

FACULTAD DE CIENCIAS EXACTAS

Presentación curso de postgrado

Año	2024	Semestre	Primero
Nombre del Curso			
Métodos Estadísticos aplicados a las Ciencias Químicas y Biológicas			
Profesorxs Responsables (indicando las horas que participa en el dictado de clases)			
Dra. María Esperanza Ruiz (60 hs) Dr. Pablo Martín Demetrio (60 hs)			
Docentes Participantes (indicando las horas que participa en el dictado de clases)			
Dr. Sebastián Scioli Montoto (60 hs)			
Duración Total (en horas)	60		
Modalidad (Teórico, teórico-práctico, seminario, etc)	Teórico-práctico (modalidad híbrida). Los docentes harán la introducción teórica de cada tema (virtual), y luego se resolverán situaciones problemáticas, dudas y consultas (modalidad presencial).		
Tipo de evaluación prevista	Escrita, presencial, a libro abierto, con uso de software y calificación numérica		
Especificación clara si se lo considera válido para cubrir exigencias del Doctorado.			
<p>El presente curso ha sido diagramado con el objetivo de cubrir las exigencias de los estudiantes de doctorado de la facultad, ya que se han intentado incluir todos aquellos métodos estadísticos de uso común en la investigación en ciencias aplicadas, especialmente útiles para quienes inician su recorrido como investigadores en formación.</p> <p>El curso inicia retomando los conocimientos brindados en la asignatura Diseño de Experimentos y avanza hacia otras técnicas estadísticas no cubiertas por la formación de grado, desde un enfoque aplicado y a través de su implementación mediante el uso de software específico y de libre acceso. Los módulos temáticos del curso incluyen un repaso inicial, especialmente pensado para contemplar a los alumnos que pudieran no haber cursado Diseño de Experimentos en su carrera de grado, tales como los alumnos de la Lic. en Biotecnología y Biología Molecular.</p>			
Fecha de dictado	Del 16/5/24 al 11/7/24 (examen en agosto)	Cupo de alumnos	Mínimo: 9 Máximo: 30 (este número se basa en la relación docente/alumno, de manera de realizar correcciones y devoluciones a tiempo, como así también para el mejor desarrollo de las consultas online)
Inscripción desde	1/4/2024	Hasta	20/4/2024

Exigencias y requisitos de inscripción					
<p>El curso se dictará únicamente para estudiantes del doctorado de la Facultad de Ciencias Exactas, UNLP ¹.</p> <p>Los interesados deberán poseer título de grado relacionado a las Cs. Químicas y Biológicas en las que se enfoca el curso, el cual acredite como mínimo un curso de estadística. <i>(link para el formulario inscripción en la página de la Facultad).</i></p>					
Arancelamiento					
NO	X	SÍ		Montos	
Breve resumen de los objetivos y contenidos					

PRESENTACIÓN Y OBJETIVOS

El presente curso ha sido concebido para ofrecer una visión general de los métodos estadísticos más comúnmente requeridos por estudiantes de postgrado de disciplinas del área de las ciencias químicas y biológicas. Se pretende preparar a los estudiantes para resolver problemas comúnmente encontrados en proyectos de investigación, tomar decisiones basadas en datos, diseñar las experiencias que generen dichos datos y asimismo convertirse en lectores críticos de análisis estadísticos presentados en la literatura científica.

El curso supone que los estudiantes tienen conocimientos matemáticos adecuados y al menos un curso previo en estadística. Dado que los problemas reales no son nunca tan sencillos que puedan resolverse mecánicamente, se pondrá énfasis en la necesidad de pensar en la naturaleza real del problema científico en sí, de reflexionar sobre los gráficos y otro tipo de representaciones, así como de entender los principios estadísticos útiles en cada caso y sus consecuencias (y limitaciones) prácticas.

Debido a que el curso está pensado para participantes con diferentes formaciones de grado, el mismo se inicia con un primer módulo de naturaleza niveladora, que a la vez permitirá iniciar a todo el grupo de alumnos en el uso del software propuesto. A continuación, el curso avanza hacia modelos de regresión lineal múltiple y diseño de experimentos avanzado, lo que resulta especialmente útil para los estudiantes de posgrado que realicen tesis doctorales u otras tareas de investigación.

Por lo tanto, el objetivo general del curso es continuar con la formación general en estadística iniciada en el grado, profundizando en las principales técnicas de análisis de datos empleadas en experimentación e investigación científica, como así también en la posibilidad de su ejecución mediante el uso de un software especializado, desarrollado en el país y accesible gratuitamente en su versión estudiantil (Infostat, U. de Córdoba).

¹ Si éstos superaran el cupo máximo del curso, se procederá a realizar un sorteo entre los interesados.

CONTENIDOS

Módulo 1: Revisión de conceptos básicos

Manejo de datos estadísticos. Descriptores de un conjunto de datos. Distribuciones de frecuencias. Variable aleatoria. Distribución de probabilidad de variables aleatorias discretas y continuas. Distribución de estadísticos muestrales. Estimación de parámetros por intervalos de confianza. Teoría general de las pruebas de hipótesis. Tipos de errores. Inferencias basadas en una y dos muestras.

Módulo 2: Análisis de la varianza y diseño de experimentos

Modelos lineales con variables categóricas: experimentos con un solo factor. Análisis de la varianza. Diseños completamente aleatorizados (DCA). Cuadrados medios y su esperanza. Verificación de los supuestos del modelo. Comparaciones a posteriori. Criterios a priori: contrastes ortogonales. Estudio de tendencia. Experimentos factoriales. Modelos de clasificación según dos o más factores cruzados. Concepto de réplica oculta e interacción. Diseño en bloques completos aleatorizados (DBCA). Medidas de eficiencia. Ventajas y aplicaciones. Efectos fijos y aleatorios. Estimación de componentes de varianza. Diseños jerárquicos con factores anidados. Aplicaciones. Modelo estadístico. Hipótesis y estadísticos de prueba para el caso de factores fijos, aleatorios y mixtos. Diseños que combinan factores anidados y cruzados, fijos y aleatorios. Cálculo de la esperanza de los cuadrados medios. Pruebas F parciales. Método de Satterthwaite para la estimación de los grados de libertad.

Módulo 3: Análisis de regresión y de correlación

Repaso de análisis de regresión lineal simple. Estimadores mínimos cuadrados. Supuestos del modelo. Pruebas de hipótesis e intervalos de confianza. Regresión en términos matriciales. Regresión lineal múltiple. Estimación de parámetros. Pruebas de hipótesis e intervalos de confianza. Bondad de ajuste y adecuación del modelo. Estudio de residuales. Pruebas formales y gráficas para la verificación de supuestos. Correlación: simple y parcial. Alternativas para la selección de variables: Método de selección paso a paso (“stepwise”). Predicciones. Modelos polinómicos de regresión. Modelos no lineales pero linealizables por transformación. Regresión lineal ponderada. Variables indicadoras o “dummy”.

Módulo 4: Otros métodos estadísticos

Inferencia no paramétrica entre dos muestras independientes: prueba U de Mann-Whitney. Inferencia no paramétrica entre dos muestras pareadas: prueba del signo, prueba de Wilcoxon. Análisis de la varianza no paramétrico: prueba de Kruskal-Wallis, Friedman.

MODALIDAD DE DICTADO

El curso será dictado en modalidad teórico-práctica (hídrida). Para ello, se implementará el esquema presentado en el Anexo, el cual incluye:

- » Videos de teoría. En general, una clase consta de uno o más videos, con una duración aproximada de 40 – 60 minutos cada uno.
- » Bibliografía. Se provee al estudiante de las presentaciones de clase, como así también del/los capítulos de libro/s correspondiente al tema de la clase, donde profundizar lo visto en el video².
- » Tutoriales de resolución de ejercicios. Uno o más videos breves, de 20 - 30 min aproximadamente, principalmente enfocados en el uso del software.
- » Guía de ejercicios correspondiente al tema tratado.
- » Encuentros presenciales de 5 h de duración, una vez por semana.

REQUISITOS PARA LA APROBACIÓN

Para obtener la acreditación del curso, el alumno deberá:

- a. Haber asistido al 80% de las clases presenciales programadas.
- b. Entregar y aprobar los trabajos prácticos correspondientes a cada uno de los módulos (en forma individual o en grupos de a dos)
- c. Aprobar un examen final presencial, individual, a libro abierto y a realizarse con el software.
- d. La calificación final tendrá en cuenta la nota obtenida en b) y c).

Contacto con los responsables del curso			
Dirección	Calle 47 y 115 - La Plata		
Teléfono	221 423 5333 int. 43	Fax	
Correo electrónico	metodos.estadisticos.unlp@gmail.com		

² Adicionalmente, se encuentran disponibles para su consulta todos los libros mencionados en el apartado bibliografía

ANEXO – Programa detallado de actividades

		Carga horaria	
		Virtuales	Presenciales
Sem 1	Se estudiarán los modelos lineales cuyas variables explicativas son categóricas. Concepto de Análisis de la Varianza, descomposición de la suma de cuadrados, planteo de hipótesis y estadísticos de prueba. Ejemplos y aplicaciones.	1,5	
	Supuestos del modelo, justificación de su necesidad. Verificación de supuestos, metodologías analíticas y gráficas. Comparaciones particulares entre los niveles de un factor: métodos a posteriori y contrastes a priori. Aplicación de contrastes para evaluar tendencias.	1,0	
	Experimentos factoriales de dos o más factores cruzados en DCA. Aplicaciones, ventajas. Concepto de réplica oculta e interacción entre factores. Modelo estadístico, descomposición de la suma de cuadrados y análisis de la varianza. Verificación de supuestos y comparaciones a posteriori.	1,5	
	Ejercicios de aplicación.		5
Sem 2	Se continuará con los modelos de clasificación según dos o más factores en diseños que involucran bloques (DBCA, Cuadrados Latinos). Aplicaciones y ventajas. Eficiencia del bloque.	1,5	
	Se estudiará la diferencia entre efectos fijos y aleatorios. Consideraremos el ANOVA como método para la estimación de los componentes de varianza.	1,0	
	Ejercicios de aplicación.		5
Sem 3	Se estudiarán los diseños jerárquicos o anidados como alternativa a los factoriales cruzados. Aplicaciones, ventajas y desventajas. Modelo estadístico. Hipótesis y estadísticos de prueba según la naturaleza de los factores (fijos, aleatorios o mixtos).	1,5	
	Modelos que combinan factores anidados y cruzados, fijos y aleatorios. Cálculo de esperanzas de CM. Construcción de pruebas F aproximadas, grados de libertad (método de Satterthwaite).	1,0	
	Ejercicios de aplicación.		5
Sem 4	Consulta - Entrega de ejercicios Módulo 1		4
Sem 5	Repaso Regresión Lineal Simple (RLS): se recordará la estimación por MCO supuestos del modelo, prueba de hipótesis e intervalos de confianza. Repaso de matrices, operaciones con matrices. RLS y MCO en notación matricial.	1,5	
	Se considerarán los estadísticos asociados con la bondad de ajuste. Se realizará el abordaje de análisis de los residuales mediante pruebas formales y gráficas para verificación de supuestos y la comprobación de la adecuación del modelo. Ejemplos aplicados con uso de software.	1,5	
	Ejercicios de aplicación.		5
Sem 6	Regresión lineal múltiple (RLM): se abordará la estimación de parámetros, pruebas de hipótesis e intervalos de confianza. Interpretación del modelo de RLM en el contexto del problema abordado. Se retomará el concepto de análisis de correlación y se presentará el concepto de correlación parcial. Ejercicios de aplicación.	1,5	
	A partir de los conceptos anteriores se abordarán posibles alternativas para la selección de variables en RLM. Se estudiará en detalle el método "stepwise" Se considerarán algunas estrategias de análisis cuando los modelos no verifican los supuestos: modelos linealizables y regresión lineal ponderada.	1,5	
	Ejercicios de aplicación.		5
Sem 7	Se introducirá el concepto de variable "dummy" o indicadora en el marco de los modelos lineales. Relación con el análisis de la covarianza.	1,5	

	Se estudiarán técnicas de inferencia no paramétrica entre dos muestras independientes, dos muestras pareadas y más de dos muestras. Ejemplos aplicados con uso de software	1,5	
	Ejercicios de aplicación.		5
	Consulta - Entrega de ejercicios Módulo 2		4,0
	Examen (fecha a coordinar)		4,0
	Horas totales =	18,0	42,0

BIBLIOGRAFÍA

DI RIENZO, J., CASANOVES, F. GONZALES, L., TABLADA, E., DÍAZ, M., ROBLEDO, C. & BALZARINI, M. (2005). *Estadística para las Ciencias Agropecuarias*. 6ª Ed. Córdoba, Argentina: Ed. Trunfar.

BOX, G., HUNTER, W. & HUNTER, J. (2005) *Estadística para investigadores. Introducción al diseño de experimentos, análisis de datos y construcción de modelos*. México [etc], Ed. Reverté.

MONTGOMERY, D. & RUNGER, G. (2012). *Probabilidad y estadística aplicadas a la ingeniería*. México, Limusa Wiley.

MONTGOMERY, D., PECK, E. & VINING, G. (2006). *Introducción al análisis de regresión lineal*, 3ª Ed. México, CECSA.

KUEHL, R. (2001). *Diseño de experimentos: principios estadísticos para el análisis y diseño de investigación*. 2ª Ed. México [etc.], Thomson Learning.

MONTGOMERY, D. (2004). *Diseño y análisis de experimentos*. 2ª Ed. México, Limusa Wiley.

OTT, R.L. & LONGNECKER, M. (2010). *An Introduction to Statistical Methods and Data Analysis*. 7ª Ed. Belmont: Brooks/Cole

WACKERLY, D., MENDENHALL, W. & SCHEAFFER, R. (2008) *Mathematical Statistics with Applications*. 7ª Ed. Australia [etc], Thomson Learning.

WALPOLE, R., MYERS, R., & MYERS, S. (2012). *Probabilidad y estadística para ingeniería y ciencias*. 9ª Ed. Distrito Federal, Pearson Educación.

CONOVER, W. J. (1999). *Practical nonparametric statistics*. 3rd Ed. New York, John Wiley & Sons Inc.