EPG3343 - Seminario de Estadística III

Clase 11

Jonathan Acosta

Pontificia Universidad Católica de Chile

Segundo Semestre, 2022

Esquema

1. Modelos Jerárquicos
2. Modelos Jerárquicos para Datos Geoestadísticos
   - Modelos de procesos espaciales estacionarios
   - Modelos Lineales Generalizados Espaciales
3. Modelos Jerárquicos para Datos de Área

Esquema

1. Modelos Jerárquicos
2. Modelos Jerárquicos para Datos Geoestadísticos
   - Modelos de procesos espaciales estacionarios
   - Modelos Lineales Generalizados Espaciales
3. Modelos Jerárquicos para Datos de Área

Modelos Jerárquicos

- Como hemos discutido en las clases previas, los modelos bayesianos combinan la información de los datos (a través de la verosimilitud) con el conocimiento externo (distribución a priori).

Modelos Jerárquicos

- Como hemos discutido en las clases previas, los modelos bayesianos combinan la información de los datos (a través de la verosimilitud) con el conocimiento externo (distribución a priori).
- Sea $\boldsymbol{y} = (y_1, \ldots, y_n)$ los datos observados, los cuales se asumen que provienen de $f(\boldsymbol{y} \mid \theta)$.

Modelos Jerárquicos

- Como hemos discutido en las clases previas, los modelos bayesianos combinan la información de los datos (a través de la verosimilitud) con el conocimiento externo (distribución a priori).

- Sea $\boldsymbol{y} = (y_1, \ldots, y_n)$ los datos observados, los cuales se asumen que provienen de $f(\boldsymbol{y} \mid \theta)$.

- $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ es un vector aleatorio desconocido, con distribución a priori $\pi(\boldsymbol{\theta} \mid \lambda)$, donde $\boldsymbol{\lambda}$ es un vector de hiperparámetros.

Modelos Jerárquicos

- Si $\boldsymbol{\lambda}$ es conocido, la inferencia sobre $\boldsymbol{\theta}$ esta basado en distribución posterior

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{\lambda}) = \frac{p(\boldsymbol{y}, \boldsymbol{\theta} \mid \boldsymbol{\lambda})}{p(\boldsymbol{y} \mid \boldsymbol{\lambda})} = \frac{f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})}{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})d\boldsymbol{\theta}}$$

Modelos Jerárquicos

- Si $\boldsymbol{\lambda}$ es conocido, la inferencia sobre $\boldsymbol{\theta}$ esta basado en distribución posterior

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{\lambda}) = \frac{p(\boldsymbol{y}, \boldsymbol{\theta} \mid \boldsymbol{\lambda}}{p(\boldsymbol{y} \mid \boldsymbol{\lambda}} = \frac{f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})}{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})d\boldsymbol{\theta}}$$

- Si $\boldsymbol{\lambda}$ es desconocido, a menudo, se utiliza otra distribución $h(\boldsymbol{\lambda})$, en este caso,

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y}, \boldsymbol{\theta}}{p(\boldsymbol{y}} = \frac{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})h(\boldsymbol{\lambda})d\boldsymbol{\lambda}}{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})h(\boldsymbol{\lambda})d\boldsymbol{\theta}d\boldsymbol{\lambda}}$$

Modelos Jerárquicos

- Si $\boldsymbol{\lambda}$ es conocido, la inferencia sobre $\boldsymbol{\theta}$ esta basado en distribución posterior

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{\lambda}) = \frac{p(\boldsymbol{y}, \boldsymbol{\theta} \mid \boldsymbol{\lambda}}{p(\boldsymbol{y} \mid \boldsymbol{\lambda}} = \frac{f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})}{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda}) d\boldsymbol{\theta}}$$

- Si $\boldsymbol{\lambda}$ es desconocido, a menudo, se utiliza otra distribución $h(\boldsymbol{\lambda})$, en este caso,

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y}, \boldsymbol{\theta}}{p(\boldsymbol{y}} = \frac{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})h(\boldsymbol{\lambda}) d\boldsymbol{\lambda}}{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})h(\boldsymbol{\lambda}) d\boldsymbol{\theta} d\boldsymbol{\lambda}}$$

- Notar la estructura jerárquica implícita, es decir, se requieren tres niveles de especificación distribucional.

Modelos Jerárquicos

- Si $\boldsymbol{\lambda}$ es conocido, la inferencia sobre $\boldsymbol{\theta}$ esta basado en distribución posterior

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{\lambda}) = \frac{p(\boldsymbol{y}, \boldsymbol{\theta} \mid \boldsymbol{\lambda}}{p(\boldsymbol{y} \mid \boldsymbol{\lambda}} = \frac{f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})}{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})d\boldsymbol{\theta}}$$

- Si $\boldsymbol{\lambda}$ es desconocido, a menudo, se utiliza otra distribución $h(\boldsymbol{\lambda})$, en este caso,

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y}, \boldsymbol{\theta}}{p(\boldsymbol{y}} = \frac{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})h(\boldsymbol{\lambda})d\boldsymbol{\lambda}}{\int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})h(\boldsymbol{\lambda})d\boldsymbol{\theta}d\boldsymbol{\lambda}}$$

- Notar la estructura jerárquica implícita, es decir, se requieren tres niveles de especificación distribucional.

- Típicamente con interés primordial en el nivel $\boldsymbol{\theta}$.

Modelos Jerárquicos

- En lugar de integrar sobre $\boldsymbol{\lambda}$, se puede reemplazar este valor por su estimador $\widehat{\boldsymbol{\lambda}}$ obtenido como

$$\widehat{\boldsymbol{\lambda}} = \operatorname{argmax} p(\boldsymbol{y} \mid \boldsymbol{\lambda}), \quad \text{donde} \quad p(\boldsymbol{y} \mid \boldsymbol{\lambda}) = \int f(\boldsymbol{y} \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda}) d\boldsymbol{\theta}$$

Modelos Jerárquicos

- En lugar de integrar sobre $\boldsymbol{\lambda}$, se puede reemplazar este valor por su estimador $\widehat{\boldsymbol{\lambda}}$ obtenido como

$$\widehat{\boldsymbol{\lambda}} = \operatorname{argmax} p(\boldsymbol{y} \mid \boldsymbol{\lambda}), \quad \text{donde} \quad p(\boldsymbol{y} \mid \boldsymbol{\lambda}) = \int f(\boldsymbol{y} \mid \boldsymbol{\theta}) \pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda}) d\boldsymbol{\theta}$$

- Luego la inferencia sobre $\boldsymbol{\theta}$ se realiza mediante la distribución posterior estimada, $p(\boldsymbol{\theta} \mid \boldsymbol{y}, \widehat{\boldsymbol{\lambda}})$.

Modelos Jerárquicos

- En lugar de integrar sobre $\boldsymbol{\lambda}$, se puede reemplazar este valor por su estimador $\widehat{\boldsymbol{\lambda}}$ obtenido como

$$\widehat{\boldsymbol{\lambda}} = \operatorname{argmax} p(\boldsymbol{y} \mid \boldsymbol{\lambda}), \quad \text{donde} \quad p(\boldsymbol{y} \mid \boldsymbol{\lambda}) = \int f(\boldsymbol{y} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{\lambda})d\boldsymbol{\theta}$$

- Luego la inferencia sobre $\boldsymbol{\theta}$ se realiza mediante la distribución posterior estimada, $p(\boldsymbol{\theta} \mid \boldsymbol{y}, \widehat{\boldsymbol{\lambda}})$.

- Este procedimiento se denomina Bayes empírico.

Esquema

1. Modelos Jerárquicos

2. Modelos Jerárquicos para Datos Geoestadísticos
   - Modelos de procesos espaciales estacionarios
   - Modelos Lineales Generalizados Espaciales

3. Modelos Jerárquicos para Datos de Área

- Considere el modelo

$$Y(\boldsymbol{s}) = \mu(\boldsymbol{s}) + Z(\boldsymbol{s}) + \epsilon(\boldsymbol{s})$$

donde la estructura de la media es $\mu(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$.

Modelos Jerárquicos para Datos Geoestadísticos

- Considere el modelo

$$Y(\boldsymbol{s}) = \mu(\boldsymbol{s}) + Z(\boldsymbol{s}) + \epsilon(\boldsymbol{s})$$

  donde la estructura de la media es $\mu(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$.

- Los residuos son particionados en; una parte espacial, $Z(\boldsymbol{s})$, y una parte no espacial, $\epsilon(\boldsymbol{s})$.

- Considere el modelo

$$Y(\boldsymbol{s}) = \mu(\boldsymbol{s}) + Z(\boldsymbol{s}) + \epsilon(\boldsymbol{s})$$

  donde la estructura de la media es $\mu(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$.

- Los residuos son particionados en; una parte espacial, $Z(\boldsymbol{s})$, y una parte no espacial, $\epsilon(\boldsymbol{s})$.

- Así, el residuo $Z(\boldsymbol{s})$ se asume que es una realización del campo aleatorio Gaussiano estacionario, de media $\mathbf{0}$ y función de covarianza isotrópica, $C(h)$.

Modelos Jerárquicos para Datos Geoestadísticos

- Considere el modelo

$$Y(\boldsymbol{s}) = \mu(\boldsymbol{s}) + Z(\boldsymbol{s}) + \epsilon(\boldsymbol{s})$$

  donde la estructura de la media es $\mu(\boldsymbol{s}) = \boldsymbol{X}(\boldsymbol{s})\boldsymbol{\beta}$.

- Los residuos son particionados en; una parte espacial, $Z(\boldsymbol{s})$, y una parte no espacial, $\epsilon(\boldsymbol{s})$.

- Así, el residuo $Z(\boldsymbol{s})$ se asume que es una realización del campo aleatorio Gaussiano estacionario, de media $\mathbf{0}$ y función de covarianza isotrópica, $C(h)$.

- Mientras que $\epsilon(\boldsymbol{s})$ es un término de error puro no correlacionado.

- Se asumirá que $\mathbb{V}[Z(\boldsymbol{s})] = \sigma^2$ denominada sill parcial, y $\mathbb{V}[\epsilon(\boldsymbol{s})] = \tau^2$ (efecto nugget).

Modelos Jerárquicos para Datos Geoestadísticos

- Se asumirá que $\mathbb{V}[Z(\boldsymbol{s})] = \sigma^2$ denominada sill parcial, y $\mathbb{V}[\epsilon(\boldsymbol{s})] = \tau^2$ (efecto nugget).
- Notar que $[Z(\boldsymbol{s}+\boldsymbol{h}) - Z(\boldsymbol{s})] \to 0$ cuando $\boldsymbol{h} \to \boldsymbol{0}$, pero $[Z(\boldsymbol{s}+\boldsymbol{h}) + \epsilon(\boldsymbol{s}+\boldsymbol{h}) - Z(\boldsymbol{s}) - \epsilon(\boldsymbol{s})]$ no tiende a 0.

Modelos Jerárquicos para Datos Geoestadísticos

- Se asumirá que $\mathbb{V}[Z(\boldsymbol{s})] = \sigma^2$ denominada sill parcial, y $\mathbb{V}[\epsilon(\boldsymbol{s})] = \tau^2$ (efecto nugget).

- Notar que $[Z(\boldsymbol{s}+\boldsymbol{h}) - Z(\boldsymbol{s})] \to 0$ cuando $\boldsymbol{h} \to \boldsymbol{0}$, pero $[Z(\boldsymbol{s}+\boldsymbol{h}) + \epsilon(\boldsymbol{s}+\boldsymbol{h}) - Z(\boldsymbol{s}) - \epsilon(\boldsymbol{s})]$ no tiende a 0.

- Suponga que tiene los datos $Y(\boldsymbol{s}_i)$, $i = 1, \ldots, n$ y sea $\boldsymbol{Y} = (Y(\boldsymbol{s}_1), \ldots,$ el vector de datos observados.

Modelos Jerárquicos para Datos Geoestadísticos

- Se asumirá que $\mathbb{V}[Z(\boldsymbol{s})] = \sigma^2$ denominada sill parcial, y $\mathbb{V}[\epsilon(\boldsymbol{s})] = \tau^2$ (efecto nugget).

- Notar que $[Z(\boldsymbol{s}+\boldsymbol{h}) - Z(\boldsymbol{s})] \to 0$ cuando $\boldsymbol{h} \to \boldsymbol{0}$, pero $[Z(\boldsymbol{s}+\boldsymbol{h}) + \epsilon(\boldsymbol{s}+\boldsymbol{h}) - Z(\boldsymbol{s}) - \epsilon(\boldsymbol{s})]$ no tiende a 0.

- Suponga que tiene los datos $Y(\boldsymbol{s}_i)$, $i = 1, \ldots, n$ y sea $\boldsymbol{Y} = (Y(\boldsymbol{s}_1), \ldots,$ el vector de datos observados.

- Para este conjunto de datos se tiene media y matriz de varianzas dadas por

$$\boldsymbol{\mu} = \boldsymbol{X\beta}; \quad \wedge \quad \boldsymbol{\Sigma} = \sigma^2 \boldsymbol{R}(\phi) + \tau^2 \boldsymbol{I} \tag{1}$$

donde $\boldsymbol{R}(\phi)$ es la matriz de correlación dada por algún modelo de correlación parametrizado por ϕ.

Modelos Jerárquicos para Datos Geoestadísticos

- Si $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \tau^2, \phi)$, una solución Bayesiana requiere una distribución a priori apropiada $p(\boldsymbol{\theta})$.

- Si $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \tau^2, \phi)$, una solución Bayesiana requiere una distribución a priori apropiada $p(\boldsymbol{\theta})$.

- Así, se tiene que

$$\boldsymbol{Y} \mid \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{R}(\boldsymbol{\phi}) + \tau^2 \boldsymbol{I})$$

Modelos Jerárquicos para Datos Geoestadísticos

- Si $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \tau^2, \phi)$, una solución Bayesiana requiere una distribución a priori apropiada $p(\boldsymbol{\theta})$.

- Así, se tiene que

$$\boldsymbol{Y} \mid \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{R}(\boldsymbol{\phi}) + \tau^2 \boldsymbol{I})$$

- Típicamente, prioris independientes son elegidas, tales como

$$p(\boldsymbol{\theta}) = p(\boldsymbol{\beta}) p(\sigma^2) p(\phi)$$

Modelos Jerárquicos para Datos Geoestadísticos

- Si $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \tau^2, \phi)$, una solución Bayesiana requiere una distribución a priori apropiada $p(\boldsymbol{\theta})$.

- Así, se tiene que

$$\boldsymbol{Y} \mid \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{R}(\boldsymbol{\phi}) + \tau^2 \boldsymbol{I})$$

- Típicamente, prioris independientes son elegidas, tales como

$$p(\boldsymbol{\theta}) = p(\boldsymbol{\beta})p(\sigma^2)p(\phi)$$

- Candidatos naturales para $\boldsymbol{\beta}$ corresponde a una Normal Multivariada, mientras que gamma inversa lo son para σ^2 y τ^2.

- La especificación de la priori de ϕ depende del modelo de correlación utilizado. Por ejemplo, para la función de correlación exponencial una priori Gamma podría ser adecuada.

Modelos Jerárquicos para Datos Geoestadísticos

- La especificación de la priori de ϕ depende del modelo de correlación utilizado. Por ejemplo, para la función de correlación exponencial una priori Gamma podría ser adecuada.

- En el caso sin nugget, $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{R}(\phi)$, se debe tener aún más cuidado. Ya que Zhang (2004) prueba que solo el producto $\sigma^2\phi^{2\nu}$ puede ser identificable, para el modelo de covarianza Matérn con parámetro de suavidad ν.

Modelos Jerárquicos para Datos Geoestadísticos

- La especificación de la priori de ϕ depende del modelo de correlación utilizado. Por ejemplo, para la función de correlación exponencial una priori Gamma podría ser adecuada.

- En el caso sin nugget, $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{R}(\phi)$, se debe tener aún más cuidado. Ya que Zhang (2004) prueba que solo el producto $\sigma^2\phi^{2\nu}$ puede ser identificable, para el modelo de covarianza Matérn con parámetro de suavidad ν.

- Para el modelo exponencial solo se puede identificar $\sigma^2\phi$ y no el rango o la varianza por separado. Es decir, solo si se fija uno el otro puede ser identificado.

Modelos Jerárquicos para Datos Geoestadísticos

- La especificación de la priori de ϕ depende del modelo de correlación utilizado. Por ejemplo, para la función de correlación exponencial una priori Gamma podría ser adecuada.

- En el caso sin nugget, $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{R}(\phi)$, se debe tener aún más cuidado. Ya que Zhang (2004) prueba que solo el producto $\sigma^2\phi^{2\nu}$ puede ser identificable, para el modelo de covarianza Matérn con parámetro de suavidad ν.

- Para el modelo exponencial solo se puede identificar $\sigma^2\phi$ y no el rango o la varianza por separado. Es decir, solo si se fija uno el otro puede ser identificado.

- Banerjee et al. (2015) recomienda una priori muy informativa para ϕ y relativamente imprecisa para σ^2. Para la primera, en lugar de la distribución Gamma, recomiendan emplear una uniforme sobre un intervalo. determinado

- A menudo el objetivo es realizar la inferencia sobre los parámetros individualmente, para esto se requiere obtener la distribución marginal posterior.

Modelos Jerárquicos para Datos Geoestadísticos

- A menudo el objetivo es realizar la inferencia sobre los parámetros individualmente, para esto se requiere obtener la distribución marginal posterior.

- Por ejemplo, para $\boldsymbol{\beta}$, se tiene

$$
\begin{aligned}
p(\boldsymbol{\beta} \mid \boldsymbol{y}) &= \int\int\int p(\boldsymbol{\beta}, \sigma^2, \tau^2, \phi \mid \boldsymbol{y}) d\sigma^2 d\tau^2 d\phi \\
&\propto p(\boldsymbol{\beta}) \int\int\int f(\boldsymbol{y} \mid \boldsymbol{\theta}) p(\sigma^2) p(\tau^2) p(\phi) d\sigma^2 d\tau^2 d\phi
\end{aligned}
$$

Modelos Jerárquicos para Datos Geoestadísticos

- A menudo el objetivo es realizar la inferencia sobre los parámetros individualmente, para esto se requiere obtener la distribución marginal posterior.

- Por ejemplo, para $\boldsymbol{\beta}$, se tiene

$$
\begin{aligned}
p(\boldsymbol{\beta} \mid \boldsymbol{y}) &= \int\int\int p(\boldsymbol{\beta}, \sigma^2, \tau^2, \phi \mid \boldsymbol{y}) d\sigma^2 d\tau^2 d\phi \\
&\propto p(\boldsymbol{\beta}) \int\int\int f(\boldsymbol{y} \mid \boldsymbol{\theta}) p(\sigma^2) p(\tau^2) p(\phi) d\sigma^2 d\tau^2 d\phi
\end{aligned}
$$

- La naturaleza jerárquica del modelamiento queda como;

  [data| process, parameters] [process| parameters] [parameters]

- Para visualizar el marco anterior, primero considere

$$\boldsymbol{Y} \mid \boldsymbol{\theta}, \boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \tau^2 \boldsymbol{I})$$

donde $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n)^\top$ es el vector de efectos aleatorios espaciales. Así, los $Y(\boldsymbol{s}_i)$ son condicionalmente independientes dado $w(\boldsymbol{s}_i)$.

Modelos Jerárquicos para Datos Geoestadísticos

- Para visualizar el marco anterior, primero considere

$$\boldsymbol{Y} \mid \boldsymbol{\theta}, \boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \tau^2 \boldsymbol{I})$$

  donde $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n)^\top$ es el vector de efectos aleatorios espaciales. Así, los $Y(\boldsymbol{s}_i)$ son condicionalmente independientes dado $w(\boldsymbol{s}_i)$.

- La especificación de la segunda etapa de $\boldsymbol{W}$ es

$$\boldsymbol{Z} \mid \sigma^2, \phi \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{R}(\phi)), \quad \text{Este es el proceso.}$$

Modelos Jerárquicos para Datos Geoestadísticos

- Para visualizar el marco anterior, primero considere

$$\boldsymbol{Y} \mid \boldsymbol{\theta}, \boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \tau^2 \boldsymbol{I})$$

  donde $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n)^\top$ es el vector de efectos aleatorios espaciales. Así, los $Y(\boldsymbol{s}_i)$ son condicionalmente independientes dado $w(\boldsymbol{s}_i)$.

- La especificación de la segunda etapa de $\boldsymbol{W}$ es

$$\boldsymbol{Z} \mid \sigma^2, \phi \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{R}(\phi)), \quad \text{Este es el proceso.}$$

- Finalmente, la especificación del modelo es completa añadiendo prioris para $\boldsymbol{\beta}$ y τ^2, así como para σ^2 y ϕ, estos últimos dos pueden ser considerados como hiperparámetros.

Modelos Jerárquicos para Datos Geoestadísticos

- Para visualizar el marco anterior, primero considere

$$\boldsymbol{Y} \mid \boldsymbol{\theta}, \boldsymbol{Z} \sim \mathcal{N}(\boldsymbol{X}\boldsymbol{\beta}, \tau^2 \boldsymbol{I})$$

  donde $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n)^\top$ es el vector de efectos aleatorios espaciales. Así, los $Y(\boldsymbol{s}_i)$ son condicionalmente independientes dado $w(\boldsymbol{s}_i)$.

- La especificación de la segunda etapa de $\boldsymbol{W}$ es

$$\boldsymbol{Z} \mid \sigma^2, \phi \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{R}(\phi)), \quad \text{Este es el proceso.}$$

- Finalmente, la especificación del modelo es completa añadiendo prioris para $\boldsymbol{\beta}$ y τ^2, así como para σ^2 y ϕ, estos últimos dos pueden ser considerados como hiperparámetros.

- Notar que el espacio paramétrico aumenta de $\boldsymbol{\theta}$ a $(\boldsymbol{\theta}, \boldsymbol{Z})$, siendo este creciente en n.

- Una vez establecido el modelo, el interés sigue en la predicción de Y en un sitio no observado $\boldsymbol{s}_0$, en el cual se conoce $\boldsymbol{x}_0 = \boldsymbol{x}(\boldsymbol{s}_0)$.

Modelos Jerárquicos para Datos Geoestadísticos

- Una vez establecido el modelo, el interés sigue en la predicción de Y en un sitio no observado $\boldsymbol{s}_0$, en el cual se conoce $\boldsymbol{x}_0 = \boldsymbol{x}(\boldsymbol{s}_0)$.
- Esta etapa predictiva se denomina kriging Bayesiano.

- Una vez establecido el modelo, el interés sigue en la predicción de Y en un sitio no observado $\boldsymbol{s}_0$, en el cual se conoce $\boldsymbol{x}_0 = \boldsymbol{x}(\boldsymbol{s}_0)$.
- Esta etapa predictiva se denomina kriging Bayesiano.
- Sea $Y_0 = Y(\boldsymbol{s}_0)$. La solución en el marco Bayesiano esta dada por la distribución predictiva

$$\begin{aligned} p(y_0 \mid \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}_0) &= \int p(y_0, \boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}_0) d\boldsymbol{\theta} \\ &= \int p(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{x}_0) p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{X}) d\boldsymbol{\theta} \end{aligned}$$

donde $p(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{x}_0)$ es una distribución condicional normal multivarida que surge desde la distribución conjunta (normal) de $\boldsymbol{Y}, \boldsymbol{Y}_0$.

Modelos Jerárquicos para Datos Geoestadísticos

- Los métodos de generación de números aleatorios pueden resolver el problema, ya que, si $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(G)} \sim p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{X})$, entonces, la integral anterior puede ser calculada mediante una mixtura de Monte-Carlo de la forma

$$\widehat{p}(y_0 \mid \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}_0) = \frac{1}{G} \sum_{g=1}^{G} p(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}^{(g)}, \boldsymbol{x}_0)$$

- Los métodos de generación de números aleatorios pueden resolver el problema, ya que, si $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(G)} \sim p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{X})$, entonces, la integral anterior puede ser calculada mediante una mixtura de Monte-Carlo de la forma

$$\widehat{p}(y_0 \mid \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}_0) = \frac{1}{G} \sum_{g=1}^{G} p(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}^{(g)}, \boldsymbol{x}_0)$$

- En la practica, se utiliza el muestreo de composición para extraer, uno por uno para cada $\boldsymbol{\theta}^{g}$, un $y_0^{(g)} \sim p\left(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}^{(g), \boldsymbol{x}_0}\right)$.

Modelos Jerárquicos para Datos Geoestadísticos

- Los métodos de generación de números aleatorios pueden resolver el problema, ya que, si $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(G)} \sim p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{X})$, entonces, la integral anterior puede ser calculada mediante una mixtura de Monte-Carlo de la forma

$$\widehat{p}(y_0 \mid \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}_0) = \frac{1}{G}\sum_{g=1}^{G} p(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}^{(g)}, \boldsymbol{x}_0)$$

- En la practica, se utiliza el muestreo de composición para extraer, uno por uno para cada $\boldsymbol{\theta}^{g}$, un $y_0^{(g)} \sim p\left(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}^{(g), \boldsymbol{x}_0}\right)$.

- La colección $\left\{y_0^{(1)}, y_0^{(2)}, \ldots, y_0^{(G)}\right\}$ es una muestra desde la densidad predictiva posterior. En consecuencia, se puede obtener un estimador puntual o intervalo de confianza para Y_0.

- Los métodos de generación de números aleatorios pueden resolver el problema, ya que, si $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(G)} \sim p(\boldsymbol{\theta} \mid \boldsymbol{y}, \boldsymbol{X})$, entonces, la integral anterior puede ser calculada mediante una mixtura de Monte-Carlo de la forma
$$\widehat{p}(y_0 \mid \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{x}_0) = \frac{1}{G} \sum_{g=1}^{G} p(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}^{(g)}, \boldsymbol{x}_0)$$
- En la practica, se utiliza el muestreo de composición para extraer, uno por uno para cada $\boldsymbol{\theta}^{g}$, un $y_0^{(g)} \sim p\left(y_0 \mid \boldsymbol{y}, \boldsymbol{\theta}^{(g), \boldsymbol{x}_0}\right)$.
- La colección $\left\{y_0^{(1)}, y_0^{(2)}, \ldots, y_0^{(G)}\right\}$ es una muestra desde la densidad predictiva posterior. En consecuencia, se puede obtener un estimador puntual o intervalo de confianza para Y_0.
- Por lo general, se requiere estimar Y en m sitios $\boldsymbol{s}_0$. El procedimiento es el mismo que el caso $m = 1$.

- El modelamiento geoestadístico clásico asume la Gaussianidad de la variable respuesta.

Modelos Lineales Generalizados Espaciales

- El modelamiento geoestadístico clásico asume la Gaussianidad de la variable respuesta.
- Este supuesto puede ser muy poco realista en algunos casos, como por ejemplo, para respuesta binaria.

Modelos Lineales Generalizados Espaciales

- El modelamiento geoestadístico clásico asume la Gaussianidad de la variable respuesta.
- Este supuesto puede ser muy poco realista en algunos casos, como por ejemplo, para respuesta binaria.
- El Generalised linear spatial model (GLSM) es presentado en Diggle et al. (1998), Zhang (2002) y Christensen and Waagepetersen (2002) como una extensión natural a variables respuesta que no pueden ser Normales.

Modelos Lineales Generalizados Espaciales

- El modelamiento geoestadístico clásico asume la Gaussianidad de la variable respuesta.
- Este supuesto puede ser muy poco realista en algunos casos, como por ejemplo, para respuesta binaria.
- El Generalised linear spatial model (GLSM) es presentado en Diggle et al. (1998), Zhang (2002) y Christensen and Waagepetersen (2002) como una extensión natural a variables respuesta que no pueden ser Normales.
- El GLSM es un modelo lineal mixto generalizado, donde el efecto aleatorio se obtiene desde un proceso espacial.

Modelos Lineales Generalizados Espaciales

Sea $\{Y(\boldsymbol{s}) : \ \boldsymbol{s} \in D\}$ la variable respuesta observada en $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$, y sea $\{Z(\boldsymbol{s}) : \ \boldsymbol{s} \in D\}$ un campo Gaussiano no observable con media 0 y función de correlación $\rho(h|\phi)$.

Modelos Lineales Generalizados Espaciales

Sea $\{Y(\boldsymbol{s}) : \ \boldsymbol{s} \in D\}$ la variable respuesta observada en $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$, y sea $\{Z(\boldsymbol{s}) : \ \boldsymbol{s} \in D\}$ un campo Gaussiano no observable con media 0 y función de correlación $\rho(h|\phi)$.

Supuesto: $Y(\boldsymbol{s})$ es un proceso condicionalmente independiente dado $\boldsymbol{\beta}$ y $Z(\boldsymbol{s}_i)$ con distribución

$$f(y(\boldsymbol{s}_i) \mid \boldsymbol{\beta}, Z(\boldsymbol{s}_i), \gamma) = h(y(\boldsymbol{s}_i), \gamma) \exp\left\{\gamma[y(\boldsymbol{s}_i)\eta(\boldsymbol{s}_i) - \psi(\eta(\boldsymbol{s}_i))]\right\},$$

donde $g(\eta) = \boldsymbol{x}^\top(\boldsymbol{s}_i)\boldsymbol{\beta} + Z(\boldsymbol{s}_i)$ para alguna función d enlace $g(\cdot)$, y γ es un parámetro de dispersión.

Sea $\{Y(\boldsymbol{s}) : \ \boldsymbol{s} \in D\}$ la variable respuesta observada en $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$, y sea $\{Z(\boldsymbol{s}) : \ \boldsymbol{s} \in D\}$ un campo Gaussiano no observable con media 0 y función de correlación $\rho(h|\phi)$.

Supuesto: $Y(\boldsymbol{s})$ es un proceso condicionalmente independiente dado $\boldsymbol{\beta}$ y $Z(\boldsymbol{s}_i)$ con distribución

$$f(y(\boldsymbol{s}_i) \mid \boldsymbol{\beta}, Z(\boldsymbol{s}_i), \gamma) = h(y(\boldsymbol{s}_i), \gamma) \exp\left\{\gamma[y(\boldsymbol{s}_i)\eta(\boldsymbol{s}_i) - \psi(\eta(\boldsymbol{s}_i))]\right\},$$

donde $g(\eta) = \boldsymbol{x}^\top(\boldsymbol{s}_i)\boldsymbol{\beta} + Z(\boldsymbol{s}_i)$ para alguna función d enlace $g(\cdot)$, y γ es un parámetro de dispersión.

Obs: Si $Z(\boldsymbol{s}_i)$ fueran iid tendríamos el tradicional generalized linear mixed effect model (Breslow ad Clayton, 1993).

- Utilizando la independencia condicional, lo que hemos hecho es crear una distribución conjunta

$$L(\boldsymbol{\beta}, \sigma^2, \phi, \gamma) = \int \prod_{i=1}^{n} f(y_i|\boldsymbol{\beta}, Z(\boldsymbol{s}_i), \gamma) p(\boldsymbol{z}|\sigma^2, \phi) d\boldsymbol{z}$$

- Utilizando la independencia condicional, lo que hemos hecho es crear una distribución conjunta

$$L(\boldsymbol{\beta}, \sigma^2, \phi, \gamma) = \int \prod_{i=1}^{n} f(y_i|\boldsymbol{\beta}, Z(\boldsymbol{s}_i), \gamma) p(\boldsymbol{z}|\sigma^2, \phi) d\boldsymbol{z}$$

- La clase de distribuciones que puede soportar un proceso estocástico es limitado, y caracterizado por la mixtura de distribuciones elípticas.

- La integral anterior es también la constante normalizadora en la distribución condicional

$$p(\boldsymbol{z}|\boldsymbol{y}, \boldsymbol{\beta}, \sigma^2, \phi) \propto \prod_{i=1}^{n} f(y_i|\boldsymbol{\beta}, Z(\boldsymbol{s}_i), \gamma) p(\boldsymbol{z}|\boldsymbol{\beta}, \sigma^2, \phi)$$

- La integral anterior es también la constante normalizadora en la distribución condicional

$$p(\boldsymbol{z}|\boldsymbol{y}, \boldsymbol{\beta}, \sigma^2, \phi) \propto \prod_{i=1}^{n} f(y_i|\boldsymbol{\beta}, Z(\boldsymbol{s}_i), \gamma) p(\boldsymbol{z}|\boldsymbol{\beta}, \sigma^2, \phi)$$

- Sea $\boldsymbol{Z} = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n))^\top$ el vector no observable del proceso y $Z_0 = Z(\boldsymbol{s}_0)$ el valor del campo en un sitio no observado $\boldsymbol{s}_0$.

Obsercaciones:

- En la práctica se utilizan las Cadenas de Markov Monte Carlo (MCMC) para resolver las integrales anteriores, con el objetivo de determinar la distribución predictiva $Z_0|\boldsymbol{y}, \boldsymbol{\beta}, \sigma^2, \phi$.

Obsercaciones:

- En la práctica se utilizan las Cadenas de Markov Monte Carlo (MCMC) para resolver las integrales anteriores, con el objetivo de determinar la distribución predictiva $Z_0|\boldsymbol{y}, \boldsymbol{\beta}, \sigma^2, \phi$.

- En **R** la libreria geoRglm permite realizar los cáculos anteriores.

Esquema

Modelos Jerárquicos para Datos de Área

- Cuando trabajamos con datos de área, la dependencia espacial se caracteriza por la estructura de vecinos.

Modelos Jerárquicos para Datos de Área

- Cuando trabajamos con datos de área, la dependencia espacial se caracteriza por la estructura de vecinos.
- En este caso $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$ simplemente se convierten en $\{1, 2, \ldots, n\}$ y denote por $\mathcal{V}(i)$ al vecindario de i.

Modelos Jerárquicos para Datos de Área

- Cuando trabajamos con datos de área, la dependencia espacial se caracteriza por la estructura de vecinos.
- En este caso $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$ simplemente se convierten en $\{1, 2, \ldots, n\}$ y denote por $\mathcal{V}(i)$ al vecindario de i.
- Bajo la propiedad de Markov el parámetro θ_i, para la i-ésima área is independiente de todos los otros parámetros, dado el conjunto de vevinos $\mathcal{V}(i)$, entonces

$$\theta_i \perp \boldsymbol{\theta}_{-i} \mid \boldsymbol{\theta}_{\mathcal{V}(i)} \quad \Longrightarrow \quad \theta_i \perp \theta_j \mid \boldsymbol{\theta}_{-(ij)} \;\Leftrightarrow\; Q_{ij} = 0$$

Es decir, los valores no cero en la mratiz de precisión están dados por la estructura del vecindario. $Q_{ij} \neq 0$ si y solo si $j \in \{i, \mathcal{V}(i)\}$.

Modelos Jerárquicos para Datos de Área

- Mapas de enfermedades son comúnmente usados con datos de área para evaluar los patrones espaciales de una particular enfermedad.

Modelos Jerárquicos para Datos de Área

- Mapas de enfermedades son comúnmente usados con datos de área para evaluar los patrones espaciales de una particular enfermedad.
- Los datos en este caso son discretos en los naturales.

Modelos Jerárquicos para Datos de Área

- Mapas de enfermedades son comúnmente usados con datos de área para evaluar los patrones espaciales de una particular enfermedad.
- Los datos en este caso son discretos en los naturales.
- Típicamente es de interés la tasa de mortalidad (o morbilidad) estandarizada (SMR), la cual es la razón entre el número de casos observados y_i y el número de casos esperados E_i en el área i-ésima.

Modelos Jerárquicos para Datos de Área

- Mapas de enfermedades son comúnmente usados con datos de área para evaluar los patrones espaciales de una particular enfermedad.
- Los datos en este caso son discretos en los naturales.
- Típicamente es de interés la tasa de mortalidad (o morbilidad) estandarizada (SMR), la cual es la razón entre el número de casos observados y_i y el número de casos esperados E_i en el área i-ésima.
- Se suele utilizar tasas de referencia estandarizadas, r_j, por edad y sexo (en general J grupos) y el censo:

$$r_j = \frac{\sum_{i=1}^n y_{ij}}{\sum_{i=1}^n n_{ij}},\ j = 1, \ldots, J; \qquad E_i = \sum_{j=1}^{J} n_{ij} * r_j$$

n_{ij} es el número de personas en riesgo del área i para el grupo j.

Modelos Jerárquicos para Datos de Área

- Asumiremos que los recuentos esperados ya están disponibles y se han calculado mediante el procedimiento anterior.

Modelos Jerárquicos para Datos de Área

- Asumiremos que los recuentos esperados ya están disponibles y se han calculado mediante el procedimiento anterior.
- Este enfoque simplista no tiene en cuenta la dependencia espacial entre las zonas.

Modelos Jerárquicos para Datos de Área

- Asumiremos que los recuentos esperados ya están disponibles y se han calculado mediante el procedimiento anterior.
- Este enfoque simplista no tiene en cuenta la dependencia espacial entre las zonas.
- Para tener en cuenta dicha información se asumirá modelo de Poisson, donde para cada zona suponemos:

$$y_i \sim Po(\lambda_i), \quad \lambda_i = E_i * \rho_i, \quad \log(\rho_i) = \eta_i$$

- En el caso clásico, el EMV de ρ_i es

$$\widehat{\rho}_i = SMR_i = \frac{y_i}{E_i}; \quad IC_{95\,\%} \approx SMR_i(e^{-2/\sqrt{y_i}}, e^{2/\sqrt{y_i}})$$

Modelos Jerárquicos para Datos de Área

- En el caso bayesiano sin dependencia espacial, un modelo que se suele emplear es el Poisson-Gamma, el cual considera

$$y_i \sim Po(\lambda_i), \quad \lambda_i = E_i * \rho_i, \quad \rho_i \sim \text{Gamma}(a, b)$$

donde la distribución Gamma es parametrizada tal que su esperanza $\mu = a/b$ y varianza $\sigma^2 = a/b^2$.

Modelos Jerárquicos para Datos de Área

- En el caso bayesiano sin dependencia espacial, un modelo que se suele emplear es el Poisson-Gamma, el cual considera

$$y_i \sim Po(\lambda_i), \quad \lambda_i = E_i * \rho_i, \quad \rho_i \sim \text{Gamma}(a, b)$$

  donde la distribución Gamma es parametrizada tal que su esperanza $\mu = a/b$ y varianza $\sigma^2 = a/b^2$.

- Se puede probar que

$$\rho_i \mid y_i \sim \text{Gamma}(y_i + a, E_i + b)$$

Modelos Jerárquicos para Datos de Área

- En el caso bayesiano sin dependencia espacial, un modelo que se suele emplear es el Poisson-Gamma, el cual considera

$$y_i \sim Po(\lambda_i), \quad \lambda_i = E_i * \rho_i, \quad \rho_i \sim \text{Gamma}(a, b)$$

donde la distribución Gamma es parametrizada tal que su esperanza $\mu = a/b$ y varianza $\sigma^2 = a/b^2$.

- Se puede probar que

$$\rho_i \mid y_i \sim \text{Gamma}(y_i + a, E_i + b)$$

- Por lo tanto, un estimador puntual para ρ_i es

$$\mathbb{E}[\rho_i \mid \boldsymbol{y}] = \mathbb{E}[\rho_i \mid y_i] = \frac{y_i + a}{E_i + b} = \alpha_i SMR_i + (1 - \alpha_i)\mu$$

donde $\alpha_i = E_i/(E_i + \mu/\sigma^2) \in [0, 1]$ para todo i.

- Un modelo log-lineal, puede ser especificado, para incorporar la dependencia espacial:

$$\eta_i = b_0 + u_i + v_i$$

Modelos Jerárquicos para Datos de Área

- Un modelo log-lineal, puede ser especificado, para incorporar la dependencia espacial:

$$\eta_i = b_0 + u_i + v_i$$

- b_0 cuantifica la tasa media de resultados en todo el estudio, v_i es el efecto específico de la zona modelado como intercambiable, mientras que u_i es otro efecto específico de la zona, que ahora modelamos espacialmente estructurado.

Modelos Jerárquicos para Datos de Área

- Un modelo log-lineal, puede ser especificado, para incorporar la dependencia espacial:

$$\eta_i = b_0 + u_i + v_i$$

- b_0 cuantifica la tasa media de resultados en todo el estudio, v_i es el efecto específico de la zona modelado como intercambiable, mientras que u_i es otro efecto específico de la zona, que ahora modelamos espacialmente estructurado.

- Sea $\boldsymbol{u} = (u_1, \ldots, u_n)$, donde varias estructuras de dependencia espacial pueden ser especificadas. Una particular es asumir que

$$u_i \mid \boldsymbol{u}_{-i} \sim \mathcal{N}(\mu_i + \sum_{j=1}^{n} r_{ij}(u_j - \mu_j), s_i^2)$$

donde μ_i es la media para cada área i, y $s_i^2 = \sigma_u^2 / \#\mathcal{V}(i)$.

Modelos Jerárquicos para Datos de Área

- Es decir, $\boldsymbol{u} = (u_1, \ldots, u_n)$ se modelo como un CAR(ϕ), es decir,

$$\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\mu}, (\boldsymbol{I} - \phi \boldsymbol{W})^{-1} \boldsymbol{S}^2), \quad \text{donde} \;\; S = \text{diag}\{s_1, \ldots, s_n\}$$

Modelos Jerárquicos para Datos de Área

- Es decir, $\boldsymbol{u} = (u_1, \ldots, u_n)$ se modelo como un CAR(ϕ), es decir,

$$\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\mu}, (\boldsymbol{I} - \phi \boldsymbol{W})^{-1} \boldsymbol{S}^2), \quad \text{donde} \;\; S = \text{diag}\{s_1, \ldots, s_n\}$$

- Debido a la dificultad para estimar ϕ, se suele utilizar $\phi = 1$, el cual es conocido como ICAR o IAR, resultando en una distribución impropia que puede ser rectificada imponiendo la restricción $\sum_{i=1}^{n} u_i = 0$.

Modelos Jerárquicos para Datos de Área

- Es decir, $\boldsymbol{u} = (u_1, \ldots, u_n)$ se modelo como un CAR(ϕ), es decir,

$$\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\mu}, (\boldsymbol{I} - \phi \boldsymbol{W})^{-1} \boldsymbol{S}^2), \quad \text{donde} \;\; S = \operatorname{diag}\{s_1, \ldots, s_n\}$$

- Debido a la dificultad para estimar ϕ, se suele utilizar $\phi = 1$, el cual es conocido como ICAR o IAR, resultando en una distribución impropia que puede ser rectificada imponiendo la restricción $\sum_{i=1}^{n} u_i = 0$.

- Además, se suele asumir que $\mu_i = 0$, para dar forma al modelo típicamente presentado en la literatura de mapas de enfermedades.

Modelos Jerárquicos para Datos de Área

- Es decir, $\boldsymbol{u} = (u_1, \ldots, u_n)$ se modelo como un CAR(ϕ), es decir,

$$\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\mu}, (\boldsymbol{I} - \phi \boldsymbol{W})^{-1}\boldsymbol{S}^2), \quad \text{donde} \;\; S = \text{diag}\{s_1, \ldots, s_n\}$$

- Debido a la dificultad para estimar ϕ, se suele utilizar $\phi = 1$, el cual es conocido como ICAR o IAR, resultando en una distribución impropia que puede ser rectificada imponiendo la restricción $\sum_{i=1}^{n} u_i = 0$.

- Además, se suele asumir que $\mu_i = 0$, para dar forma al modelo típicamente presentado en la literatura de mapas de enfermedades.

- Finalmente, si el modelo incorpora covariables, este se conoce como Ecologial regresión, es decir, se modifica

$$\eta_i = \boldsymbol{x}_i^{\top}\boldsymbol{\beta} + u_i + v_i$$

Bibliografia

Banerjee S., Carlin B., and Gelfand A. (2015) *Hierarchical Modeling and Analysis for Spatial Data*, Boca Raton: Chapman Hall/CRC.

Blangiardo M., and Cameletti M. (2015) *Spatial and Spatio-temporal Bayesian Models with R-INLA*, United Kingdom: Wiley & Sons.

Cressie, N. (1993) *Statistics for Spatial Data*. New York: Wiley.

Matheron, G., (1965) Les variables régionalisés et leur estimation. Masson, Paris.

Shabenberger, O., Gotway, C. A. (2005) *Statistical Methods for Spatial Data Analysis*. Boca Raton, FL: Chapman & Hall/CRC.

¿Alguna Consulta?