

SOL114 ANÁLISIS DE DATOS II

Segundo Semestre 2024

Instructor:	Mauricio Bucca	Horario Clase:	LU/MI 9:40 - 10:50
Email:	mebucca@uc.cl	Horario Consulta:	MI 14:30 - 15:30

Ayudantes:

- Elisa Tagle (elisa.tagle@uc.cl), Martes 14:50 - 17:20
- Benjamín Puran (bpuran@uc.cl), Martes 14:50 - 17:20

Páginas del curso: <https://github.com/mebucca/ad2-sol114>

Descripción: Este curso entrega las bases matemáticas y conceptuales de teoría de la probabilidad e inferencia estadística, esenciales para el análisis cuantitativo en ciencias sociales. La primera mitad del curso cubrirá la teoría básica de la probabilidad, que sirve como fundamento de la teoría estadística. La segunda mitad se dedicará a temas de inferencia estadística, incluyendo estimación puntual y de intervalos, prueba de hipótesis y asociación estadística.

Resultados esperados: Al completar este curso satisfactoriamente se espera que los estudiantes sean capaces de:

- Entender los fundamentos conceptuales y matemáticos de teoría de la probabilidad e inferencia estadística
- Usar datos y herramientas estadísticas para estimar cantidades de interés y cuantificar la incertidumbre en torno a las estimaciones
- Usar datos y herramientas estadísticas para testear hipótesis
- Usar datos y herramientas estadísticas para medir asociación entre variables
- Implementar estas herramientas en software estadístico
- Aplicar los conocimientos descritos para responder preguntas empíricas de relevancia para las ciencias sociales

Prerequisitos sugeridos:

- Un curso de introducción a la estadística de nivel pregrado.

Estrategias metodológicas: En el curso los alumnos serán evaluados por su capacidad de comprender sustantivamente, resolver matemáticamente e implementar en software estadístico problemas relacionados con los contenidos del curso. En paralelo el profesor usará métodos de simulación para ilustrar resultados estadísticos teóricos (ej., teorema del límite central). La aplicación empírica de los contenidos teóricos se cubrirá con tareas que involucren datos relevantes para las ciencias sociales y uso de software estadístico. El código que acompaña cada clase será puesto a disposición de los estudiantes. Este curso hará uso intensivo de formalización matemática y, de manera expositiva, conceptos básicos de cálculo. Si bien no es un requisito formal, se sugiere familiaridad con contenidos de pre-cálculo y cálculo en una variable.

Contenidos:

- Experimentos aleatorios, eventos, probabilidades
- Probabilidad condicional, teorema de Bayes, Independencia
- Variables aleatorias y distribuciones de probabilidad
- Distribuciones discretas: Bernoulli y Binomial
- Distribuciones continuas: Uniforme y Normal
- Momentos: valor esperado & varianza
- Estandarización: puntaje Z
- Distribución Normal Standard
- Ley de los grandes números
- Teorema del límite central
- Estimación puntual
- Distribución muestral de un estimador
- Distribución muestral via Experimento Monte Carlo
- Estimación de intervalos & intervalos de confianza
- IC: ejemplos tradeoff entre precisión y confianza
- IC para media y proporción muestral
- IC con varianza desconocida & margen de error
- Estimación de IC via método de Bootstrap
- Test de hipótesis: paradigma Fisher & Neyman-Pearson

- Test de proporción en una muestra
- Test de media en una muestra
- Test de media/proporción en dos muestras
- Test de hipótesis via Test de Permutaciones
- Asociación estadística y causalidad
- Asociación con datos categóricos
- Test de independencia χ^2
- Asociación con datos continuos & covarianza
- Correlación de Pearson

Evaluaciones: Tareas cortas (30%), Pruebas (40%), Exámen (30%).

- Tareas cortas: Se realizarán 7 tareas cortas. Estas tareas medirán el conocimiento de conceptos y técnicas específicas de estadística, así como la capacidad de los estudiantes para implementarlos utilizando software estadístico. Los estudiantes podrán descartar la tarea con nota más baja, con lo cual la ponderación de las 6 tareas restantes será de 5% sobre la calificación final del curso.
- Pruebas: Se llevarán a cabo 2 pruebas de mediana extensión, cada una con una ponderación del 20%. Estas pruebas evaluarán el dominio matemático y conceptual de los contenidos y herramientas vistos en módulos específicos del curso.
- Examen Final: El examen final del curso tendrá una ponderación del 30%. Este examen será de carácter comprehensivo, abarcando todos los temas y contenidos estudiados. Los estudiantes deberán demostrar su comprensión, conocimiento y capacidad de aplicación de los conceptos y técnicas aprendidas.
- Prueba Bonus: Esta prueba es opcional y abarcará todos los temas y contenidos estudiados hasta el momento. Si la nota obtenida en la prueba bonus es mayor que la más baja entre las pruebas 1 y 2, entonces cada una (pruebas 1, 2 y bonus) tendrá una ponderación del 13.3% en la nota final. Si la nota de la prueba bonus es menor que la más baja entre las pruebas 1 y 2, dicha nota no se considerará en el cálculo de la nota final. Los estudiantes que, por una razón justificada, no hayan rendido una de las pruebas (1 o 2), deberán obligatoriamente rendir la prueba bonus.

Bibliografía mínima: El curso es auto-contenido, sin embargo el siguiente libro ofrece una cobertura profunda y accesible de los contenidos del curso y será utilizado en ayudantías.

- Imai, Kosuke. (2018). Quantitative Social Science: An Introduction in tidyverse. Princeton University Press.

Bibliografía complementaria:

- DeGroot, M. H. and Schervish, M. J. (2002). Probability and Statistics, Addison-Wesley.

Políticas del curso:

- **Asistencia:** La asistencia regular a clases es importante, esperada, pero no obligatoria.
- **Computadores en clases:** No.
- **Entregas:** Las tareas y trabajos deben ser entregados a mas tardar el día indicado en el calendario del curso, a la hora señalada.
- **Horario de consulta:** El profesor destinará una hora semanal para responder preguntas relacionadas con el curso. La participación es voluntaria.
- **Email:** Escribir primariamente para agendar reunión en el horario de consulta. Incluir SOL114 en el asunto del correo. Respondo correos una vez al día (en la mañana), de lunes a viernes.
- **Honestidad Académica:** El desconocimiento de la política de honestidad académica de la universidad - incluyendo el uso de IA - no es una explicación razonable para su incumplimiento.