
• Información de los profesores y del monitor

Nombre profesor: Adolfo J. Quiroz

Correo electrónico: ajquiroz@gmail.com,aj.quiroz1079@uniandes.edu.co

Horario y lugar de atención:

• Introducción y descripción general del curso

Este curso pretende introducir al estudiante a las ideas, métodos fundamentales y bases teóricas de la Estadística No Paramétrica, tanto en lo referente a la teoría clásica basada en rangos, tal como se presenta en Randles y Wolfe, 1981, como en lo concerniente a métodos basados en el remuestreo no paramétrico y remuestreo suavizado, tal como aparecen en Davison y Hinkley, 1999. Además, se pretende que el estudiante alcance una cierta familiaridad con algunos métodos en desarrollo actual tales como métodos basados en grafos o nociones de profundidad para datos multivariados y datos funcionales.

• Objetivos de la asignatura y competencias a desarrollar

- Que el estudiante pueda enunciar las ventajas y posibles desventajas de utilizar procedimientos estadísticos no-paramétricos en el análisis de datos.
- Que el estudiante pueda describir, implementar, utilizar y evaluar los procedimientos no-paramétricos para el problema de estimación de localización, para el problema de dos muestras, y para el problema de comparación de parámetros de dispersión, incluyendo los procedimientos de rango signado de Wilcoxon, del signo de Fisher, el de Wilcoxon-Mann-Whitney y el de Ansari-Bradley.
- Que el estudiante sea capaz de producir e interpretar los intervalos de confianza no-paramétricos tipo Hodges-Lehman asociados a los estadísticos mencionados en el punto anterior.
- Que el estudiante conozca la teoría asintótica correspondiente a los procedimientos mencionados en los puntos anteriores y pueda evaluar, mediante simulaciones, la validez de la aproximación asintótica en muestras finitas.
- Que el estudiante pueda manejar el principio de permutación para desarrollar estadísticos condicionalmente no paramétricos.
- Que el estudiante conozca el significado de las medidas no-paramétricas de asociación de Kendall (τ) y de Spearman (ρ), pueda calcular dichas medidas en conjuntos de datos y producir los correspondientes intervalos de confianza no-paramétricos para τ y ρ y los correspondientes estadísticos para la hipótesis nula de independencia.
- Que el estudiante comprenda el principio de sustitución y la utilidad del remuestreo en la estimación no-paramétrica.
- Que el estudiante pueda aplicar el remuestreo para la estimación no-paramétrica del sesgo o la varianza de un estimador, incluyendo la estimación por intervalos.
- Que el estudiante conozca y pueda deducir las propiedades de los estimadores de densidades por núcleos
- Que el estudiante pueda aplicar el remuestreo para la estimación no-paramétrica de curvas de regresión, incluyendo la estimación por intervalos.

- Que el estudiante pueda explicar cómo las propiedades de ciertos grafos asociados a la matriz de distancias entre los datos de una muestra, pueden servir para extender diversos procedimientos no paramétricos al contexto multivariado.
- Que el estudiante pueda explicar la teoría que justifica la validez del remuestreo no paramétrico y el remuestreo suavizado, al menos en algunos casos.
- Que el estudiante pueda describir, implementar y aplicar a datos multivariados los procedimientos basados en grafos, incluyendo los descritos en Quiroz (2006) que generalizan al contexto multivariado algunos de los métodos no paramétricos univariados.
- Que el estudiante pueda aplicar análisis de datos relevantes en el contexto de datos funcionales, incluyendo métodos no paramétricos para el problema de dos muestras.

- **Contenido de la asignatura**

RW se refiere al texto de Randles y Wolfe, ET al de Efron y Tibshirani, DH al de Davison y Hinkley.

Semana	Contenido	Secciones de algún texto
1	Motivación. Porqué estadística no paramétrica. Ejemplos. Estadísticos de orden. La técnica de igualdad en distribución	RW: 1.1 a 1.3
2	Estadísticos no paramétricos o de "distribución libre". Estadísticos de conteo. Estadísticos de rangos.	RW: 2.1 a 2.2
3	Estadísticos de conteo y rangos. U-estadísticos de una muestra.	RW: 2.3 y 3.1
4	Teoremas auxiliares para U-estadísticos. TLC para U-estadísticos de una muestra.	RW: 3.2 y 3.3
5	TLC para U-estadísticos de dos muestras. Estadísticos asintóticamente no paramétricos.	RW: 3.4 y 3.5
6	Intervalos de Confianza tipo Hodges-Lehman para estadísticos no paramétricos. Parcial 1, jueves 2 de marzo, 20%.	RW: 6.1
7	Intervalos de Confianza tipo Hodges-Lehman para estadísticos no paramétricos, continuación. Propiedades asintóticas de los intervalos Hodges-Lehman.	RW: 6.1 y 6.2
8	Entrega Proyecto 1, lunes 13 de marzo, 20% Estimación puntual asociada a estadísticos no paramétricos. Propiedades de muestra pequeña y propiedades asintóticas.	RW: 7.1 a 7.3

9	Estadísticos lineales de rango. Definición y Teoría Asintótica (este tema es opcional)	RW: 8.1 a 8.4
10	Estadísticos condicionalmente no paramétricos. Medidas de correlación.	RW 11.1, 12.3
11	Idea del remuestreo. Teorema de Glivenko-Cantelli. El principio de sustitución y la función de distribución acumulativa empírica. El remuestreo básico. Fuentes de error.	DH: 2.1 a 2.5 ET: Caps. 4 y 5.
12	Aproximación no paramétrica para varianza y sesgo. Parcial 2, jueves 20 de abril, 20%.	DH: 2.7
13	Métodos basados en submuestras. Jackknife y cross-validación. Estimación de densidades por núcleos. Remuestreo suavizado.	DH: 2.8, 3.4
14	Entrega Proyecto 2, martes 2 de mayo, 20%. Remuestreo en modelos de regresión	6.1 a 6.3
15	Exposición de temas de lectura. Entrega de informe escrito, domingo 14 de mayo. Las exposiciones pueden ocupar la primera semana de finales también. Posibles temas: Métodos basados en grafos para el problema de dos-muestras multivariado, para "clustering" y para estimación de dimensión. Nociones de profundidad y el problema de dos muestras para datos funcionales.	

- **Criterios de evaluación y aspectos académicos**

a. Porcentajes de evaluación:

Evaluación	Porcentaje de la nota total
Dos exámenes parciales	20% cada uno
Dos proyectos computacionales	20% cada uno
Una exposición de tema de lectura	20%

b. Fechas Importantes:

- Inicio de clases:** semana del 23 al 27 de enero.
- Parciales:** Jueves 2 de marzo y jueves 20 de abril.
- Entrega de proyectos:** lunes 13 de marzo y martes 2 de mayo.

- d. **Exposiciones:** del 8 al 19 de mayo.
- e. **Entrega del 30% de la nota del curso:** hasta el viernes 17 de marzo.
- f. **Último día para retiro de cursos:** viernes 24 de marzo.

- **Bibliografía Principal**

- ✓ Randles, R. H. y Wolfe, D. A. Introduction to the Theory of Nonparametric Statistics. Kreiger Publishing, 1991.
- ✓ Hollander, M. y Wolfe, D. A. Nonparametric Statistical Methods. Wiley. 1999.
- ✓ Efron, B. y Tibshirani, R.: An Introduction to the Bootstrap, CRC Press, 1998.
- ✓ Davison, A. C. y Hinkley, D. V. Bootstrap Methods and Their Application. Cambridge University Press, 2002.
- ✓ Quiroz, A. J. Graph Theoretic Methods. En Encyclopedia of Statistical Sciences, 2nd Edition, S. Kotz, C. B. Read, N. Balakrishnan y B. Vidakovic, editores. Vol 5, pp 2910-2916, Wiley and Sons, New York. 2006.

Nota: Todos los libros de texto se encuentran en la Biblioteca "Ramón de Zubiría" y serán colocados en reserva para que todos puedan tener acceso.

- **Reclamos**

Según el Régimen Académico de la Universidad, si se trata de una prueba escrita, el estudiante deberá dirigir el reclamo por escrito, dentro de los ocho (8) días hábiles siguientes al día en que se da a conocer la calificación. El profesor cuenta con diez (10) días hábiles para responderle.

Política de aproximación de notas

Las notas de los parciales, proyectos y exposiciones, se darán sobre 20 puntos. Estas notas se suman para obtener un total sobre 100, que se llevará, dividiendo entre 20, a la escala del 1 al 5. Esta definitiva se redondeará a la décima más cercana, por ejemplo, 4.37 se redondeará a 4.4, mientras que 3.53 se redondeará a 3.5
La Nota Definitiva aprobatoria mínima es 3 puntos.