

# Estadística Aplicada 1

Rodrigo Zepeda

20 de enero de 2020

## 1. Temario

### 1. INTRODUCCIÓN

- a)* Definición de Estadística y conceptos fundamentales
- b)* Las ramas de la Estadística
- c)* Los elementos de un Problema Estadístico
- d)* La calidad de los Datos
- e)* Censo vs. Muestra

### 2. ANÁLISIS EXPLORATORIO DE DATOS (AED)

- a)* Objetivos del AED
- b)* Tipos de Variables y Escalas de Medición
- c)* Análisis Exploratorio de Datos Cualitativos
- d)* Análisis Exploratorio de Datos Cuantitativos
- e)* El Problema de Comparación
- f)* El Problema de Asociación
- g)* Paquetes Computacionales para AED

### 3. INTRODUCCIÓN AL MUESTREO

- a)* Introducción
- b)* Poblaciones Finitas vs. Poblaciones Infinitas
- c)* Ventajas del Muestreo
- d)* Las 4 preguntas Fundamentales del Muestreo
- e)* Definición de Diseño Muestral
- f)* Muestreo Probabilístico vs. Muestreo No Probabilístico
- g)* Estimación de los parámetros más usuales

### 4. MUESTREO ALEATORIO SIMPLE (MAS)

- a) Definición y Notación. Población simple
  - b) Estimación de Medidas y Totales
  - c) Estimación de Proporciones y Totales de Categoría
  - d) Estimación de la Varianza
  - e) Estimación por Intervalos
  - f) Tamaños de Muestra
  - g) Ventajas y Desventajas del muestreo Aleatorio Simple
5. MUESTREO ALEATORIO ESTRATIFICADO (MAE)
- a) Definición y Notación. Población estratificada
  - b) Estimación de Medidas y Totales
  - c) Estimación de Proporciones y Totales de Categoría
  - d) Estimación de la Varianza
  - e) Estimación por Intervalos
  - f) Tamaños de Muestra
  - g) Asignación de la Muestra a los Distintos Estratos
  - h) Ventajas y Desventajas del Muestreo Aleatorio Estratificado
6. MUESTREO ALEATORIO BIETÁPICO (MAB)
- a) Definición y Notación. Población en etapas
  - b) Estimación de Medidas y Totales
  - c) Estimación de Proporciones y Totales de Categoría
  - d) Estimación de la Varianza
  - e) Estimación por Intervalos
  - f) Tamaños de Muestra
  - g) Casos Especiales: Muestreo Aleatorio Bietápico por Conglomerados y Muestreo Aleatorio Sistemático
  - h) Comparación de MAS, MAE, MAB
  - i) Muestreo Multietápico
  - j) Ventajas y Desventajas del Muestreo Aleatorio Bietápico
7. ESTIMADORES DE RAZÓN Y DE REGRESIÓN
- a) El Estimador de Razón
  - b) El Estimador de Regresión

## 2. Evaluación

Para la evaluación de la materia se realizarán dos exámenes parciales y un examen final así como tres proyectos y tres tareas. Toma en cuenta las siguientes observaciones

1. Para acreditar la materia es necesario acreditar tanto el proyecto final como el examen final.
2. Para que un proyecto se considere válido y se califique debe presentarse en la fecha determinada por el calendario la *presentación del proyecto* así como la tarea correspondiente.
3. La presentación del proyecto dura dos minutos. En él se tiene que presentar:
  - a) Una breve descripción de los datos (responder a las preguntas ¿qué se midió? ¿cómo se midió? y ¿cuál es la fuente?)
  - b) Una breve evaluación de la calidad de los datos.
  - c) Una gráfica ó tabla que describa brevemente el tema a analizar.

Ejemplo: Mis datos son de la Encuesta Nacional de Adicciones realizada por la Secretaría de Salud y el tema que me interesa estudiar es el consumo de marihuana. Se midió el consumo preguntándole a las personas .<sup>En</sup> las últimas dos semanas ¿has consumido marihuana?<sup>en</sup> caso de que respondieran que sí, "¿cuántos gramos?". La siguiente tabla muestra la media, varianza, mediana y cuartiles de los gramos reportados.

4. Para los proyectos los equipos son de una, dos ó tres personas.
5. Las tareas son individuales a menos que se indique lo contrario. Tareas o proyectos con similitudes mayores al 25 % se considerarán plagio lo que conlleva desde la pérdida de puntos hasta la pérdida de la materia con reporte a la coordinación.

La puntuación final está dada por la tabla 1.

### 2.1. Criterios específicos de los exámenes

Los exámenes pueden tener preguntas de teoría (ej. *Explica qué es una población objetivo*) o de práctica matemática (ej. *determina si el estimador de  $\bar{x}$  es insesgado*).

Las preguntas de teoría se evalúan de la siguiente manera:

1. (25 %) Comprensión básica del concepto.
2. (25 %) Precisión y exactitud. La definición del concepto no debe contener un concepto distinto (no se puede dar una definición tan vaga que incluya otras cosas).

Entregable	Porcentaje (%)	Fecha
Proyecto 1 (presentación de objetivos)	REQUISITO	28 de enero
Tarea 1	5	14 de febrero
Proyecto 1	10	25 de febrero
Parcial 1	15	3 de marzo
Proyecto 2 (presentación de objetivos)	REQUISITO	10 de marzo
Tarea 2	5	20 de marzo
Proyecto 2	10	14 de abril
Parcial 2	15	21 de abril
Proyecto Final (presentación de objetivos)	REQUISITO	23 de abril
Tarea Final	5	1 de mayo
Proyecto Final	15	19 de mayo
Examen Final	20	Por definir

Cuadro 1: Forma de evaluación

3. (25 %) Ejemplo claro del concepto.
4. (25 %) Claridad de la redacción.

Las preguntas de práctica matemática se evalúan de la siguiente manera:

1. (25 %) Comprensión básica del concepto y su notación.
2. (25 %) Claridad en cuanto a cómo resolver el ejercicio.
3. (25 %) Ningún error de signo / conteo / suma (sólo aplicable si se sabe cómo resolver el ejercicio).
4. (25 %) Solución completa del ejercicio.

Algunos errores numérico o algebraicos que resultan en ejercicios de menor dificultad son penalizados con hasta el 90 % del puntaje total (por terminar haciendo un ejercicio más simple). Mientras que errores que resultan en ejercicios de mayor dificultad y son resueltos correctamente pueden, inclusive, obtener el 100 % de los puntos<sup>1</sup>.

## 2.2. Criterios específicos de los proyectos

### 2.2.1. Proyecto 1

El primer proyecto consiste en realizar una visualización y una tabla de una base de datos elegida por el equipo de alumnxs y aprobada por el instructor *sin tomar en cuenta el diseño de encuestas*. Dicha visualización deberá ser elaborada en **R** ó **Python**. Para la visualización de los datos ésta debe presentarse en a lo más 10 minutos estableciendo:

<sup>1</sup>Por ejemplo si al copiar los ejercicios copiaste un 5 donde era 7 pero todo lo demás fue correcto y con la misma dificultad.

1. El objetivo del análisis que se presenta.
2. La fuente y confiabilidad de los datos evaluada como en la Tarea 1.
3. Tantas visualizaciones de datos y tablas como se consideren necesarias para respaldar las conclusiones que se presentan. Es completamente válido concluir que "con estos datos no se puede concluir" siempre y cuando esto esté respaldado.
4. Una conclusión.
5. Al profesor deberá entregársele en USB o por correo:
  - a) Archivo .R con el código de R versión 3.6.2 que pueda correr en mi máquina. EL código deberá tener un comentario de qué hace cada enunciado. Así mismo, deberá seguir las guías de formato de Google disponibles en: <https://google.github.io/styleguide/Rguide.html>. Código que no siga esas guías, no corra o no tenga comentarios resulta en la pérdida de 2 puntos en el proyecto.
  - b) Base de datos en formato .RDS con la información analizada.

Se pueden utilizar datos de sus trabajos siempre y cuándo estos puedan ser compartidos con el instructor.

6. En caso de no presentarse todos los miembros del equipo el día asignado de la presentación el equipo deberá enviar:
  - a) Un documento pdf con toda la información previa y excelente redacción (se resta un punto completo por errores gramaticales, ortográficos o de la manera de citar las referencias). Dicho documento debe entregarse a lo más dos días después de la fecha de la presentación en pdf a mi correo.
  - b) Los miembros del equipo que sí asistan a la presentación deberán presentar todo y responder todas las preguntas (no vale decir "eso le tocó al que no vino").
  - c) Deberá, así mismo, realizarse un video de duración máxima de 10 minutos donde se realice la presentación que iba a ocurrir durante la clase. Dicho video debe ser entregado a lo más dos días después de la fecha de la presentación a mi correo.

### **2.2.2. Proyecto 2**

SE ANUNCIARÁ

### **2.2.3. Proyecto Final**

SE ANUNCIARÁ

### 3. Reglamento de clase

Durante la clase queda terminantemente prohibido realizar comentarios xenófobos, misóginos, albueros, racistas o de cualquier otra naturaleza que pueda ofender, humillar o faltar al respeto de alguna manera a cualquiera de los asistentes. Las sanciones van desde salir del salón hasta perder la materia y tener un reporte ante las autoridades del ITAM.

El profesor también es sujeto de estas sanciones y puedes reportarme con mi coordinador `ruben.hernandez.cid@itam.mx`

### 4. Código de Honor

Al tomar esta materia te comprometes a que todo el trabajo entregado en clase (como pueden ser tareas, exámenes, etcétera) es de tu autoría. Trabajo cuya fuente primaria no sólo seas tú (por ejemplo, de otra persona, de un equipo del cual eres miembro o de un programa computacional) debe estar claramente indicado como tal.

Te recuerdo que el reglamento del ITAM establece sanciones pesadas en caso de deshonestidad académica ¡no te arriesgues!

### 5. Bibliografía

El libro principal que estaremos usando es *Sampling: Design and Analysis* de Lohr (2a edición). En algunas ocasiones nos apoyaremos de *Model Assisted Survey Sampling* de Sarndal, de *Statistical Design* de Casella. En caso de estar usando otra referencia para algún tema en específico el profesor lo indicará por correo.

### 6. Recursos adicionales

Hay dos páginas que necesitas conocer para apoyarte en Aplicada I. La primera es Wolfram Alpha: éste es un programa que te resuelve casi todos los ejercicios *de talacha* de la materia. Para la talacha es la cosa más útil, no lo dudes; empero, no todos los ejercicios son de talacha (y algunos Wolfram no los sabe hacer o no da la respuesta correcta como veremos) entonces no te confíes. Wolfram es genial para checar respuestas: úsalo para eso.

Las otras páginas que necesitas conocer son Math Stack Exchange y Cross Validated. Estos son sitios donde puedes hacer preguntas de matemáticas y la gente te contesta ¡de manera gratuita! (De hecho si sabes buscarle, varios problemas de los libros de texto ya están ahí). Éste es un muy buen recurso para preguntar dudas / pedir apoyo en las tareas. ¡No lo desprecies!

Éstas son las mejores formas de hacer trampa (o pedir ayuda, depende cómo lo hagas). Úsalas bajo tu propio riesgo.