los ciudadanos de Cartago que no votarán. (ligada a los parámetros de promedio y variancia de la población 1)

- Edad de los ciudadanos de San José que no votarán. (ligada a los parámetros de promedio y variancia de la población 2)
- Ingreso de los ciudadanos de Cartago que no votarán (necesario para que se tengan dos variables cuantitativas asociadas a una misma población)

Variables cualitativas:

- Candidato por el que votará el ciudadano de Cartago.
- Cantón en el que votará el ciudadano de Cartago

Variables binarias:

- Si el ciudadano de Cartago vota o no por el candidato X (ligada a la proporción de la población 1)
- Si el ciudadano de San José vota o no por el candidato X (ligada a la proporción de la población 1)

No puede iniciar el trabajo hasta que el profesor le apruebe el problema a investigar.

II parte: Recolección de Datos

- Elaborar el instrumento para recolectar la información (encuesta, entrevista,...)
- Debe recopilar datos de al menos 500 individuos para las variables identificadas
- Muestreo. Aplicar el instrumento. Debe indicar la fecha, lugar y hora a la que aplicó el instrumento, si fue aleatorio en la escogencia de los individuos y si utilizó cuotas para ciertos subgrupos de individuos (ejemplo: 50 hombres y 30 mujeres). Investigar sobre los tipos de muestreo
- Tabulación de Datos en Excel

III parte: Análisis Inferencial (IC)

Realice intervalos de confianza que involucren una y dos poblaciones (Al menos un IC de: un promedio, una proporción, una variancia, diferencia de promedios, diferencia de proporciones y cociente de variancias)

Debe indicar las conclusiones obtenidas y también las hipótesis asumidas (por ejemplo, la población X se asume que sigue una distribución normal)

IV parte: Análisis Inferencial (pruebas de hipótesis de una y dos poblaciones)

Realice pruebas de hipótesis que involucren una y dos poblaciones (Al menos una prueba de hipótesis de: un promedio, una proporción, una variancia, diferencia de promedios, diferencia de proporciones y cociente de variancias)

Debe indicar las conclusiones obtenidas y también las hipótesis asumidas (por ejemplo, la población X se asume que sigue una distribución normal)

V parte: Análisis Inferencial (Otras pruebas de hipótesis)

1. Realice una Bondad de Ajuste para determinar si una de las variables cuantitativa sigue o no una distribución normal.
2. Realice una Bondad de Ajuste para determinar si una de las variables cuantitativa sigue o no una distribución no normal (escoja una que no sea normal).
3. Realice una prueba de independencia entre las variables cualitativas de una misma población.
4. Realice un ANOVA entre una variable cuantitativa y otra cualitativa de una misma población.
5. Debe indicar las conclusiones obtenidas

VI parte: Análisis de regresión

1. Realice un análisis de regresión lineal entre las variables cuantitativas de una misma población. (estimación puntual, estimación por intervalo de los parámetros, intervalo de predicción, coeficiente de correlación)
2. Escoja una regresión NO lineal (de las vistas) que se ajuste más para las variables cuantitativas de una misma población. Justifique su escogencia.
3. Realice un análisis de regresión no lineal escogida entre las variables cuantitativas de una misma población. (estimación puntual, estimación por intervalo de los parámetros, intervalo de predicción, coeficiente de correlación)

V parte: Fases del proyecto y detalles de entrega

	Fase 1	Fase 2	Fase 3
Valor total	15%	15%	20%
Entrega de informe detallado del análisis realizado que involucra las partes:	De la I a la III Valor 12%	IV Valor 12%	V y IV Valor 16%
Fecha de entrega del informe:	A más tardar el 14 de octubre	A más tardar el 18 de noviembre.	A más tardar el 9 de diciembre
Exposición del trabajo realizado	Valor 3%	Valor 3%	Valor 4%

(ver observaciones en la siguiente página)

Observaciones:

- Los temas definidos en la primera fase junto con los datos obtenidos se mantienen para las siguientes fases. Sin embargo, cambian los integrantes de cada grupo. Por lo tanto, se pretende en cada fase analice un tema diferente.
- Los grupos, en cada fase, serán conformados por el profesor.
- Sobre la nota del informe en cada fase.
 - El profesor dará una nota inicial del proyecto.
 - Cada estudiante deberá valorar con un porcentaje el trabajo de cada uno de sus compañeros y su propio trabajo.
 - Porcentaje promedio: es el promedio de los porcentajes obtenidos para cada estudiante.
 - La nota de cada estudiante se obtiene sacando el porcentaje promedio obtenido de la nota inicial.
 - Ejemplo: suponga que la fase I el grupo 1 estuvo conformado por los estudiantes A,B, C y D:

	Nota inicial del informe da por el profesor	Valoración Del trabajo realizado dada por				Promedio de valoraciones	Nota obtenida del informe de la fase I
		A	B	C	D		
A	90	100	100	100	100	100	90
B	90	70	70	70	90	75	67,5
C	90	10	60	50	40	40	36
D	90	100	80	90	100	92,5	83,25

- Sobre la exposición.
 - Las exposiciones se realizan en el horario de clases posterior a la entrega del informe. La exposición de la fase 1 y 2 se realizará en el período de clases asincrónico, para no afectar las clases. La exposición de la fase 3 se realizará en la última clase de la semana 16.
 - Deben estar todos los integrantes del grupo, y pueden preparar una presentación no mayor a los 5 minutos.
 - El profesor del curso elegirá uno de los miembros grupo para que realice la exposición.
 - Finalizada la presentación realizará un par de preguntas teóricas sobre el trabajo realizado a alguno de los integrantes.
 - Desglose de la nota (será la misma para todos los integrantes del grupo):
 - Presentación y manejo del tiempo: 40%
 - Respuesta a la pregunta 1: 30%
 - Respuesta a la pregunta 2: 30%