



UNIVERSIDADE DA CORUÑA *Universidade de Vigo*

## MÁSTER EN TÉCNICAS ESTADÍSTICAS

### *DATOS IDENTIFICATIVOS DE LA MATERIA*

Nombre de la materia: **Contrastes de Especificación**

Número de créditos ECTS: 5

Curso académico: 2023/2024

Profesorado:

Juan Carlos Pardo Fernández (3 ECTS, coordinador)

Wenceslao González Manteiga (2 ECTS)

### *OBJETIVOS DE LA MATERIA*

En esta materia se pretende dar a conocer las técnicas clásicas y recientes de contrastes de bondad de ajuste para la distribución y para la función de regresión. Se analizarán para ello las metodologías más importantes, que incluyen a los contrastes basados en procesos empíricos, los contrastes basados en técnicas de suavizado y otros tipos de contrastes. Se estudiarán los métodos más conocidos para resolver cada problema de contraste, al tiempo que se buscará una visión global sobre los múltiples trabajos existentes dentro de la temática de esta asignatura, de modo que se desarrolle la capacidad para la búsqueda, comprensión y profundización en líneas más específicas.

### *CONTENIDOS DE LA MATERIA*

#### **1. Introducción.**

Elementos de un contraste de hipótesis. Contrastes paramétricos y no paramétricos. Propiedades del p-valor. El problema de la multiplicidad de contrastes y posibles

soluciones. Diseño de estudios de Monte Carlo.

## **2. Contrastes de bondad de ajuste para la distribución.**

Revisión de herramientas gráficas: pp-plots y qq-plots. Contrastes basados en la función de distribución. Contrastes basados en la función de densidad. Contrastes basados en la función cuantil. Contrastes basados en la función característica.

## **3. Contrastes de normalidad.**

Contrastes de especificación para modelos paramétricos particulares. Contrastes de normalidad univariante. Contrastes de normalidad multivariante.

## **4. Contrastes de independencia y otros contrastes sobre la distribución.**

Herramientas gráficas para detectar dependencia. Contrastes de independencia. Otros contrastes: contrastes de simetría, contraste de un posible punto de cambio.

## **5. Contrastes de especificación para modelos de regresión basados en la estimación de la función de regresión.**

Visión general de las técnicas de suavizado en problemas de regresión. Aplicación a los contrastes sobre la función de regresión. Aproximaciones bootstrap.

## **6. Contrastes de especificación para modelos de regresión basados en la función de regresión integrada.**

La función de regresión integrada. Descripción del test. Convergencia en distribución del proceso de contraste. Aproximaciones bootstrap de la distribución del proceso.

## **7. Otros contrastes sobre la regresión.**

Contrastes de igualdad de curvas de regresión. Contrastes de significación de variables. Contrastes de homocedasticidad. Contrastes para la varianza condicional.

## ***BIBLIOGRAFÍA BÁSICA Y COMPLEMENTARIA***

### **Libros:**

Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley.

Billingsley, P. (1999). *Convergence of Probability Measures (2ª edición)*. Wiley.

Claeskens, G.; Hjort, N.L. (2008). *Model Selection and Model Averaging*. Cambridge University Press.

Conover, W.J. (1999). *Practical Nonparametric Statistics (3ª edición)*. Wiley.

D'Agostino, R.B.; Stephens, M.A., eds. (1986). *Goodness-of-Fit Techniques*. Marcel Dekker, Inc.

- Efron, B.; Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall.
- Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.
- Härdle, W.; Müller, M.; Sperlich, S.; Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Springer.
- Hart, J. D. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*. Springer-Verlag, New York.
- Huber-Carol, C.; Balakrishnan, N.; Nikulin, M.S.; Mesbah, M., eds. (2002). *Goodness-of-Fit Tests and Model Validity*. Birkhäuser.
- Kvam, P.H.; Vidakovic, B. (2007). *Nonparametric Statistics with Applications to Science and Engineering*. Wiley.
- Rayner, J.C.W.; Thas, O.; Best, D.J. (2009). *Smooth Tests of Goodness-of-Fit. Using R*. Wiley.
- Rohatgi, V.K. (1976). *An Introduction to Probability Theory and Mathematical Statistics*. Wiley.
- Rohatgi, V.K. (1984). *Statistical Inference*. Wiley (Reeditado por Dover en 2003).
- Thas, O. (2010). *Comparing Distributions*. Springer.
- Thode, H.C. (2002). *Testing for Normality*. Marcel Decker, Inc.
- Vélez Ibarrola, R.; García Pérez, A. (2012). *Principios de Inferencia Estadística*. UNED.
- Wasserman, L. (2006). *All of Statistics. A Concise Course in Statistical Inference*. Springer.
- Zhu, L.-X. (2005). *Nonparametric Monte Carlo Tests and Their Applications*. Lecture Notes in Statistics, vol. 182. Springer.

### **Artículos:**

- Ahmad, I.A.; Li, Q. (1997). Testing symmetry of an unknown density function by kernel method. *Journal of Nonparametric Statistics* 7, 279-293.
- Bickel, P.J.; Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *The Annals of Statistics* 1, 1071-1095.
- Bierens, H.J. (1990). A consistent conditional moment test of functional form. *Econometrica* 58, 1443-1458.
- Bierens, H.J.; Ploberger, W. (1997). Asymptotic theory of integrated conditional moment tests. *Econometrica* 65, 1129-1152.

- Butler, C.C. (1969). A test for symmetry using the sample distribution function. *The Annals of Mathematical Statistics* 40, 2209-2210.
- Cao, R.; Lugosi, G. (2005). Goodness-of-fit tests based on the kernel density estimator. *Scandinavian Journal of Statistics* 32, 599-616.
- Csörgö, M.; Horváth, L. (1998). Invariance principles for changepoint problems. *Journal of Multivariate Analysis* 27, 151-168.
- Delgado, M.A. (1993). Testing the equality of nonparametric regression curves. *Statistics and Probability Letters* 17, 199-204.
- Delgado, M.A.; González Manteiga, W. (2001). Significance testing in nonparametric regression based on the bootstrap. *The Annals of Statistics* 29, 1469-1507.
- Dette, H. (1999). A consistent test for the functional form of a regression based on a difference of variance estimators. *The Annals of Statistics* 27, 1012-1040.
- Dette, H.; Munk, A. (2003). Some methodological aspects of validation of models in nonparametric regression. *Statistica Neerlandica*, 57 (2), 207-244.
- Fan, Y. (1994). Testing the goodness-of-fit of a parametric density function by kernel method. *Econometric Theory* 10, 316-356.
- Fan, Y.; Li, Q. (2000). Consistent model specification tests: kernel-based tests versus Bierens' ICM tests. *Econometric Theory* 16, 1016-1041.
- Fisher, N.I.; and Switzer, P. (2001). Graphical assessment of dependence: is a picture worth 100 tests? *The American Statistician*, 55, 233-239.
- Genest, C.; Boies, J.-C. (2003). Detecting dependence with Kendall plots. *The American Statistician*, 57, 275-284.
- Gozalo, P. (1993). A consistent model specification test for nonparametric estimation of regression function models. *Econometric Theory* 9, 451-477.
- González-Manteiga, W.; Crujeiras, R. M. (2013). An updated review of goodness-of-fit tests for regression models. *TEST*, 22, 361-411.
- Härdle, W.; Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *The Annals of Statistics* 21, 1926-1947.
- Henze, N. (2002). Invariant tests for multivariate normality: a critical review. *Statistical Papers*, 43, 467-506.
- Horowitz, J.; Härdle, W. (1994). Testing a parametric model against a semi-parametric alternative. *Econometric Theory* 10, 821-848.
- Li, Q.; Wang, S. (1998). A simple consistent bootstrap test for a parametric regression function. *Journal of Econometrics* 87, 145-165.

Lilliefors, H.W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62, 399-402.

Lilliefors, H.W. (1969). On the Kolmogorov-Smirnov test for the Exponential distribution with mean unknown. *Journal of the American Statistical Association*, 64, 387-389.

Mann, H.B.; Whitney, D.R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18, 50-60.

Massey, F.J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46, 68-78.

Massey, F.J. (1951). The distribution of the maximum deviation between two sample cumulative step functions. *The Annals of Mathematical Statistics*, 22, 125-128.

Massey, F.J. (1952). Table for the deviation between two sample cumulatives. *The Annals of Mathematical Statistics*, 23, 435-441.

Miles, D.; Mora, J. (2003). On the performance of nonparametric specification tests in regression models. *Computational Statistics and Data Analysis* 42, 477- 490.

Pettitt, A.N. (1979). A nonparametric approach to the change-point problem. *Journal of the Royal Statistical Society, Series C*, 28, 126-135.

Rothman, E.D.; Woodroffe, M. (1972). A Cramér von-Mises type statistic for testing symmetry. *The Annals of Mathematical Statistics*, 43, 2035-2038.

Shapiro, S.S.; Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52, 591-611.

Shapiro, S.S., Wilk, M.B.; Chen, H.J. (1968). A comparative study of various tests for normality. *Journal of the American Statistical Association* 63, 1343- 1372.

Stute, W. (1997). Nonparametric model checks for regression. *The Annals of Statistics* 25, 613-641.

Stute, W., González Manteiga, W.; Presedo Quindimil, M. (1998). Bootstrap approximations in model checks for regression. *Journal of the American Statistical Association* 93, 141-149.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics*, 1, 80-83.

Zheng, J.X. (1996). A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics* 75, 263-289.

## **COMPETENCIAS BÁSICAS, TRANSVERSALES, GENERALES Y ESPECÍFICAS**

En esta materia se trabajarán las competencias básicas, generales y transversales recogidas en la memoria del título. Se indican a continuación cuáles son las competencias específicas que se potenciarán en esta materia:

[CE1] Conocer, identificar, modelar, estudiar y resolver problemas complejos de estadística e investigación operativa, en un contexto científico, tecnológico o profesional, surgidos en aplicaciones reales.

[CE3] Adquirir conocimientos avanzados de los fundamentos teóricos subyacentes a las distintas metodologías de la estadística y la investigación operativa, que permitan su desarrollo profesional especializado.

[CE4] Adquirir las destrezas necesarias en el manejo teórico-práctico de la teoría de la probabilidad y las variables aleatorias que permitan su desarrollo profesional en el ámbito científico/académico, tecnológico o profesional especializado y multidisciplinar.

[CE5] Profundizar en los conocimientos en los fundamentos teórico-prácticos especializados del modelado y estudio de distintos tipos de relaciones de dependencia entre variables estadísticas.

[CE6] Adquirir conocimientos teórico-prácticos avanzados de distintas técnicas matemáticas, orientadas específicamente a la ayuda en la toma de decisiones, y desarrollar la capacidad de reflexión para evaluar y decidir entre distintas perspectivas en contextos complejos.

[CE8] Adquirir conocimientos teórico-prácticos avanzados de las técnicas destinadas a la realización de inferencias y contrastes relativos a variables y parámetros de un modelo estadístico, y saber aplicarlos con autonomía suficiente un contexto científico, tecnológico o profesional.

## **METODOLOGÍA DOCENTE: ACTIVIDADES DE APRENDIZAJE Y SU VALORACIÓN EN CRÉDITOS ECTS**

1.4 ECTS se dedicarán a docencia presencial (35 horas). De estas 35 horas, 25 se dedicarán a lección expositiva por parte del profesor, 8 a la realización de prácticas de ordenador para la ejecución de procedimientos de contraste de bondad de ajuste con datos reales y simulados, y 2 a la exposición de ejercicios por parte de los alumnos. Los 3.6 ECTS restantes serán no presenciales. De ellos, 1.6 ECTS estaría destinado al aprendizaje de los contenidos metodológicos, 1 ECTS a la resolución de los problemas prácticos, el manejo del software (esencialmente R) y la revisión bibliográfica, y 1 ECTS a la elaboración de los trabajos propuestos, que son parte del proceso de evaluación.

## **CRITERIOS Y MÉTODOS DE EVALUACIÓN**

La evaluación se realizará de forma **continua** a través de la entrega de trabajos y ejercicios a lo largo del curso y de exposiciones orales. Las ponderaciones de las distintas actividades de evaluación son las siguientes:

- Primera entrega de ejercicios: 25%
- Segunda entrega de ejercicios: 25%
- Primera exposición oral: 10%
- Segunda exposición oral: 40%

El alumnado que opte por la **evaluación global** realizará un examen teórico práctico sobre los contenidos de la materia con el que podrán optar por el 100% de la calificación.

Las actividades propuestas para la evaluación serán de distinta naturaleza, de forma que permitan evaluar las distintas competencias a desarrollar en la materia:

- Resolución de problemas teóricos del ámbito de los contrastes de especificación que permitirán evaluar las competencias CB7, CB9, CG4, CE4 y CE8.
- Aplicaciones de los contrastes de especificación a problemas reales, las cuales permitirán evaluar las competencias CB6, CB7, CB8, CB9, CG1, CG2, CG5, CT2, CT3, CE1, CE3, CE4, CE5 y CE8.
- Diseño y análisis de estudios de Monte Carlo, los cuales permitirán evaluar las competencias CB6, CB8, CB9, CG2, CG3, CG5, CT1, CT2, CT3, CE6 y CE8.
- Estudio de publicaciones recientes sobre contrastes de especificación, lo que permitirá evaluar las competencias CB8, CB9, CB10, CG3, CT1, CT2, CT3, CT4, CT5, CE4, CE6 y CE8.
- Exposiciones orales individuales y en grupo que permitirán evaluar las competencias CB9, CT1, CT2, CT4 y CT5.

## **TIEMPO DE ESTUDIO Y DE TRABAJO PERSONAL QUE DEBE DEDICAR UN ESTUDIANTE PARA SUPERAR LA MATERIA**

Docencia presencial: 35 horas (25 horas de lección expositiva por parte del profesor, 8 horas de prácticas, 2 horas de tutorización).

Estudio y trabajo personal: 90 horas.

Total: 125 horas (5 ECTS)

### **RECOMENDACIONES PARA EL ESTUDIO DE LA MATERIA**

Conviene acudir a esta materia con conocimientos medios de cálculo de probabilidades e inferencia estadística, con especial énfasis en los métodos de regresión, en la estimación de curvas y en los métodos de remuestreo. También es recomendable disponer de unas habilidades medias en el manejo de ordenadores, y en concreto de lenguajes de programación y software estadístico (esencialmente R). Para un mejor aprendizaje de la materia, conviene tener presente una clasificación básica de los múltiples métodos de contraste, un conocimiento detallado de algunos métodos fundamentales y mucha flexibilidad para la asimilación de métodos novedosos.

### **RECURSOS PARA EL APRENDIZAJE**

Esta materia requiere de aula ordinaria dotada de pizarra y retroproyector de transparencias, así como de otra aula dotada de ordenadores, en las cuales habrá lenguajes de programación y el software estadístico adecuado para la ejecución de procedimientos de contraste de especificación.

### **OBSERVACIONES**

Para los casos de realización fraudulenta de ejercicios o pruebas, será de aplicación lo recogido en las respectivas normativas de las universidades participantes en el Máster en Técnicas Estadísticas.

Esta guía y los criterios y metodologías en ella descritos están sujetos a las modificaciones que se deriven de normativas y directrices de las universidades participantes en el Máster en Técnicas Estadísticas.