

## Modelización de regresión de series temporales con R.

### XIV Summer School MESIO UPC-UB

# meta-analisis y dlnm

Carmen Iñíguez  
Depto. Estadística i I.O., UV



- 1 Introducción
- 2 Formulación
- 3 Aplicación con DLNM
- 4 Ejecución en **R**

# META-ANALISIS



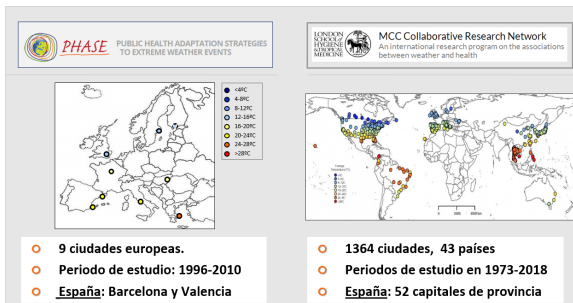
**Meta-análisis:** procedimiento estadístico para **combinar la evidencia** de  $m$  estudios independientes que examinan la misma hipótesis de investigación.

- Un meta-análisis tiene dos **objetivos**:
  - Obtener un estimador de efecto **promedio**, que tenga en cuenta la **precisión** de cada estudio.
  - Cuantificar el grado de **heterogeneidad** entre estudios y, si se dispone de información, explorar sus fuentes.
- Asociaciones complejas que dependen de varios coeficientes requieren una aproximación **multivariante**
- Herramienta clave en estudios de **temperatura y mortalidad**.

## RELACIÓN TEMPERATURA-MORTALIDAD



- Relación dependiente de la localización
- Evidencia científica basada en **estudios multi-céntricos**. Ejemplos:



- Obvia utilidad/necesidad de meta-análisis
- No linealidad y decalaje (**dlnm**), exige **meta-análisis multivariante**

## APARICIÓN DEL META-ANÁLISIS MULTIVARIANTE



- En origen surge para promediar resultados ya publicados: meta-análisis **bibliográfico**
  - Higgins PT & Green J. Cochrane Handbook for Systematic Reviews of Interventions. *John Wiley & Sons*, UK, 2008 .
- En los noventa se desarrolla como **metodología** de análisis para síntesis de estimadores.
  - DerSimonian R & Laird N. (1986). Meta-analysis in clinical trials. *Control Clin Trials*. **7** (3): 177–88.
  - Higgins JPT & Thompson SG. (1986). Quantifying heterogeneity in a meta-analysis. *Stat Med*. **21** (11): 1539–58.
- En la última década se desarrolla la versión multivariante y su integración en el **contexto dlrm**
  - Jackson D, et al. (2011) **Multivariate meta-analysis**: potential and promise. *Stat Med*. **30** (3): 2481–98.
  - Gasparrini et al (2012). Multivariate meta-analysis for non- linear and other multi-parameter associations. *Stat Med*. **31** (29): 3821–39.
  - Gasparrini A (2013)& Armstrong B. Reducing and **meta-analysing estimates from distributed lag non-linear models**. *BMC Med Res Methodol*. **13**: 1.
  - Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019) M. An extended **mixed-effects framework for meta-analysis**. *Stat Med*. **38**(29): 5429–44.

## META-ANÁLISIS. FORMULACIÓN.



- Supongamos una relación analizada en **m estudios** que viene determinada por **k coeficientes**:

$$\theta_i = (\theta_{i1}, \dots, \theta_{ik})^t, \quad i = 1, \dots, m$$

- El **modelo de meta-análisis multivariante**, como el univariante, tiene dos versiones:

- Modelo de **efectos fijos**

$$\hat{\theta}_i \sim N_k(\theta, S_i), \quad (1)$$

- Modelo de **efectos aleatorios**

$$\hat{\theta}_i | \mathbf{u}_i \sim N_k(\theta + \mathbf{u}_i, S_i), \quad \mathbf{u}_i \sim N_k(\mathbf{0}, \Psi), \quad (2)$$

dónde:

- $\theta$  vector  $k \times 1$  de efectos **combinados** que se desea estimar
- $\mathbf{u}_i$  vector  $k \times 1$  de efectos **aleatorios** para los coeficientes del estudio  $i$
- $S_i$  matrix  $k \times k$  de covarianzas **intra** estudio para el estudio  $i$
- $\Psi$  matrix  $k \times k$  de covarianzas **entre** estudios, completamente definida por  $k(k+1)/2$  variables  $\xi$ .

- En definitiva, modelo lineal de ef. fijos o aleatorios: Estimación por ML o REML

## META-REGRESIÓN MULTIVARIANTE



- La heterogeneidad entre ciudades en los coeficientes podría deberse a características estudio-específicas: variables sociodemográficas, económicas, climáticas, etc. (**meta-predictores**)

$$\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^t, \quad i = 1, \dots, m$$

- El modelo de meta-análisis puede ser extendido para modelizar la relación entre los coeficientes y **p predictores** a nivel de estudio: **meta-regresión**
- El **modelo de meta-regresión multivariante** de efectos aleatorios, sería:

$$\hat{\theta}_i \mid \mathbf{u}_i \sim \mathbf{N}_k(\mathbf{X}_i \boldsymbol{\beta} + \mathbf{u}_i, \mathbf{S}_i), \quad \mathbf{u}_i \sim \mathbf{N}_k(\mathbf{0}, \boldsymbol{\Psi}) \quad (3)$$

dónde:

- $\boldsymbol{\beta}$  vector  $kp \times 1$  de coeficientes a estimar  $i$
- $\mathbf{X}_i = \mathbf{I}_k \otimes \mathbf{x}_i^t$

$$\mathbf{X}_i = \begin{pmatrix} x_{i1} & x_{i2} & \cdots & x_{ip} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & x_{i1} & x_{i2} & \cdots & x_{ip} \end{pmatrix}_{k \times kp}$$

## MEDIDA DE HETEROGENEIDAD Y COMPARACIÓN



- ¿Existe Heterogeneidad?:  $H_0 : \Psi = \mathbf{0}$

Test de heterogeneidad basado en la  $Q$  de Cochran multivariante:

$$Q = \sum_{i=1}^m (\hat{\theta}_i - \mathbf{X}_i \hat{\beta})^t \mathbf{S}_i^{-1} (\hat{\theta}_i - \mathbf{X}_i \hat{\beta}) \sim \chi_{km-q}^2 \quad (4)$$

- $q$ :  $n^{\circ}$  parámetros de efectos fijos a estimar
- Crítica: Test conservador que depende del número de estudios y predictores.

- ¿Cuanta?: Índice  $I^2$  de Higgins:

$$Q = \frac{H^2 - 1}{H^2}, \quad H^2 = \max\{1, Q/(km - q)\} \quad (5)$$

- Interpretación: la proporción de variabilidad entre estudios no debida al azar
- $I^2 \in [0, 100]$  %, permite comparaciones entre modelos
- Crítica: puede verse afectada por la precisión de los estudios

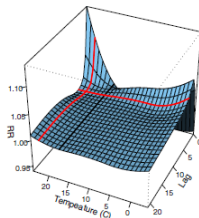
- ¿ $X_j$  explica heterogeneidad?:  $H_0 : \beta^j = 0$

- En definitiva, lme: AIC, BIC y el test de Wald multivariante aplican



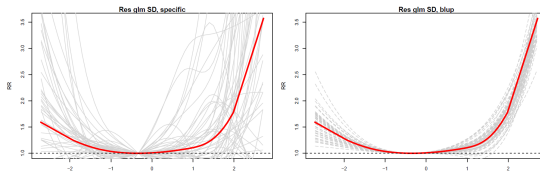
## APLICACIÓN EN EL CONTEXTO DLNM

- Meta-analizamos los coeficientes **eta** obtenidos en el dlrm. (coefs. de splines)
- Escasa interpretabilidad de los betas individuales para los meta-predictores. Hay que **predecir**.
- Para evitar problemas numéricos (acondicionamiento de  $\Sigma$ ) conviene trabajar con **bases reducidas** (una sola dimensión, dosis o lag).
- Los objetivos más comunes: curva dosis-respuesta overall y curva lag-respuesta para una exposición concreta.



## APLICACIÓN EN EL CONTEXTO DLNM

- Interés especial de la predicción **BLUP** (Best Linear Unbiased Prediction): suma de los parámetros combinados y de la desviación estudio específica.
- Los BLUPS permiten a las ciudades pequeñas o con NAs en las series aprender del conjunto (el medio camino entre originales y combinados)



- Idealmente los análisis dlnm debería realizarse con la misma crossbasis para que el promedio de coeficientes tenga sentido, obliga a veces a utilizar una **escala relativa**

Más matemáticas en Gasparrini et al, *BMC Med Res Methodol.* 13: 1

## META-ANÁLISIS MULTIVARIANTE EN R



- Guardamos coefs,  $\beta_i$ , y las matrices de varianza,  $S_i$ , de cada estudio obtenidos a partir de la base reducida que corresponda: **coef**, **vcov**
  - ① Objetivo: Overall  

```
red <- crossreduce(cb, model, at= tpred)
```
  - ② Objetivo: Distribución en lags  

```
red <- crossreduce(cb, model, type= "var", value= p95, cen= mmt)
```
- Se usará la función **mvmeta** (o **mixmeta**) sobre coeficientes y varianza :
  - ① Meta-análisis de efectos aleatorios\* :  

```
mv <- mvmeta(coef ~ 1, vcov, method="reml")
```
  - ② Meta regresión de efectos aleatorios\*:  

```
mv <- mvmeta(coef ~ metap1+ metap2+ ..., vcov, method="reml")
```
  - (\*) Efectos fijos si `method="ml"`
- Para obtener los blups basta aplicar la función **blup**:  

```
coefBlup <- blup(mv, vcov=T)
```

## META-ANÁLISIS MULTIVARIANTE EN R



- La predicción se obtiene utilizando la base que corresponda según el objetivo, en coherencia con `crossreduce`:

- Objetivo: Overall

```
bvar <- do.call("onebasis", c(list(x=tpred), attr(cb,"argvar")))
```

- Objetivo: Distribución en lags

```
blag <- do.call("onebasis", c(list(x=xlag), attr(cb,"arglag")))
```

- y mediante los comandos `crosspred` y `predict`

- Meta-análisis:

```
metapred<-crosspred(basis=bvar,coef=coef(mv),vcov=vcov(mv),at=tpred)
```

- Meta-regresión\*:

```
predict(mv, newdata= data.frame(metap10, metap20)vcov=T,  
interval="confidence")
```

- (\*) Sólo coefs, habría que aplicarlos a la base.