
TEMARIO

Modelos lineales generalizados

Descripción

En muchas áreas es de interés analizar la relación entre un conjunto de variables explicativas o inputs con una variable respuesta u output. Así, mismo puede ser relevante obtener predicciones del output a partir de los inputs. Los modelos lineales más conocidos para realizar estas tareas son los de regresión simple y múltiple; sin embargo, existen una gran variedad de modelos similares y el uso del modelo apropiado según nuestro problema puede ayudarnos a obtener mejores resultados. Estos corresponden a los Modelos Lineales Generalizados (MLG).

En regresión simple hay una sola variable explicativa, mientras que en múltiple hay un conjunto de variables explicativas, todas ellas cuantitativas. Cuando hay una sola variable explicativa y esta es categórica se estudia un ANOVA de un factor. Este modelo puede generalizarse y se obtienen modelos que se relacionan con diseños factoriales en experimentos. En diseño de experimentos se plantea el experimento apropiado para comprobar una hipótesis, por ejemplo, si un medicamento es efectivo. Finalmente, en el caso general es posible variables explicativas categóricas y continuas.

Por otra parte, cuando la variable respuesta corresponde a un conteo o se quiere modelar una tasa, la distribución asociada es Poisson o multinomial (regresión Poisson y modelos loglineales). Los modelos loglineales permiten además identificar relaciones de independencia marginal o condicional en un conjunto de variables categóricas (tablas de contingencia) y se ligan con teoría de gráficas.

Cuando la variable respuesta es categórica, de tal modo que la distribución asociada es por ejemplo Bernoulli, Binomial o multinomial, pueden usarse un MLG como por ejemplo la regresión logística. Esta permite construir un clasificador aún más general que en un análisis discriminante. Los MLG en su planteamiento más general, en los cuales los outputs tienen distribuciones asociadas a la familia exponencial, abarcan otros modelos, e. g. Probit, log-log, etc.

Los MLG son empleados para explicar y predecir en áreas muy diversas: Econometría, Bioestadística, Geografía, Finanzas y Seguros, etc. En este curso se introducen cada uno de los distintos tipos de modelos y se dan ejemplos de aplicaciones.

1. Introducción
 - a) Motivación y uso de los modelos lineales generalizados (MLG).
2. MLG con variable respuesta distribuida normalmente: Regresión lineal múltiple
 - a) Modelo de regresión lineal múltiple y supuestos del modelo
 - b) Ajuste por mínimos cuadrados ordinarios.
 - c) Propiedades de los estimadores.
 - d) Estimadores por máxima verosimilitud
 - e) Intervalos de confianza y pruebas de hipótesis.
 - f) Predicción.
 - g) Diagnósticos de los supuestos estadísticos.
 - h) Métodos de selección de variables.
 - i) Ejemplos prácticos usando paquetes de cómputo.
3. MLG con variable respuesta distribuida normalmente: Modelos con variables explicativas cuantitativas y cualitativas
 - a) Modelos tipo ANOVA de un solo factor como generalización de pruebas de diferencias de medias.
 - b) Supuestos y parametrización del modelo ANOVA de un solo factor.
 - c) Relación de los modelos ANOVA de un factor con Diseño de Experimentos.
 - d) Modelos con más de una variable explicativa discreta: Diseños factoriales e interacciones.
 - e) Modelos lineales generales: variables explicativas continuas (covariables) y discretas (factores).
 - f) Diagnósticos de los supuestos estadísticos.
 - g) Ejemplos prácticos usando paquetes de cómputo.
4. Modelos lineales cuya variable dependiente es un conteo
 - a) Modelos loglineales y tablas de contingencia
 - b) Modelos loglineales y su relación con modelos gráficos probabilísticos
 - c) Regresión Poisson
 - d) Obtención de parámetros e interpretación.
 - e) Bondad de ajuste y sobredispersión.
 - f) Selección de modelos.
 - g) Análisis residuales.
 - h) Ejemplos prácticos usando paquetes de cómputo.
5. Modelos lineales cuya variable dependiente es cualitativa.
 - a) Modelos logísticos.
 - b) Estimación de parámetros e interpretación.
 - c) Selección de modelos.
 - d) Regresión logística como modelo de clasificación.

- e) Regresión logística para varias categorías de respuesta (multinomial)
 - f) Bondad de ajuste y sobredispersión.
 - g) Otros modelos en respuestas binarias: probit y log-log
 - h) Ejemplos prácticos usando paquetes de cómputo.
 - i)
6. Familia exponencial. Planteamiento general de un MLG, ecuaciones de verosimilitud y su solución.

Informes: lumialearning@gmail.com