

Modelización de regresión de series temporales con R. XIV Summer School MESIO UPC-UB.

Ejercicio de evaluación

Este ejercicio de evaluación consiste en analizar la asociación entre temperatura y mortalidad cardiovascular en la ciudad de Chicago entre 1987 y 2000, utilizando los fundamentos de regresión de series temporales aprendidos durante el curso.

El alumno deberá enviar sus **respuestas en un archivo MSWord y un archivo script de R** con la sintaxis que haya utilizado para analizar los datos, **o alternativamente un archivo R markdown** que incluya las respuestas y la sintaxis de R conjuntamente.

Las respuestas se enviarán por email a aurelio.tobias@idaea.csic.es antes del **jueves 8 de julio de 2021**.

Los datos para realizar el ejercicio se encuentran en el *data frame* **chicagoNMMAPS** de la librería **dlnm**.

```
library(dlnm)
data <- chicagoNMMAPS
str(data)
'data.frame':5114 obs. of 14 variables:
 $ date : Date, format: "1987-01-01" "1987-01-02" "1987-01-03" ...
 $ time : int 1 2 3 4 5 6 7 8 9 10 ...
 $ year : num 1987 1987 1987 1987 1987 ...
 $ month: num 1 1 1 1 1 1 1 1 1 1 ...
 $ doy : int 1 2 3 4 5 6 7 8 9 10 ...
 $ dow : Factor w/ 7 levels "Sunday","Monday",...: 5 6 7 1 2 3 4 5 6 7 ...
 $ death: int 130 150 101 135 126 130 129 109 125 153 ...
 $ cvd : int 65 73 43 72 64 63 72 51 62 90 ...
 $ resp : int 13 14 11 7 12 12 12 13 7 11 ...
 $ temp : num -0.278 0.556 0.556 -1.667 0 ...
 $ dptp : num 31.5 29.9 27.4 28.6 28.9 ...
 $ rhum : num 95.5 88.2 89.5 84.5 74.5 ...
 $ pm10 : num 27 NA 32.8 40 NA ...
 $ o3 : num 4.38 4.93 3.75 4.29 4.75 ...
```

Pregunta 1. Realiza un análisis descriptivo de la mortalidad diaria por causas cardiovasculares (**cvd**)

- ¿Qué componentes de tendencia y estacionalidad observas?
 - ¿Existe alguna variación a corto plazo entre los días de la semana?
-

Pregunta 2. Calcula la sobredispersión y autocorrelación de la mortalidad diaria por causas cardiovasculares (**cvd**)

- ¿Cuanta sobredispersión tienen los datos de mortalidad?
 - ¿Crees que los datos de mortalidad se distribuyen independientemente?
-

Pregunta 3. Utiliza un modelo de regresión de Poisson para ajustar la estructura temporal de mortalidad diaria por causas cardiovasculares (**cvd**)

- ¿Qué modelización encuentras más adecuada: variables de calendario, funciones periódicas (y de qué orden), o funciones flexibles (y con cuantos grados de libertad)?
 - ¿Cuánto mejora la bondad de ajuste del modelo (qAIC), la sobredispersión y la autocorrelación residual con el modelo que has utilizado finalmente?
-

Pregunta 4. Ahora estudia los efectos a corto plazo, hasta 21 días, de la temperatura media (**temp**) sobre la mortalidad por causas cardiovasculares (**cvd**)?

- ¿Qué parametrización has utilizado para definir la **crossbasis** de temperatura?
 - ¿Cómo interpretas la asociación entre temperatura y mortalidad?
 - ¿Cuánto vale la temperatura de mínima mortalidad (**mmt**)?
-

Pregunta 5. A partir del modelo anterior, calcula el efecto del frío considerando en percentil del 5% de la distribución de la temperatura en comparación con el **mmt**, y el efecto del calor con el percentil del 95% de la distribución de temperatura en comparación con el **mmt**.

- ¿Cómo interpretas el efecto del frío? ¿y del calor?
 - ¿Hasta cuantos días se extiende el efecto de frío? ¿y del calor?
-

Preguntas Opcionales

Las siguientes preguntas del ejercicio son opcionales, para aquellos que quieran subir nota. Esta parte del ejercicio consiste en ampliar el análisis anterior evaluando la asociación entre contaminación atmosférica por partículas y mortalidad cardiovascular.

Pregunta 6. Considerando el modelo en el que ya has ajustando la estructura temporal de la serie de mortalidad y la temperatura, estudia ahora el efecto a corto plazo, hasta 7 días, de la contaminación atmosférica por partículas (**pm10**) .

- ¿Qué parametrización has utilizado ahora para definir la **crossbasis** de contaminación?
 - ¿Cómo interpretas la asociación entre contaminación y mortalidad?
 - ¿Hasta que día observas un efecto significativo de la contaminación sobre la mortalidad?
-

Pregunta 7. Por último, utiliza un diseño de casos cruzados con estratificación por tiempo para estudiar el efecto a corto plazo, hasta 7 días, de la contaminación atmosférica por partículas (**pm10**)

- ¿Cómo interpretas ahora la asociación entre contaminación y mortalidad?
 - ¿Existen diferencias con los resultados obtenidos utilizando regresión de series temporales?
-