

Estadística Bayesiana

Clase 7: Modelos Jerárquicos

Isabel Cristina Ramírez Guevara

Escuela de Estadística
Universidad Nacional de Colombia, Sede Medellín

Medellín, 24 de agosto de 2020

Modelos Jerárquicos

Estos son modelos de múltiples parámetros que están relacionados o conectados por la estructura del problema que se está estudiando; como su nombre lo indica, se supone un modelo de probabilidad conjunto para los parámetros formando una jerarquía que refleja la dependencia entre ellos.

Modelos Jerárquicos

Estos son modelos de múltiples parámetros que están relacionados o conectados por la estructura del problema que se está estudiando; como su nombre lo indica, se supone un modelo de probabilidad conjunto para los parámetros formando una jerarquía que refleja la dependencia entre ellos.

Una ventaja de los métodos de simulación Bayesianos, como el MCMC, es que permiten realizar inferencias relativamente fáciles para modelos jerárquicos complejos.

Modelos Jerárquicos

Estos son modelos de múltiples parámetros que están relacionados o conectados por la estructura del problema que se está estudiando; como su nombre lo indica, se supone un modelo de probabilidad conjunto para los parámetros formando una jerarquía que refleja la dependencia entre ellos.

Una ventaja de los métodos de simulación Bayesianos, como el MCMC, es que permiten realizar inferencias relativamente fáciles para modelos jerárquicos complejos.

Ejemplo

En un estudio de efectividad de tratamientos cardíacos en varios hospitales, se considera que los pacientes del hospital j tienen una tasa de sobrevivencia θ_j , y que los estimadores de los θ_j provenientes de una muestra de hospitales deben relacionarse entre sí. Se podría utilizar una distribución a priori que considere a los θ_j 's como una muestra de una población común.

Modelos Jerárquicos-Intercambiabilidad

La idea de intercambiabilidad es crucial para establecer la probabilidad conjunta de todos los parámetros θ de un modelo jerárquico, esto es, suponer que dicha probabilidad conjunta $p(\theta_1, \dots, \theta_J)$ es invariante a permutaciones en los índices $(1, \dots, J)$. La forma más simple de un modelo intercambiable es suponer que los θ_j 's son una muestra independiente de una distribución poblacional gobernada por un parámetro desconocido ϕ ,

$$p(\theta_1, \dots, \theta_J) = \int p(\theta_1, \dots, \theta_J | \phi) p(\phi) d\phi = \int \prod_{j=1}^J p(\theta_j | \phi) p(\phi) d\phi$$

donde $p(\phi)$ es la distribución inicial del hiperparámetro ϕ .

Modelos Jerárquicos- Tratamiento Bayesiano

Se considera la distribución a priori conjunta:

$$p(\phi, \theta) = p(\theta|\phi)p(\phi).$$

Modelos Jerárquicos- Tratamiento Bayesiano

Se considera la distribución a priori conjunta:

$$p(\phi, \theta) = p(\theta|\phi)p(\phi).$$

La distribución posterior conjunta viene dada por:

$$p(\phi, \theta|\mathbf{y}) \propto p(\mathbf{y}|\phi, \theta)p(\phi, \theta)$$

$$= p(\mathbf{y}|\theta)p(\phi, \theta)$$

$$= p(\mathbf{y}|\theta)p(\theta|\phi)p(\phi).$$

porque la distribución de los datos sólo depende de θ . El hiperparámetro ϕ afecta a y a través de θ .

Modelos Jerárquicos- Tratamiento Bayesiano

Se considera la distribución a priori conjunta:

$$p(\phi, \theta) = p(\theta|\phi)p(\phi).$$

La distribución posterior conjunta viene dada por:

$$\begin{aligned} p(\phi, \theta|\mathbf{y}) &\propto p(\mathbf{y}|\phi, \theta)p(\phi, \theta) \\ &= p(\mathbf{y}|\theta)p(\phi, \theta) \\ &= p(\mathbf{y}|\theta)p(\theta|\phi)p(\phi). \end{aligned}$$

porque la distribución de los datos sólo depende de θ . El hiperparámetro ϕ afecta a y a través de θ .

En las secciones anteriores los hiperparámetros fueron considerados como conocidos, pero esta consideración en una buena parte de las aplicaciones no es realista; los modelos jerárquicos lo que hacen es incluir la incertidumbre asociada a ϕ .

Modelos Jerárquicos - Ejemplos

Modelo Normal

Verosimilitud (primera etapa):

$$Y_i | \theta_i \sim \text{Normal}(\theta_i, \tau_i), \text{ con } \tau_i \text{ conocida.}$$

Modelos Jerárquicos - Ejemplos

Modelo Normal

Verosimilitud (primera etapa):

$$Y_i | \theta_i \sim \text{Normal}(\theta_i, \tau_i), \text{ con } \tau_i \text{ conocida.}$$

Distribuciones a priori (segunda y tercera etapa):

$$\theta_i | \phi \sim \text{Normal}(\phi, \lambda), \text{ con } \lambda \text{ conocida.}$$

$$\phi \sim \text{Normal}(\phi_0, \lambda_0), \text{ con } \phi_0 \text{ y } \lambda_0 \text{ conocida.}$$

Modelos Jerárquicos - Ejemplos

Modelo Normal

Verosimilitud (primera etapa):

$$Y_i | \theta_i \sim \text{Normal}(\theta_i, \tau_i), \text{ con } \tau_i \text{ conocida.}$$

Distribuciones a priori (segunda y tercera etapa):

$$\theta_i | \phi \sim \text{Normal}(\phi, \lambda), \text{ con } \lambda \text{ conocida.}$$

$$\phi \sim \text{Normal}(\phi_0, \lambda_0), \text{ con } \phi_0 \text{ y } \lambda_0 \text{ conocida.}$$

Es posible que en la primera etapa se incluyan covariables, por

ejemplo:

$$\mathbf{Y} | \boldsymbol{\beta} \sim \text{Normal}(\mathbf{X}\boldsymbol{\beta}, \tau^2) \text{ con } \boldsymbol{\beta} | \phi \sim \text{Normal}(\phi, \lambda) \text{ y } \phi \sim \text{Normal}(\phi_0, \lambda_0).$$

Modelos Jerárquicos - Ejemplos

Ejemplo

Se realiza un estudio de niños con leucemia. Se cuenta con la información del número de muertes en dos regiones del Reino Unido en el año de 1950. Las tasas de muerte están clasificadas por edad (1-0:5, 2-6:14), el tipo de residencia del niño (1-rural, 2-urbano) y el tipo de cancer (1- Lymphoblastic, 2-Myeloblastic). Sea O_i el número de muertes observadas, P_i es el número total de niños (en millones) y θ_i la tasa de mortalidad en la clase i .

Modelos Jerárquicos - Ejemplos

Ejemplo

Se realiza un estudio de niños con leucemia. Se cuenta con la información del número de muertes en dos regiones del Reino Unido en el año de 1950. Las tasas de muerte están clasificadas por edad (1-0:5, 2-6:14), el tipo de residencia del niño (1-rural, 2-urbano) y el tipo de cancer (1- Lymphoblastic, 2-Myeloblastic). Sea O_i el número de muertes observadas, P_i es el número total de niños (en millones) y θ_i la tasa de mortalidad en la clase i .

Se proponen tres modelos para estos datos:

- Modelo de efectos constantes: $Y_i | \theta \sim \text{Poisson}(\theta P_i)$, $\theta \sim \text{Gamma}(\alpha, \beta)$ con α y β conocidas.

Modelos Jerárquicos - Ejemplos

Ejemplo

Se realiza un estudio de niños con leucemia. Se cuenta con la información del número de muertes en dos regiones del Reino Unido en el año de 1950. Las tasas de muerte están clasificadas por edad (1-0:5, 2-6:14), el tipo de residencia del niño (1-rural, 2-urbano) y el tipo de cancer (1- Lymphoblastic, 2-Myeloblastic). Sea O_i el número de muertes observadas, P_i es el número total de niños (en millones) y θ_i la tasa de mortalidad en la clase i .

Se proponen tres modelos para estos datos:

- Modelo de efectos constantes: $Y_i|\theta \sim \text{Poisson}(\theta P_i)$, $\theta \sim \text{Gamma}(\alpha, \beta)$ con α y β conocidas.
- Modelo de efectos independientes: $Y_i|\theta_i \sim \text{Poisson}(\theta_i P_i)$, $\theta_i \sim \text{Gamma}(\alpha, \beta)$ con α y β conocidas.

Modelos Jerárquicos - Ejemplos

Ejemplo

Se realiza un estudio de niños con leucemia. Se cuenta con la información del número de muertes en dos regiones del Reino Unido en el año de 1950. Las tasas de muerte están clasificadas por edad (1-0:5, 2-6:14), el tipo de residencia del niño (1-rural, 2-urbano) y el tipo de cancer (1- Lymphoblastic, 2-Myeloblastic). Sea O_i el número de muertes observadas, P_i es el número total de niños (en millones) y θ_i la tasa de mortalidad en la clase i .

Se proponen tres modelos para estos datos:

- Modelo de efectos constantes: $Y_i|\theta \sim \text{Poisson}(\theta P_i)$, $\theta \sim \text{Gamma}(\alpha, \beta)$ con α y β conocidas.
- Modelo de efectos independientes: $Y_i|\theta_i \sim \text{Poisson}(\theta_i P_i)$, $\theta_i \sim \text{Gamma}(\alpha, \beta)$ con α y β conocidas.
- Modelo de efectos intercambiables: $Y_i|\theta_i \sim \text{Poisson}(\theta_i P_i)$, $\theta_i|\alpha, \beta \sim \text{Gamma}(\alpha, \beta)$, $\alpha \sim \text{Gamma}(a, b)$ y $\beta \sim \text{Gamma}(c, d)$ con a, b, c, d conocidas.

Modelos Jerárquicos

La inferencia Bayesiana requiere encontrar la distribución posterior de los parámetros o cantidades desconocidas de los modelos. Obtener expresiones analíticas de la distribución posterior sólo es posible en casos particulares que usualmente representan modelos sencillos. Por tal razón es necesario estudiar métodos de simulación útiles para encontrar la distribución posterior de los parámetros.