



## DOCTORADO EN CIENCIAS AGRARIAS

NOMBRE DEL CURSO	
<b>MODELOS LINEALES: MÉTODOS Y APLICACIONES EN LAS CIENCIAS AGROPECUARIAS</b>	
<i>Prerrequisitos</i>	<i>Sin prerrequisitos</i>
<i>Creditaje</i>	<i>4 SCT-Chile</i>
<i>Horas presenciales y no presenciales</i>	<i>36 horas presenciales y 72 no presenciales</i>
<i>Horario</i>	<i>Jueves 09.00 a 13.00 horas</i>
<i>Profesor responsable</i>	<i>Dr. Francisco Zamudio</i>
<i>Profesores participantes</i>	
DESCRIPCIÓN DEL CURSO	
<p>El curso se orienta principalmente a revisar los procedimientos analíticos que permiten (a) enunciar modelos lineales (MLs) y (b) analizarlos adecuadamente. Estos procedimientos están basados principalmente en el álgebra de matrices y la teoría de probabilidades. Entonces, el enfoque está orientado a la construcción de modelos y el análisis de datos.</p> <p>El énfasis es más bien conceptual, concreto y de aspecto aplicado más que abstracto y teórico. Esto significa que el curso no estará orientado a desarrollar demostraciones de teoremas o hipótesis estadístico-matemáticas. Sin embargo, gran parte del curso se basará en el desarrollo de expresiones algebraicas que utilizan fuertemente el álgebra de matrices. Estas demostraciones serán dadas sólo para generar en los estudiantes las habilidades intelectuales y conocimientos analíticos – adecuados y suficientes – para que ellos puedan comprender claramente los procedimientos de cálculo que realizan softwares estadísticos más adecuados para el análisis de datos, basados en el análisis de modelos lineales.</p> <p>El énfasis en el enunciado de ejemplos y desarrollo analítico será orientado al análisis de datos desbalanceados y diseños experimentales incompletos.</p> <p>El primer capítulo realiza una revisión general del modelo de regresión simple. Se muestra brevemente el uso del procedimiento de mínimos cuadrados para resolver las ecuaciones normales y obtener los parámetros de la ecuación. Se plantea la forma tradicional de análisis de varianza y dócima de hipótesis. En el segundo capítulo incluye una revisión de algunos elementos del álgebra de matrices que es relevante para el análisis estadístico. El uso del álgebra matricial simplifica significativamente la formulación de MLs para diseños experimentales complejos de enunciar y analizar. Además, un conocimiento adecuado de los principales elementos y procedimiento de álgebra matricial permiten al investigador tener un completo conocimiento y control del tipo de análisis numérico que el software estadístico deberá utilizar para realizar las operaciones estadísticas que demandará el MLs seleccionado, durante el análisis de los datos derivados de la investigación inicial.</p> <p>El tercer capítulo extiende los principios revisados en el primer capítulo hacia la formulación de modelos de regresión múltiple. Aunque, el enfoque aquí es el planteamiento de modelos y su análisis mediante el uso de matrices. En particular, se mostrará y enfatizará el uso de formas cuadráticas en el análisis de</p>	

varianzas. Aunque algunos buenos estudiantes puede que no tengan un fuerte conocimiento matemático, ellos rápidamente podrán entender la racionalidad que une la formulación de un modelo lineal (de regresión) con el análisis de varianza utilizando formas cuadráticas. Al final del capítulo, se generalizará los resultados hacia la formulación de modelos complejos que involucren múltiples variables y se discutirá la racionalidad en el proceso de selección de variables, en función del uso de formas cuadráticas.

El cuarto capítulo está íntegramente dedicado a lo que comúnmente llamamos “análisis de varianza” (ANDEVA). Sin embargo, el capítulo está específicamente orientado a que el estudiante comprenda la relación entre (a) el diseño de un experimento, (b) la formulación de un ML asociado al experimento, (c) la estructuración inmediata de una tabla de ANDEVA que corresponda al ML y (d) el tipo de procedimientos matemáticos realizados por software estadísticos para obtener parámetros asociados al modelo y realizar las dójimas de hipótesis. Para obtener esta comprensión, se extenderá la formulación matricial de un ML asociado a una regresión hacia el caso de MLs que incluyen variables de clase (tratamientos, bloques, etc.). Luego se mostrará cómo se resuelven las ecuaciones normales y estructura un ANDEVA asociado a un ML con variables de clase, usando formas cuadráticas. Finalmente, se discutirá el tipo de hipótesis que pueden ser evaluadas en función del tipo de ML planteado inicialmente. Esta secuencia se utilizará para revisar los modelos: (a) simple en una vía (ejemplo: sólo efecto de tratamiento); (b) de dos vías (ejemplo: tratamientos y bloques); (c) de dos vías con interacción (ejemplo: bloques con repeticiones). El énfasis en el desarrollo de los procedimientos matriciales y analíticos estará enfocado en el análisis de datos desbalanceados, provenientes de diseños aleatorizados completos e incompletos. Posteriormente, se extenderá la formulación de MLs y análisis de varianza hacia la inclusión tanto de variables continuas o “covariantes” como al uso de variables “de clase” El uso de MLs que incluyen ambos tipos de variables es comúnmente denominado “análisis de covarianza” Igualmente, se utilizará álgebra de matrices y formas cuadráticas para comprender los procedimientos numéricos de análisis de los datos y pruebas de hipótesis asociados a este tipo de MLs. Durante las diferentes secciones que cubren el capítulo se discutirán ejemplos de modelos experimentales de utilidad para la investigación agropecuaria. El capítulo termina con el planteamiento de MLs de utilidad en el establecimiento de ensayos con variedades vegetales en diversos sitios o regiones edafoclimáticas.

El quinto capítulo está enfocado en la estimación de componentes de varianza. Inicialmente, se introducen los conceptos de MLs fijos, aleatorios y mixtos. Se discute cuidadosamente los supuestos de cada tipo de modelo y se entregarán ejemplos para el uso de cada tipo en diferentes disciplinas que cubren las ciencias agropecuarias. Se discutirán específicamente dos MLs: (a) simple o de clasificación en una vía y (b) de dos vías, con y sin interacción. El análisis incluirá la formulación matricial de cada modelo y la estimación de componentes de varianza siguiendo tres métodos de estimación: (a) de momentos (obtención de la esperanza de los cuadrados medios del ML); (b) de máximo verosimilitud (“maximum likelihood estimation” o MLE) y (c) de máximo verosimilitud restringida (“restricted maximum likelihood estimation” o REML). En función de los procedimientos analíticos presentados en el capítulo, se discutirán los criterios con los cuales se pueden determinar el número adecuado de bloques y sitios de ensayo, o tipo y número de tratamientos, para el establecimiento de los experimentos. Al igual que en el capítulo anterior, se enfatizará el desarrollo de los procedimientos matriciales y analíticos enfocados en el análisis de datos desbalanceados, provenientes de diseños aleatorizados completos e incompletos. Finalmente, se extenderá la aplicación de los diferentes métodos hacia modelos otros modelos tales como el anidado y el anidado-factorial.

El sexto capítulo está dedicado a introducir el uso de modelos lineales mixtos en el caso de mediciones repetidas. Inicialmente, se discutirá los diferentes tipos de mediciones repetidas y los diferentes tipos de diseños experimentales que los originan. Se enfatizará en los errores conceptuales y prácticos asociados al uso de procedimientos analíticos tradicionales para el estudio de mediciones repetidas. Posteriormente, se analizarán diferentes tipos de diseños experimentales en los cuales se debe o puede estudiar mediciones repetidas. Finalmente, se estudiará en detalle el uso de la teoría de los modelos lineales mixtos (MLM) para el análisis de este tipo de datos. En particular, se describirá en algoritmo de trabajo orientado a estimar las matrices de varianzas-covarianzas asociadas al MLM y el procedimiento recomendado para la dócima de hipótesis.

### **OBJETIVOS**

#### Objetivo General

Desarrollar un entendimiento conceptual y práctico de “qué es” y “para qué sirve” y “cómo puedo usar” un modelo estadístico lineal.

#### Objetivos específicos

1. Conocer cuáles son los principales métodos numéricos y estadísticos utilizados en la formulación y análisis de los modelos lineales.
2. Enunciar ejemplos de diseños experimentales aplicables en las ciencias agropecuarias, y analizar la forma de estudiarlos según los principios estadísticos y matemáticos que caracterizan a los modelos lineales.

### **CONTENIDOS**

Capítulo 1. Revisión de la regresión simple.

Capítulo 2. Álgebra de matrices útil para la estadística.

Capítulo 3. Regresión múltiple en notación matricial.

Capítulo 4. Variables de clase en regresión.

Capítulo 5. Estimación de componentes de varianza.

Capítulo 6. Introducción al uso de modelos lineales mixtos en el caso de mediciones repetidas.

### **MODALIDAD DE EVALUACIÓN**

Los alumnos serán evaluados a través de dos tareas que deberán ser desarrolladas individualmente. Para este efecto, los alumnos deberán trabajar durante el curso con dos bases de datos. La primera deberá ser generada mediante simulación estocástica. La segunda deberá ser obtenida con datos reales a partir de la investigación en la cual se encuentre trabajando el alumno o de otro origen. Las tareas serán entregadas a mediados y a fines del trimestre y evaluadas según una pauta que será entregada al comienzo del curso.

El curso utilizará fuertemente el software SAS, en su versión 9.1, 9.2, 9.3 o 9.4. El profesor entregará los programas en código SAS que serán utilizados para desarrollar las tareas asignadas durante el curso. De común acuerdo, se desarrollarán 2 o tres talleres en los cuales se revisará el estado de avance de las tareas y discutirá el uso del software SAS, aplicado al desarrollo del curso.

### **BIBLIOGRAFÍA:**

#### **Básica**

El profesor entregará a los alumnos los archivos digitales que contienen todos los apuntes que son utilizados en el curso. Estos apuntes son de su propia autoría.

#### **Recomendada**

- Draper, N.R.; Smith, H. 1981. Applied regression analysis. 2nd Ed. John Wiley & Sons. New York. 709 p.

- |  |  |
|--|--|
|  | <ul style="list-style-type: none"><li>- Searle, S.R. 1971. Linear Models. John Wiley &amp; Sons. New York. 532 p.</li><li>- Searle SR, Casella G, McCulloch CE (1992). Variance components. John Wiley &amp; Sons. New York. 501 p.</li><li>- Rawlings, J.O.; Pantula, S.G.; Dickey, D.A. 1998. Applied regression analysis. A research tool. 2nd. edition. Springer-Verlag. 657 pp.</li></ul> |
|--|--|