

Instituto Tecnológico Autónomo de México
EST-46114: ESTADÍSTICA MULTIVARIADA Y DATOS CATEGÓRICOS

Prof. Juan Carlos Martínez Ovando
Primavera 2019

Objetivos

- Comprender los fundamentos de los métodos estadísticos aplicados a datos multidimensionales (e.g. vectores, matrices, tensores y formas).
- Implementar procedimientos predictivos asociados con estos modelos.
- Interpretar resultados analíticos prácticos.

Temas

Parte 1: Preliminares

1. Álgebra matricial
2. Distribuciones multidimensionales
3. Análisis de regresión múltiple
4. Comparación y selección de modelos

Parte 2: Análisis multivariado

5. Análisis de componentes principales
6. Análisis de factores
7. Escalamiento multidimensional
8. Análisis de correlación canónica
9. Clasificación supervisada
10. Agrupamiento no supervisado

Parte 3: Datos categóricos

11. Modelos log-lineales
12. Modelos de regresión multinomial y ordinal
13. Árboles de regresión
14. Datos perdidos
15. Selección y reducción de dimensionalidad
16. Tensores y modelos de grafos probabilísticos

Dinámica

- Cada tema se cubrirá en dos sesiones de 1.5 horas.
 - La primera sesión consistirá en revisión teórica.
 - La segunda sesión consistirá en implementaciones e interpretación.
- Cada sesión vendrá acompañada por ejercicios a casa.
- Cada sesión tendrá asociada lecturas complementarias de libros o artículos de investigación acerca de extensiones de temas.¹

Librerías especiales en R²

- MVA (<https://cran.r-project.org/web/packages/MVA/index.html>)
- AppliedPredictiveModeling (<https://cran.r-project.org/web/packages/AppliedPredictiveModeling/index.html>)
- gRbase (<https://cran.r-project.org/web/packages/gRbase/index.html>)
- rTensor (<https://cran.r-project.org/web/packages/rTensor/index.html>)

Datos para proyectos

- EMOVI 2011 (<https://ceey.org.mx/contenido/que-hacemos/emovi/>)
- EMIF-N y S 2015 (<https://www.colef.mx/emif/>)
- ENOE 2018 (<https://www.inegi.org.mx/programas/enoe/15ymas/>)

Horarios de atención

- Martes y Jueves: 8:30-10:00 h.
- Martes y Jueves: 11:30-13:00 h.
- Previa cita via email: `juan.martinez.ovando@itam.mx`

Evaluación

- Tareas: 15 %
- Examen parcial 1: 20 %
- Examen parcial 2: 20 %
- Examen parcial 3: 20 %
- Examen/proyecto final: 25 %

Fechas importantes

- Examen parcial 1: Martes 5 de marzo
- Examen parcial 2: A casa con entrega el jueves 25 de abril
- Examen parcial 3: A casa con entrega el jueves 16 de mayo
- Examen/proyecto final: Fecha determinada por control escolar en el periodo 18 de mayo-1 de junio

¹Aunque serán de interés, los temas cubiertos en las lecturas complementarias no formarán parte de la evaluación del curso.

²Para usarlas en la tercera parte dle curso.

Referencias

- Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. Springer, New York. [Parte 3](#).
- Agresti, A. (2002). *Categorical Data Analysis*. John Wiley & Sons, Berlin. [Parte 3](#).
- Bhisop, Y., et al. (2007). *Discrete Multivariate Analysis: Theory and Practice*. Springer, London. [Parte 3](#).
- Coppi, R. (1994). An introduction to multiway data and their analysis. *Computational Statistics and Data Analysis*, (18):3-13. [Parte 3](#).
- Everytt, B. y Hothorn, T. (2011). *An Introduction to Applied Multivariate Analysis with R*. Springer, London. [Parte 2](#).
- Friendly, M. (2016). *Visualizing Categorical Data with SAS and R*. SAS Inc. [Parte 3](#).
- Højsgaard, S. et al. (2012). *Graphical Models with R*. Springer, London. [Parte 3](#).
- Izenman, A.J. (2008). *Modern Multivariate Statistical Techniques*. Springer, London. [Parte 2](#).
- Johnson, V.E. y Albert, J. (1999). *Ordinal Data Modeling*. Springer, London. [Parte 3](#).
- Kuhn, M. y Johnson, K. (2013). *Applied Predictive Modeling*. Springer, New York. [Parte 2](#).
- Press, J.S. (2005). *Applied Multivariate Analysis*. Krieger Pub., New York. [Partes 1-2](#).

Versión: 5 de abril de 2019