

Programa del curso MC-6104

Diseño de Experimentos

Escuela de Computación
Maestría en Ciencias de la Computación, Plan 472

I parte: Aspectos relativos al plan de estudios

1 Datos generales

Nombre del curso:	Diseño de Experimentos
Código:	MC-6104
Tipo de curso:	Teórico-práctico
Electivo o no:	Sí
Nº de créditos:	4
Nº horas de clase por semana:	3
Nº horas extraclase por semana:	9
% de las áreas curriculares:	-
Ubicación en el plan de estudios:	Bloque 2
Requisitos:	Álgebra Lineal, Cálculo, Probabilidades
Correquisitos:	-
El curso es requisito de:	MC-7203 – Seminario de Investigación - Tesis II.
Asistencia:	Obligatoria.
Suficiencia:	No.
Posibilidad de reconocimiento:	No.
Vigencia del programa:	I Semestre del 2023

2 Descripción general

“It doesn’t matter how beautiful your theory is, it doesn’t matter how smart you are. If it doesn’t agree with experiment, it’s wrong.”

—Richard Feynman (1918 - 1988)

Por sus características particulares, la Ciencia de la Computación ha sido históricamente débil en sus procedimientos experimentales. Este curso entrenará al estudiante en el uso de las técnicas clásicas de Diseño de Experimentos para validar o refutar las afirmaciones, hipótesis o teorías de las disciplinas computacionales, así como para ejercer controles de calidad en la práctica del desarrollo de software.

3 Objetivos Objetivo general.

Utilizar el Diseño de Experimentos en el desarrollo de investigación en Ciencia de la Computación.

En la siguiente tabla se describen los objetivos específicos del curso:

Objetivo(s) del curso
Al finalizar el curso, el estudiante estará en capacidad de:
Establecer hipótesis respecto a sistemas computacionales que puedan ser verificadas experimentalmente.
Diseñar experimentos científicos aplicados a la Ciencia de la Computación.
Usar software apropiado para analizar los resultados obtenidos de los experimentos dando niveles de confianza de la certeza de sus conclusiones.
Aplicar técnicas de Diseño de Experimentos para el control de calidad en el desarrollo de productos de software.
Usar las técnicas experimentales aprendidas en este curso en su propuesta de tesis, en el desarrollo de su tesis, y en investigaciones científicas en general.

4 Contenidos

A. Introducción al Diseño de Experimentos

1. Motivación.
2. Historia (Aristóteles, Escolásticos, Bacon, Galileo, Diderot, Lavoisier, Pasteur, Fisher, Coz, Taguchi).
3. ¿Qué es un experimento?
4. ¿Por qué hacer experimentos?
5. Aplicaciones científicas e industriales.
6. Estado experimental de la Ciencia de la Computación y la Ingeniería de Software.

B. Conceptos de Experimentos

1. Método Científico.
2. Ciencia vs. Pseudociencia (Karl Popper).
3. Hipótesis y experimentos.
4. Experimentos cuantitativos y cualitativos.
5. Grupo experimental y grupo de control.
6. Tipos de experimento.
7. Aleatorización - replicaciones.
8. Protocolos médicos.
9. Pruebas con animales y humanos - consideraciones éticas.
10. Experimentos y Control de Calidad.
11. Experimentos comparativos simples.

C. Repaso de Estadística

1. Conceptos generales de probabilidad y estadística.
2. Distribuciones de Probabilidad.
3. Media, varianza, desviación estándar.
4. Muestreo.
5. Inferencia Estadística.
6. Prueba de Hipótesis.

D. Experimentos Monofactoriales

1. Variable de Respuesta
2. Factores - Niveles – Tratamientos
3. Análisis de Varianza
4. Modelo de Efectos Fijos
5. Tamaño de muestra
6. Datos no balanceados
7. Verificación del modelo.
8. Procedimiento
9. Interpretación de resultados

E. Diseño de Bloques

1. Diseño de bloque completo aleatorio.
2. Cuadros latinos.
3. Cuadros greco-latinos.
4. Otros diseños.

F. Diseños Multifactoriales

1. Principios y Definiciones.
2. Eficiencia del modelo.
3. Modelo de 2 factores - ANOVA.
4. Interacciones de factores.
5. Modelo de 3 factores.
6. Modelo general.
7. Curvas de respuesta.
8. Diseños no balanceados.

G. Diseño factorial 2^k

1. Introducción.
2. Diseño 2^2 .
3. Diseño 2^3 .
4. Diseño general.
5. Análisis con réplica única.
6. Puntos centrales.

H. Diseños fraccionales

1. Introducción.
2. $\frac{1}{2}$ de 2^k .
3. $\frac{1}{4}$ de 2^k .
4. Modelo 2^{k-p} .
5. Factores confundidos.

| parte: Aspectos operativos

5 Metodología de enseñanza y aprendizaje

A. Calendario

El curso es de 3 horas lectivas por semana. Para propósitos oficiales, la asistencia se considerará obligatoria.

B. Exámenes Parciales

Todo el material será evaluado en 2 exámenes parciales no acumulativos durante el semestre, en los que se permite un "forro" oficial de una página tamaño carta.

C. Material de Clase

No hay un libro de texto para este curso. Sin embargo, hay excelentes libros respecto a este tema que se recomiendan en la Bibliografía.

D. Presentación Oral

Todos los estudiantes del curso deben presentar durante el semestre un paper asignado por el profesor. Dicha presentación oral debe seguir el formato de una conferencia científica (ver notas de clase de "Introducción a la Investigación").

Por respeto a la clase, se recomienda que cada presentador ensaye previamente y haga su mejor esfuerzo para explicar el material asignado. Ninguna presentación puede durar más de 20 minutos ni menos de 16 minutos (no improvise elementos no ensayados). Una ausencia el día de una presentación implica un 0 en el rubro respectivo. Los presentadores de cada fecha son responsables de verificar que el equipo necesario funcione correcta y oportunamente (consejo: revisar todo antes de empezar la clase).

E. Proyectos

Un componente esencial del aprendizaje y de la evaluación de este curso son los proyectos. Se debe trabajar en grupos de 2 o 3 personas (depende de la cantidad de estudiantes del grupo). Los grupos de trabajo requieren la aprobación del profesor.

Usualmente, se debe realizar un experimento, recolectar los datos, analizar los resultados y escribir un reporte. Es posible que algunos de estos proyectos requieran el desarrollo de software y el uso de algún paquete de software para analizar los resultados. Los proyectos deben haberse realizado estrictamente en las condiciones establecidas por el profesor, así que cualquier alteración de las reglas deja al proyecto sujeto al juicio unilateral del profesor (quien posiblemente le dará un 0 de calificación).

Los reportes deben estar escritos en inglés y se deben confeccionar con LATEX (formato de IEEE Transactions).

F. Tareas cortas y quices

Habrà una serie de quices y tareas cortas durante el semestre. Ningùn quiz será repuesto bajo ninguna circunstancia. Al final del semestre se eliminarán las dos notas menores.

G. Resúmenes de papers.

Todas las semanas se asignará al menos un paper para la lectura y comprensión del estudiante. El estudiante deberá entregar un resumen no menor a 2 páginas, ni mayor a 4. El resumen únicamente podrá ser entregado en la semana siguiente a la que se dejó la lectura. En todo el semestre se deben entregar al menos 10 resúmenes. Dentro del resumen se debe contener un análisis sobre el experimento propuesto por los autores para mostrar sus resultados. Dicho análisis debe contener mejoras, o propuestas de experimentos que permitan de una manera cuantificable y estadísticamente válida, demostrar los resultados de los autores.

H. Apuntes de clase

En cada clase habrá al menos 2 apuntadores oficiales, cuyo trabajo es tomar nota de todo lo sucedido en clase. Son inaceptables transcripciones o un montón de fotografías o capturas de pantalla de la clase. No se trata de un resumen ejecutivo, sino que se quiere que el apuntador comprenda el material de la clase y lo plasme en un documento, detallado, de mucha calidad, enriquecido con imágenes, gráficos, datos adicionales.

Los apuntes se subirán al Tec Digital en la carpeta denominada "Apuntes de clase", el nombre será de la forma DOE23mmdd-k.pdf,

donde mm es el mes, dd la fecha del día, y k un número secuencial para distinguir entre los apuntes de una misma clase. Si se falla en el formato del nombre, su nota será severamente penalizada. Los apuntes deben subirse al Tec Digital antes de la siguiente clase (se cuenta con 7 días). Todos deben ser apuntadores durante el semestre pues es parte de la nota final.

I. Tareas de análisis

En este rubro, a los estudiantes se les proporcionará un conjunto de datos que deben ser analizados utilizando las herramientas y métodos estudiados en el curso.

Requisitos:

Es necesario haber recibido por lo menos un curso de Probabilidad y Estadística en el bachillerato. También se requiere un buen conocimiento de UNIX, del lenguaje de programación C, y de R.

6 Evaluación

En resumen, la evaluación se desglosa según el siguiente cuadro:

Rubro	Valor (%)
R cheatsheet	1 %
Tarea probabilidad y estadística.	3 %
Tarea ANOVA, 2 ^k	4 %
Exposición oral	7 %
Proyecto distribución estadística	10 %
Resúmenes papers/ diseño del experimento	10 %
Quices/tareas cortas	7 %
2 Exámenes parciales (20% el mayor)	20% y 15%
Proyecto DOE Investigación/paper	20%
Trabajo en clase	3%

7 Bibliografía

Libros de texto (Sugerencias):

- **M. Anderson, P. Whitcomb**, "*DOE Simplified – Practical Tools for Effective Experimentation*", Second Edition, CRC Press, U.S.A., 2007.
- **D. Ariely**, "*The (Honest) Truth About Dishonesty*", Harper Collins Publishers, U.S.A., 2012.
- **S. W. Huck**, "*Reading Statistics and Research*", Pearson Education, Inc., Sixth Edition, U.S.A., 2012.
- **N. Juristo, A. Moreno**, "*Basics of Software Engineering Experimentation*", 2da edición. Kluwer Academic Publishers, Boston, U.S.A., 2010.
- **J. Lazar et al.**, "*Research Methods in Human- Computer Interaction*", John Wiley & Sons, U.S.A., 2010.

- **D. Montgomery**, "*Design and Analysis of Experiments*", Eight Edition, John Wiley & Sons, U.S.A., 2013.
- **J. Zobel**, "*Writing for Computer Science*", Second Edition, Springer Verlag, London, 2004.

8 Cronograma

Semana 1: Introducción y presentación del curso.

Semana 2: Quiz o tarea corta 1, sesión de exposiciones 1.

Semana 3: Quiz o tarea corta 2, sesión de exposiciones 2.

Semana 4: Quiz o tarea corta 3, sesión de exposiciones 3.

Semana 5: Tarea 1 probabilidad y estadística.

Semana 6: Quiz o tarea corta 4, sesión de exposiciones 4.

Semana 7: Proyecto distribución estadística.

Semana 8: Primer Parcial.

Semana 9: Quiz o tarea corta 5, sesión de exposiciones 5.

Semana 10: Quiz o tarea corta 6, sesión de exposiciones 6.

Semana 11: Quiz o tarea corta 7, sesión de exposiciones 7.

Semana 12: Tarea 2 Anova, 2^k.

Semana 13: Quiz o tarea corta 8, sesión de exposiciones/videos 8.

Semana 14: Quiz o tarea corta 9, sesión de exposiciones/videos 9.

Semana 15: Presentación Investigación-paper.

Semana 16: Segundo Parcial.