

Módulo / Curso REG
**Modelos de Regresión para respuestas continuas,
categóricas, de recuento y de supervivencia**

Guía del curso 2023/24

Contenidos

Presentación.....	2
Requisitos de acceso	2
Elementos necesarios para realizar el curso	2
El Campus Virtual: soporte Internet del curso.....	2
Guía del Campus virtual.....	2
Foro	3
Pestaña <i>Evaluaciones</i>	3
Material de trabajo	3
Cálculos	3
Organización	3
Registro de erratas.....	4
Trabajo en equipo	4
Stages.....	4
Pruebas de evaluación continuada (PEC)	4
Calificación final.....	4
Programa calendario	5
Bloque REa: Regresión lineal múltiple con predictores cuantitativos y categóricos.....	5
Bloque REb: Regresión logística binaria, multinomial y de Poisson.....	6
Bloque REc: Análisis de la supervivencia y modelo de regresión de Cox.....	6
Advertencias importantes	6
Secretaría de los estudios	6
Contenidos del módulo.....	7
Bloque REa. Regresión lineal múltiple con predictores cuantitativos y categóricos.....	7
Bloque REb. Regresión logística binaria, multinomial y de Poisson.....	9
Bloque REc. Análisis de la supervivencia y modelo de regresión de Cox.....	11

Presentación

Este módulo a distancia está formado por tres bloques que proporcionan una formación aplicada en modelos de regresión. El conjunto de contenidos tiene una duración de 7 meses y exige 375 horas de dedicación que corresponden a 15 créditos ECTS. Es la **continuación** de los módulos iniciales *Fundamentos de Diseño y Estadística (FDE)* y *Proceso de datos sanitarios con Stata (PDS)*.

Se enseña de forma práctica a construir con *Stata*, y a interpretar, modelos de regresión lineal múltiple (bloque REa), de regresión logística, ordinal, multinomial y para recuentos (bloque REb) y de riesgos proporcionales (Cox) para datos de supervivencia (bloque REc). La finalidad de estos modelos puede ser tanto *predictiva* como de *ajuste* para medir el efecto de una exposición en presencia de variables de *confusión* y de variables *modificadoras*.

Requisitos de acceso

Es necesario tener los conocimientos del módulo FDE: Fundamentos de Diseño y Estadística y del módulo PDS: Proceso de datos sanitarios con *Stata* porque todas las pruebas de evaluación se hacen exclusivamente con *Stata*.

Elementos necesarios para realizar el curso

Los alumnos deben tener acceso a un PC conectado a Internet, con *Adobe Acrobat Reader* actualizado, disponer de correo electrónico y **leerlo diariamente** para estar al corriente de la información que se va enviando. Para seguir el curso es imprescindible **acceder regularmente** al Foro del Campus Virtual y leer **todos** los mensajes.

El módulo **requiere** saber trabajar de forma ágil con *Stata* y tener acceso a **Stata 15** o superior, porque los ejercicios y **todas las evaluaciones se deben realizar con Stata**.

El Campus Virtual: soporte Internet del curso

Todos los cursos de postgrado de los estudios en **Metodología de la investigación: Diseño y Estadística en Ciencias de la Salud** incorporan recursos de Internet para dotar a los alumnos de canales de comunicación y colaboración que eliminan las distancias físicas y facilitan la interacción con el profesorado.

El acceso a todos los recursos docentes y de evaluación está en: <https://metodo.uab.cat>. Una vez cargada la página principal de la Web debe hacer clic en **Campus Virtual** que es la puerta de entrada para alumnos.

Seguidamente se solicita su identificación. Por ejemplo, si su DNI es 01.234.567-Y, o su NIE es X01.234.567-Y, deberá escribir **sin dejar ningún blanco**:

	Usuario: 01234567Y	←	Escribir su DNI con la letra final en MAYÚSCULA.
	Contraseña: 01234567	←	Escribir el número del DNI sin la letra final .
<i>Ejemplo con NIE</i>	Usuario: X01234567Y	←	Escribir su NIE con las letras en MAYÚSCULA.
	Contraseña: X01234567	←	Escribir su NIE sin la letra final .

Si su identificación es correcta entrará en una página web con un panel izquierdo con varias opciones:

- “Menú”: Permite acceder a diferentes opciones relacionadas con aspectos generales de funcionamiento como consultar/modificar sus datos personales, su expediente académico, los envíos de material realizados, los comunicados generales sobre los estudios, etc.
- “Cursos”: Presenta la lista de módulos / cursos en que esté matriculado. Seleccionando el curso deseado entrará directamente al Foro de debate del campus virtual de dicho del curso.
- “Libros”: Presenta la lista de libros *online* a que tiene acceso, separando los del año actual y de años previos. Aunque el funcionamiento de los libros *online* es muy intuitivo, se explica detalladamente en otro documento junto al resto de Recursos digitales a que tiene acceso como alumno de la UAB.

Guía del Campus virtual

La parte superior de la página Web del curso contiene pestañas con enlaces a diferentes recursos; acceda a la pestaña **Guías y Normas**, descargue el documento PDF con la **Guía del Campus virtual**. Es conveniente imprimirla e **imprescindible leerla detenidamente antes de iniciar el curso**.

Foro

El **Foro** de debate es un espacio interno de la Web en el que se almacenan las preguntas y respuestas de los alumnos y profesores exclusivamente sobre los **contenidos científicos** del curso que se desean aclarar o ampliar, y también informaciones relacionadas con los contenidos del curso que puedan ser de interés general. Se debe **acceder regularmente y leer todos los mensajes** ya que es imprescindible para seguir el curso. Este acceso regular es especialmente importante en los períodos en que se realizan las pruebas de evaluación continuada (PEC).

El objetivo de este foro es que, en una primera instancia, sean los participantes quienes resuelvan las dudas a sus colegas. El Foro tiene asignado un profesor que actúa de moderador y que interviene si el grupo llega a conclusiones erróneas o si ningún alumno sabe contestar la pregunta. El profesor también podrá ofrecer, a través de este foro, información adicional sobre contenidos del curso.

El Foro recoge de forma ordenada los dos tipos de mensajes que se envían a lo largo del curso: *Mensaje nuevo* y *Responder* a un mensaje. Para que el foro de debate sea de utilidad es necesario **seguir de forma estricta** las reglas de conducta que se detallan en la **Guía del Campus virtual**.

Pestaña Evaluaciones

Esta pestaña abre una página Web con las pestañas **Autoevaluaciones** y **PEC (Pruebas de evaluación continuada)** que permiten gestionar las evaluaciones del curso.

En la **Guía del Campus virtual** se encuentra explicada la manera de contestar las autoevaluaciones y la prueba de evaluación continuada, y cómo obtener la solución comentada de las mismas.

Material de trabajo

El material básico está formado por 3 libros de texto que recibirá impresos en papel y que también se pueden consultar *online*:

Doménech JM, Navarro JB. Regresión múltiple con predictores categóricos y cuantitativos.
Barcelona: Graunt 21; 2023.

Doménech JM, Navarro JB. Regresión logística binaria, multinomial, Poisson y binomial negativa.
Barcelona: Graunt 21; 2023.

Doménech JM, Navarro JB. Análisis de la supervivencia y modelo de riesgos proporcionales de Cox.
Barcelona: Graunt 21; 2023.

El material complementario para realizar los cursos que forman este módulo se descarga en formato PDF desde la pestaña **Material** del Campus virtual.

Cálculos

Todos los análisis estadísticos se deben realizar con los comandos de *Stata*. Para operaciones sencillas, que deben tener la máxima precisión, se recomienda usar preferentemente el comando `display` de *Stata* aunque como segunda opción se puede usar Excel o, para operaciones elementales, la calculadora del sistema operativo Windows o Mac. Es recomendable leer el documento *Uso de calculadoras y reglas de redondeo - Cómo evitar errores de exactitud en los cálculos* que se descarga de la pestaña **Material** del Campus virtual.

Organización

El libro de texto de cada curso está formado por Unidades Didácticas (UD). Para preparar cada UD dispone, en general, de dos semanas. En el Programa calendario (pág. 5) se indican las fechas propuestas para enviar la autoevaluación de cada Unidad; estas fechas sólo son indicativas y su objetivo es orientar el ritmo de trabajo a lo largo del curso. Cada Unidad Didáctica comporta realizar los siguientes pasos:

- a) **Teoría.** Estudiar la parte teórica que se encuentra al inicio de la Unidad didáctica.
- b) **Ejercicios.** Realizar los ejercicios que se proponen durante la lectura de la parte teórica. La solución de estos ejercicios se encuentra en un Apéndice al final del libro.
- c) **Autoevaluación.** Contestar la prueba de autoevaluación que se encuentra al final de la Unidad didáctica marcando las respuestas en el libro de texto. Si no conoce la respuesta es mejor dejarla en blanco. Se trata de ejercicios de **autocomprobación** que tienen como objetivos básicos:
 1. Ayudar a la organización temporal de los estudios.
 2. Determinar el grado de comprensión de los contenidos.
 3. Ayudar a preparar las pruebas finales.

Es importante ir realizando estos tests. El **calendario facilitado es meramente orientativo** y su objetivo es ayudar a planificar temporalmente el curso, pero cualquier alumno puede organizar su propio calendario, enviar las autoevaluaciones en fechas posteriores a las previstas y formular preguntas de cualquier Unidad en cualquier momento del curso. Sin embargo, nuestra experiencia indica que los alumnos que realizan las autoevaluaciones de forma regular acostumbran a obtener los mejores resultados finales.

d) Envío respuestas autoevaluación. En la **Guía del Campus virtual** (disponible en la pestaña **Guías**) se explica cómo enviar las respuestas y cómo obtener la solución comentada de las mismas.

e) Comprobar respuestas. Debe comprobar sus respuestas con la solución comentada de la prueba y repasar las preguntas mal contestadas. Si después de este repaso le queda alguna duda pregunte al Foro.

Registro de erratas

Si encuentra algún elemento de los documentos que crea que pueda ser una errata, conviene comunicarlo **enviando un mensaje al Foro**. Si la errata es menor, le agradeceremos que nos lo comunique a través de la pestaña **Incidencias** del Campus virtual. Su colaboración nos será de gran utilidad para la confección de la **Fe de erratas** que está permanentemente disponible y actualizada en el enlace **Material** del Campus Virtual.

Trabajo en equipo

Aunque el esfuerzo personal es imprescindible, la posibilidad de trabajar las Unidades didácticas con algún colega facilita la comprensión y el seguimiento regular del curso.

Sin embargo, la prueba de evaluación continuada (PEC) **no debe** realizarse en colaboración o con ayuda de otras personas porque su resultado conduce a una **nota individual**. No obstante, se podrán discutir aspectos de la PEC a través del Foro del Campus Virtual.

Stages

El *stage* es una **conferencia de síntesis** para ayudar a consolidar los conceptos difíciles de entender con la única lectura de los textos y necesarios para el resto de los cursos. El primer stage resume los Bloques 1 y 2, y el segundo resume el bloque 3. En el Campus virtual encontrará la ciudad (Barcelona o Madrid) y fecha asignada para asistir a los *stages* y el enlace para confirmar su asistencia.

Asistir es muy recomendable, incluso para los alumnos que no han podido seguir todas las Unidades, porque proporciona una visión general de los temas tratados.

Pruebas de evaluación continuada (PEC)

Cada bloque del curso finaliza con la **Prueba de evaluación continuada (PEC)**. El presente módulo incluye 3 PEC's. A partir del día fijado en el calendario del curso (pág. 5) puede acceder al enlace **Evaluaciones** del Campus Virtual y en la pestaña **Pruebas de evaluación continuada** encontrará el enlace para descargar el formulario PDF con la PEC correspondiente.

Las respuestas se introducen en el mismo formulario PDF, cuya primera página contiene las instrucciones para rellenarlo y el **Compromiso de Honor** en el que declara que la prueba de evaluación ha sido realizada exclusivamente por usted sin la ayuda de otras personas.

Durante el tiempo fijado en el calendario del curso para realizar la PEC se pueden formular preguntas al Foro para aclarar dudas y el resto de alumnos pueden contestarlas, pero el profesor sólo interviene si hay alguna errata en el enunciado.

A diferencia de las Autoevaluaciones que permiten flexibilidad en las fechas de respuesta, la fecha límite de entrega de la PEC es la indicada en el calendario del curso (pág. 5) ya que **no es posible enviar la prueba una vez publicada la solución**. Se debe tener en cuenta que no se programan pruebas de recuperación.

Cuando la PEC está corregida se envía un e-mail para informar que ya puede descargar un pdf con la nota y las preguntas mal contestadas. También se descarga el pdf con la solución completa.

¡Atención! Durante el periodo en que realice una PEC **es imprescindible consultar regularmente el Campus Virtual** ya que el redactado de la PEC puede contener erratas o preguntas imprecisas que los profesores corrigen, modifican o aclaran **sólo** en el Campus.

Calificación final

La PEC de cada bloque representa 1/3 de la nota final sobre 10 (3.33 puntos). Además de las notas obtenidas en las PECs, la proporción de intervenciones realizadas en el Campus **en pestañas que no correspondan a la PEC** permite aumentar hasta 0.6 puntos, y la proporción de mensajes leídos permite aumentar otros 0.4, puntos la nota final de un curso.

a) **Cálculo de la proporción de intervenciones.** Si #Int es el total de intervenciones realizadas por el alumno y #mInt el total de intervenciones del alumno que más ha realizado, la proporción de intervenciones se calcula con la función max aplicando la siguiente fórmula: $pInterven = \#Int / \max(10; \#mInt)$

b) **Cálculo nota módulo:** La nota final del módulo resulta de la expresión (acotada a 10 puntos):

$$\text{NotaMod} = \text{PEC1} + \text{PEC2} + \text{PEC3} + 0.6 \times pInterven + 0.4 \times pMensajes$$

c) Para superar el módulo **NotaMod** debe ser **igual o superior a 5 puntos** y ninguna de las notas de las tres PEC debe ser inferior a 3 puntos.

Una vez realizada y superada la evaluación final se envía un Documento, firmado por el Director de los estudios, con el programa del curso, los créditos obtenidos y el número de horas mínimas que se estiman necesarias para seguir el curso. Este documento acreditativo es aceptado por la mayor parte de organismos.

Adicionalmente, si se ha matriculado a Diplomatura o Máster podrá solicitar el título oficial de sus estudios una vez completados los créditos necesarios. Si se ha matriculado como Curso de especialización recibirá el correspondiente Diploma de Experto.

En caso de **no aprobar el módulo**, para seguir con los estudios se deberá matricular de nuevo, y volver a pagar las tasas de inscripción y, en caso de tener otros módulos superados (por ejemplo los módulos FDE y PDS de primer año), las tasas de convalidación (el importe de ambas tasas es alrededor de 62 €), porque se cambiará la edición de la diplomatura / máster matriculado.

Como este módulo está formado por varios cursos, conservaremos internamente las calificaciones de los cursos aprobados, que no deberán ser evaluados de nuevo. Por ejemplo, si no se supera el módulo, no se aprueba la PEC del curso REa pero sí las PEC de los cursos REb y REc, en el futuro solo debería realizar las PEC del curso REa, y conservaríamos la nota obtenida en REb y REc.

Es más, si quisiera podría volver a evaluarse de REb y REc aun teniéndolos superados, y nos quedaríamos con la mejor de las notas obtenidas entre este curso y el que viene. Por ejemplo, si en el presente curso en REb obtiene un 6 y el curso que viene se vuelve a examinar, si obtiene peor nota mantenemos el 6, y si obtiene mejor nota mantenemos esta mejor nota.

Programa calendario

Junto a cada Unidad, entre paréntesis, figura el período de tiempo previsto para desarrollarla; el último día de cada período coincide con la fecha límite prevista de respuesta de la prueba de autoevaluación.

¡Atención! Este calendario es orientativo; su objetivo es ayudar a planificar temporalmente el curso, pero el alumno puede organizar su propio calendario, enviar las autoevaluaciones en fechas posteriores y formular preguntas sobre cualquier Unidad en cualquier momento.

Las fechas de disponibilidad y entrega de las PECs pueden variar ligeramente. Por este motivo, a medida que están disponibles se envía un *eMail*. La fecha límite de entrega de la PEC es **a las 24:00 horas del día que aparece junto al enlace para descargarla**. Si surge cualquier problema en la entrega envíe un correo a través del enlace Incidencias, pero tenga presente que **no es posible enviar la prueba una vez publicada la solución**.

Las fechas de los *stages* (y la del examen final) son **fijas** para poderlas bloquear en su agenda. Se asignan según el lugar de residencia, pero puede cambiarla para realizar el *stage* en la otra ciudad.

Bloque REa: Regresión lineal múltiple con predictores cuantitativos y categóricos

UD1 (30 Oct - 5 Nov 2023)

Regresión lineal simple

UD2 (6 - 19 Nov 2023)

Regresión lineal múltiple con predictores continuos y categóricos.

UD3 (20 Nov - 3 Dic 2023)

Diagnósticos del modelo

UD4 (4 - 17 Dic 2023)

Modelo de regresión lineal múltiple con fines predictivos: selección de la mejor ecuación

UD5 (18 Dic - 31 Dic 2023)

Modelo de regresión lineal múltiple para evaluar un efecto en presencia de confusión e interacción

Repaso (1 - 7 Ene 2024)

PEC1 (8 - 21 Ene 2024): Prueba de Evaluación Continuada

Bloque REb: Regresión logística binaria, multinomial y de Poisson

UD1 (22 - 28 Ene 2024)

El modelo de regresión logística

UD2 (29 Ene - 11 Feb 2024)

Construcción de un modelo de regresión logística

UD3 (12 - 25 Feb 2024)

Regresión logística multinomial y ordinal.

UD4 (26 Feb - 10 Mar 2024)

Regresión de Poisson y binomial negativa

Repaso (11 - 17 Mar 2024)

PRIMER STAGE: 11 Marzo 2024 de 10 a 15h (Barcelona) y 15 Marzo 2024 de 10 a 15h (Madrid)

PEC2 (18 – 31 Mar 2024): Prueba de Evaluación Continuada

Bloque REc: Análisis de la supervivencia y modelo de regresión de Cox

UD1 (1 - 14 Abr 2024)

Análisis de la supervivencia: métodos de Kaplan-Meier y Actuarial

UD2 (15 Abr - 5 May 2024)

Modelos para datos de supervivencia: Modelo de riesgos proporcionales de Cox

UD3 (6 - 19 May 2024)

Análisis de eventos recurrentes y riesgos competitivos

SEGUNDO STAGE: 17 Mayo 2024 de 10 a 14h (Madrid) y 24 Mayo 2024 de 10 a 14h (Barcelona)

Repaso (20 - 26 May 2024)

PEC3 (27 May - 9 Jun 2024): Prueba de Evaluación Continuada

Advertencias importantes

1. Los títulos de las Unidades didácticas de cada curso y su contenido puede que no correspondan exactamente con los que se impartirán porque están en permanente actualización y a criterio de los profesores pueden ser retocados antes de iniciar el curso.
2. Las fechas indicadas en el programa calendario pueden sufrir pequeñas modificaciones en función del ritmo del curso.
3. Es imprescindible acceder regularmente al Foro y leer todos los mensajes ya que, a través del Foro, los profesores aclaran conceptos y comunican posibles erratas u omisiones que puedan contener los libros y, en especial, las pruebas de evaluación continuada.
4. Se ha desarrollado un protocolo para detectar aquellas pruebas de evaluación continuada con probabilidad prácticamente nula de haber sido realizadas de forma independiente, lo que implica una violación del Compromiso de Honor. En este caso, para todos los alumnos implicados la calificación del curso será suspenso, con opción a realizar un examen *presencial* de reevaluación en el campus la UAB, que permite obtener como máximo un 5 de nota final.
5. Debe descargar de <https://metodo.uab.cat/docs/Advertencias.pdf> y leer el documento que contiene las advertencias más importantes que se asumen cuando se matricula en estos estudios de postgrado.

Secretaría de los estudios

Laboratori d'Estadística Aplicada

Universitat Autònoma de Barcelona

Edificio B – Carrer de la Fortuna

08193 Bellaterra (Barcelona)

España

Teléfono: 935 811 632 (9 a 14 horas)

Operador Web: 935 813 141 (9 a 14 horas)

Web: <https://metodo.uab.cat>

eMail: leam@uab.cat

Mensajes WhatsApp: 659 379 593

Contenidos del módulo

Bloque REa. Regresión lineal múltiple con predictores cuantitativos y categóricos

UD 1: Regresión lineal simple

1 Regresión y correlación

- 1.1 Presentación
- 1.2 Regresión y correlación
- 1.3 Estudio de la asociación lineal entre dos variables cuantitativas
- 1.4 Ajuste de una recta a una nube de puntos
 - 1.4.1 Ecuación de una recta
 - 1.4.2 Criterios de ajuste de una recta
 - 1.4.3 Ajuste de una recta con el criterio de mínimos cuadrados

2 Descripción de la relación entre dos variables con una recta de regresión

- 2.1 Estudio de la relación entre ingesta de alcohol y tiempo de reacción
- 2.2 Partición de la suma de cuadrados
- 2.3 Valoración del ajuste de la recta: coeficiente de determinación r^2
- 2.4 Interpretación de la recta de regresión ajustada
- 2.5 Mejora de ajuste

3 El modelo lineal de regresión

- 3.1 Especificación del modelo
- 3.2 Estimación de los parámetros del modelo
- 3.3 Prueba de significación de los coeficientes de regresión
- 3.4 Prueba de significación global: análisis de la variancia de la regresión
- 3.5 Comando regress
- 3.6 Comprobación de los supuestos del modelo
- 3.7 Aplicaciones del modelo de regresión
 - 3.7.1 Uso del modelo para medir un efecto
 - 3.7.2 Uso del modelo para realizar predicciones
 - 3.7.3 Predicciones con Stata
- 3.8 Generalización del modelo para predictores categóricos: la prueba t es un caso particular de la regresión lineal

UD 2: Regresión lineal múltiple

1 Modelo de regresión múltiple

- 1.1 Regresión múltiple
- 1.2 Ingesta de alcohol y tiempo de reacción
- 1.3 Preparación del modelo: exploración de las relaciones entre las variables
- 1.4 Especificación y estimación de un modelo de regresión múltiple
 - 1.4.1 Condiciones de aplicación
- 1.5 Estudio de la significación del modelo
 - 1.5.1 Análisis de las variaciones explicadas por cada variable
 - 1.5.2 Intervalos de confianza de los coeficientes de regresión
- 1.6 Regresión múltiple y ajuste estadístico
- 1.7 Coeficientes de correlación semiparcial y parcial

2 Codificación de variables categóricas

- 2.1 Variables predictoras categóricas en un modelo de regresión
- 2.2 Codificación de variables binarias
- 2.3 Codificación de variables con k categorías
 - 2.3.1 Codificación respecto a una categoría de referencia
 - 2.3.2 Codificación respecto al promedio de categorías
 - 2.3.3 Codificación en categorías anidadas
 - 2.3.4 Codificación con polinomios ortogonales
 - 2.3.5 Factor variables

UD 3: Diagnósticos del modelo

1 Diagnósticos de un modelo de regresión

- 1.1 Presentación
- 1.2 Identificación de valores alejados y de valores influyentes sobre las estimaciones de los parámetros del modelo
 - 1.2.1 Detección de valores alejados de la variable respuesta Y
 - 1.2.2 Detección de valores alejados de las variables predictoras X
 - 1.2.3 Detección de los valores alejados que afectan las estimaciones de los parámetros del modelo: valores influyentes
- 1.3 Comprobación del supuesto de normalidad
- 1.4 Comprobación de la linealidad y de la homogeneidad de variancias: examen de los gráficos de residuales
- 1.5 Comprobación del supuesto de independencia
- 1.6 Detección de la colinealidad

2 Transformaciones de variables

- 2.1 Representación gráfica de las funciones más usuales
- 2.2 Efecto de las transformaciones

3 Descripción y estimación de valores desconocidos

- 3.1 Clasificación de los valores desconocidos
- 3.2 Programas para analizar e imputar valores desconocidos
- 3.3 Comandos para analizar valores desconocidos
- 3.4 Alternativas para abordar la presencia de datos desconocidos
 - 3.4.1 Exclusión de los sujetos con algún valor desconocido
 - 3.4.2 Estimación de cada elemento de la matriz de variancias-covariancias a partir de todos los pares de valores válidos
 - 3.4.3 Estimación de cada valor desconocido y sustitución por el valor estimado
 - 3.4.3.1 Imputación del valor medio
 - 3.4.3.2 Imputación múltiple
- 3.5 Otras dimensiones de la calidad de los datos

UD 4: Construcción de un modelo con fines predictivos

1 Construcción de un modelo de regresión múltiple

- 1.1 Presentación de un estudio sobre obesidad
- 1.2 Especificación del modelo
- 1.3 Criterios de selección
- 1.4 Valores desconocidos en las variables predictoras
 - 1.4.1 Comparación de los sujetos incluidos y excluidos del modelo

2 Selección de variables

- 2.1 Selección de variables por pasos
 - 2.1.1 Inclusión secuencial (Forward selection)
 - 2.1.2 Exclusión secuencial (Backward elimination)
 - 2.1.3 Regresión paso a paso (Forward Stepwise y Backward Stepwise)
 - 2.1.4 Inconvenientes de los procedimientos automáticos paso a paso
- 2.2 Selección del mejor modelo guiada por el investigador
- 2.3 Selección del mejor modelo a partir de todas las posibles ecuaciones
 - 2.3.1 Presentación de un estudio sobre riesgo de bacteriemia
 - 2.3.2 Cribado de los predictores
 - 2.3.2.1 Cribado con regresiones simples
 - 2.3.2.2 Cribado con métodos de regresión por pasos
- 2.4 Estimación del modelo final

3 Validación del modelo

- 3.1 Valoración externa de la fiabilidad del modelo
- 3.2 Valoración interna de la fiabilidad del modelo
- 3.3 Recomendaciones importantes para construir un modelo

4 Predicciones con el modelo

- 4.1 Predicción de una respuesta. Índice pronóstico
- 4.2 Predicciones con Stata
 - 4.2.1 Generación de casos con patrones de valores de predictores
- 4.3 Problemática de la edad como variable predictora

UD 5: Construcción de un modelo para evaluar un efecto

1 Confusión e interacción

- 1.1 Presentación
- 1.2 Confusión
- 1.3 Interacción: variables modificadoras del efecto
- 1.4 Variables a registrar para controlar la confusión

2 Modelado estadístico de la confusión e interacción

- 2.1 Control estadístico de variables de confusión
 - 2.1.1 Criterios para el ajuste
- 2.2 Modelado estadístico de la interacción
 - 2.2.1 Criterio de interacción basado en la magnitud del cambio del efecto
 - 2.2.2 Interacciones de orden superior

3 Regresión múltiple como técnica de control

- 3.1 Estudio sobre dieta sin sal y presión arterial
- 3.2 Estimación de un efecto en presencia de confusión y de interacción
 - 3.2.1 Valoración de la interacción a partir del cambio producido en los efectos estimados
 - 3.2.2 Cálculo con ordenador de los intervalos de confianza
 - 3.2.2.1 Procedimiento para variables modificadoras continuas
 - 3.2.2.2 Procedimiento para variables modificadoras binarias
 - 3.2.2.3 Procedimiento para variables modificadoras categóricas

4 Medida de un efecto en presencia de confusión y de interacción

- 4.1 Modelado estadístico en presencia de confusión e interacción
- 4.2 Estudio sobre tabaco y presión arterial
- 4.3 Construcción de un modelo de regresión múltiple para estimar un efecto
- 4.4 Especificación de las variables de confusión y de interacción
- 4.5 Evaluación de los términos de interacción
- 4.6 Evaluación de los términos de confusión
 - 4.6.1 Evaluación de la confusión con el comando de usuario confound

- 4.7 Estimación del efecto de la exposición con el modelo finalista
- 4.8 Diagnósticos del modelo
- 4.9 Conclusión final

Apéndice: Codificación de variables categóricas

Bloque REb. Regresión logística binaria, multinomial y de Poisson

UD 1: El modelo de regresión logística

1 Modelo de regresión logística

- 1.1 Presentación
- 1.2 Función logística: ejemplo sobre hipertensión y consumo de tabaco
- 1.3 El modelo de regresión logística
- 1.4 Estimación de los parámetros del modelo
 - 1.4.1 Bondad de ajuste del modelo estimado
 - 1.4.1.1 Significación global del modelo: prueba de la razón de verosimilitud
 - 1.4.2 Significación de los parámetros del modelo
 - 1.4.3 Calibración del modelo: prueba de bondad de ajuste de Hosmer y Lemeshow
- 1.5 Interpretación de los parámetros estimados del modelo
- 1.6 Predicciones con el modelo
 - 1.6.1 Predicciones con el comando lincom
 - 1.6.2 Clasificación de sujetos
- 1.7 Índices de validez del modelo
 - 1.7.1 Curva ROC

2 Regresión logística con variables predictoras categóricas

- 2.1 Factor variables
- 2.2 Codificación de variables predictoras categóricas
- 2.3 Variable predictora binaria
 - 2.3.1 Análisis de una tabla de contingencia de 2x2
- 2.4 Variable predictora con k categorías
 - 2.4.1 Análisis de una tabla de contingencia de 2xk
- 2.5 Variable predictora con k categorías ordenadas
 - 2.5.1 Codificación con polinomios ortogonales
 - 2.5.2 Análisis dosis-respuesta

UD 2: Construcción de un modelo de regresión logística

1 Construcción de un modelo para medir un efecto en presencia de confusión y de interacción

- 1.1 Presentación
- 1.2 Confusión
- 1.3 Interacción: variables modificadoras del efecto
- 1.4 Variables a registrar para controlar la confusión
- 1.5 Estudio sobre aspirina e infarto
- 1.6 Estrategia para construir un modelo de regresión logística
 - 1.6.1 Especificación de las variables de confusión y de las modificadoras
- 1.7 Evaluación de los términos de interacción
 - 1.7.1 Evaluación de los diferentes modelos finalistas
- 1.8 Evaluación de los términos de confusión
 - 1.8.1 Estrategia de ajuste de la confusión
 - 1.8.2 Evaluación de los términos de confusión en un modelo con interacciones
 - 1.8.3 Evaluación de los términos de confusión en base al sesgo en un modelo sin interacciones
 - 1.8.4 Evaluación de la confusión con el comando de usuario confound
- 1.9 Estimación del riesgo relativo en lugar de la odds ratio

2 Proceso de construcción de un modelo predictivo de regresión logística

- 2.1 Elección de las potenciales variables a incluir en el modelo
- 2.2 Proceso de selección de las variables
 - 2.2.1 Selección del mejor modelo a partir de todas las posibles ecuaciones
 - 2.2.1.1 Comparación de curvas ROC
 - 2.2.2 Selección de variables por pasos
 - 2.2.3 Valoración de los modelos obtenidos
- 2.3 Interpretación de un modelo ajustado
 - 2.3.1 Interpretación de la constante
 - 2.3.2 Predicciones
 - 2.3.3 Interpretación de los coeficientes b
- 2.4 Uso del modelo de regresión logística para efectuar predicciones
 - 2.4.1 Cálculo de un índice pronóstico indicador del riesgo
 - 2.4.2 Construcción con Stata de una tabla de índices pronósticos
- 2.5 Valoración de la fiabilidad del modelo
 - 2.5.1 Valoración externa de la fiabilidad
 - 2.5.2 Valoración interna de la fiabilidad
 - 2.5.3 Estimación del modelo

3 Regresión logística condicional

- 3.1 Presentación de un estudio
- 3.2 Análisis mediante regresión logística condicional
 - 3.2.1 Incorporación de variables de confusión y modificadoras

4 Diagnóstico de un modelo de regresión logística

- 4.1 Supuestos de modelo
- 4.2 Detección de valores alejados de la variable de respuesta Y
- 4.3 Detección de valores alejados de las variables predictoras X
- 4.4 Detección de observaciones influyentes
- 4.5 Integración de diagnósticos
- 4.6 Evaluación de la linealidad
- 4.7 Detección de la colinealidad
- 4.8 El problema de la sobredispersión

UD 3: Regresión logística multinomial y ordinal

1 Regresión logística multinomial

- 1.1 Estudio sobre atención asistencial de los pacientes de EPOC
- 1.2 Asociación entre el hábitat de residencia y el tipo de asistencia recibida
- 1.3 El modelo con un predictor binario
- 1.4 Método de estimación de parámetros
- 1.5 Ajuste del modelo estimado
- 1.6 Significación global del modelo
- 1.7 Significación de los parámetros del modelo
- 1.8 Interpretación de los parámetros estimados
- 1.9 Predicciones con el modelo
- 1.10 Predictor categórico con más de 2 categorías
- 1.11 Predictor cuantitativo
- 1.12 Modelo con múltiples predictores
 - 1.12.1 Ajuste del modelo
 - 1.12.2 Interpretación del modelo seleccionado
 - 1.12.3 Tabla con las categorías reales y predichas
 - 1.12.4 Predicciones con el comando lincom

2 Regresión logística ordinal

- 2.1 Ejemplo sobre percepción de salud de los pacientes de EPOC
- 2.2 El modelo de odds proporcionales
- 2.3 El modelo con un predictor binario
- 2.4 Interpretación de los parámetros estimados del modelo
- 2.5 Estimación, bondad de ajuste y significación del modelo
- 2.6 Predicciones con el modelo
- 2.7 Predictor categórico con más de 2 categorías
- 2.8 Predictor cuantitativo

UD 4: Regresión de Poisson y binomial negativa

1 Regresión de Poisson

- 1.1 Ejemplo sobre número de hospitalizaciones evitables en los pacientes de EPOC
- 1.2 El modelo con un predictor cuantitativo
 - 1.2.1 Indicador de exposición
 - 1.2.2 Estimación del modelo
- 1.3 Interpretación de los parámetros estimados del modelo
- 1.4 Estimación, bondad de ajuste y significación del modelo
- 1.5 El problema de la sobredispersión
- 1.6 Predicciones con el modelo
- 1.7 Modelo con un predictor binario
- 1.8 Predictor categórico con más de 2 categorías

2 Modelos para recuentos con sobredispersión

- 2.1 Estudio sobre el número de hospitalizaciones evitables de los pacientes de EPOC
 - 2.2 El modelo de Poisson con el error estándar corregido
 - 2.3 El modelo de regresión binomial negativa
 - 2.3.1 Predicciones con el modelo
 - 2.4 Análisis de los residuales de desviación
 - 2.5 El modelo de regresión de Poisson con ceros inflados
 - 2.6 El modelo de regresión binomial negativa con ceros inflados
 - 2.7 Selección del mejor modelo en presencia de sobredispersión
- Apéndice:** Cálculo matricial del intervalo de predicción de las odds y proporciones predichas

Bloque REc. Análisis de la supervivencia y modelo de regresión de Cox

UD 1: Análisis de la supervivencia

1 Estimación de la función de supervivencia

- 1.1 Introducción
- 1.2 Terminología básica
- 1.3 Estudio sobre deshabitación de heroinómanos
- 1.4 Función de supervivencia
 - 1.4.1 Preparación de los datos para el análisis
 - 1.4.2 Método de Kaplan-Meier
 - 1.4.2.1 Media, mediana y cuartiles
 - 1.4.2.2 Vida residual mediana
 - 1.4.2.3 Error estándar e intervalo de confianza de las proporciones de supervivencia
 - 1.4.3 Método actuarial
 - 1.4.3.1 Cálculo de la mediana y de los cuartiles
 - 1.4.3.2 Vida residual mediana
- 1.5 Función de supervivencia y función de incidencia
 - 1.5.1 Función de incidencia
- 1.6 Calidad del seguimiento: pérdidas y duración

2 Comparación de curvas de supervivencia

- 2.1 Comparación de curvas de supervivencia
- 2.2 Comparación de dos proporciones de supervivencia
- 2.3 Prueba de Mantel-Cox (logrank test)
- 2.4 Razón de tasas
- 2.5 Comparación de varios grupos
- 2.6 Prueba de Mantel-Haenszel estratificada

UD 2: Modelos de regresión para datos de supervivencia: Regresión de Cox

1 Modelo de regresión de riesgos proporcionales

- 1.1 Presentación
- 1.2 Modelo de regresión de riesgos proporcionales de Cox
 - 1.2.1 Interpretación de los parámetros del modelo
 - 1.2.2 Condiciones de aplicación del modelo
 - 1.2.3 Introducción de variables categóricas en modelos de regresión de riesgos proporcionales: métodos de codificación
 - 1.2.4 Estimación de los parámetros del modelo
- 1.3 Modelo de regresión de Cox con una variable pronóstica binaria
 - 1.3.1 Contribución propia de cada variable predictora
- 1.4 Modelo de regresión de Cox con una variable pronóstica con k categorías
- 1.5 Modelo de regresión de Cox con una variable pronóstica con k categorías ordenadas
 - 1.5.1 Codificación con polinomios ortogonales
- 1.6 Modelo de regresión de Cox con una variable pronóstica cuantitativa
- 1.7 Capacidad predictiva: Índice C de Harrell

2 Construcción de un modelo para medir un efecto

- 2.1 Confusión e interacción
- 2.2 Construcción de un modelo de regresión de riesgos proporcionales para valorar un efecto
 - 2.2.1 Estimación de un efecto en presencia de interacción y confusión
 - 2.2.1.1 Evaluación de la confusión con el comando de usuario confound
 - 2.2.2 Cálculo del NNT con datos de supervivencia
 - 2.2.3 Comprobación de los supuestos del modelo

3 Construcción de un modelo con finalidad predictiva

- 3.1 Elección de las potenciales variables a incluir en el modelo
- 3.2 Proceso de selección de las variables
 - 3.2.1 Construcción del modelo a partir de todas las posibles ecuaciones
 - 3.2.2 Selección de variables por pasos
 - 3.2.3 Interpretación del modelo seleccionado
- 3.3 Uso del modelo de Cox para efectuar predicciones
 - 3.3.1 Índice pronóstico y tasa de riesgo
 - 3.3.2 Razón de riesgos
 - 3.3.3 Construcción de una tabla pronóstica de razones de riesgos
 - 3.3.4 Estimación de la probabilidad de supervivencia de un sujeto
- 3.4 Diagnósticos del modelo de regresión de riesgos proporcionales
 - 3.4.1 Supuesto de proporcionalidad
 - 3.4.1.1 Métodos gráficos
 - 3.4.1.2 Tests basados en la interacción predictor-tiempo
 - 3.4.1.3 Tests basados en los residuales de Schoenfeld
 - 3.4.1.4 Alternativas ante el incumplimiento del supuesto de proporcionalidad
 - 3.4.2 Supuesto de relación log-lineal
 - 3.4.3 Observaciones influyentes

4 Modelo de riesgos proporcionales con variables dependientes del tiempo

4.1 Presentación

4.2 Análisis del estudio sobre deshabituación de heroinómanos con variables que cambian durante el seguimiento

UD 3: Análisis de eventos recurrentes y de riesgos competitivos

1 Modelo de Cox para eventos recurrentes

1.1 Eventos recurrentes

1.1.1 Estudio sobre deshabituación de heroinómanos

1.2 Modelos de regresión de riesgos proporcionales para eventos recurrentes

1.2.1 Modelos de Proceso de recuento (Counting process)

1.2.2 Modelos de Riesgo condicional (Conditional risk)

1.2.2.1 Estimación del modelo CR1

1.2.3 Modelos de Riesgo marginal (Marginal Risk)

2 Modelo de riesgos competitivos

2.1 Presentación

2.2 Modelo de decrementos múltiples

2.3 Comparación de la incidencia acumulada entre grupos

2.4 Modelo de regresión de Fine-Gray

Apéndice A: Cálculo del tiempo transcurrido entre dos fechas con Stata

Apéndice B: Cálculo matricial del intervalo de confianza de la razón de riesgos

Apéndice C: Transformación de la matriz de datos para analizar eventos recurrentes