

Apêndice Metodológico

Imaginemos, portanto, um modelo de análise realizado neste quinto capítulo, sendo Y a variável dependente que estamos tentando prever e X as variáveis preditoras. Uma outra variável, de natureza diferente, é YC, que é a junção das palavras em inglês 'Year' (ano) e 'Country' (país), sendo um identificador do caso. Essa variável não é exatamente um preditor, mas desempenha o papel de variável agrupadora. Como cada país tem diferentes ondas de dados, criamos essa variável que combina o ano da pesquisa, com o código numérico do país. Por exemplo, se um país é codificado como 20 e teve uma onda em 2006, seu código na variável agrupada será 202006¹. Isso funciona da seguinte forma: no software estatístico, após a formatação das variáveis Y e X, deve ser inserido o argumento de que a variável dependente varia entre os grupos, neste caso, os países, usando $\sim 1/YC$. Dessa maneira, YC desempenha sua função agrupadora.

Vamos agora descrever a lógica dos preditores nos diferentes níveis de análise, destacando exemplos dos modelos que realmente adotamos. Para facilitar a compreensão, apresentaremos um exemplo heurístico. No nível 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{11ij} + \beta_{2j}X_{12ij} + \epsilon_{ij}$$

Na modelagem realizada neste capítulo, a variável dependente (denominada Y_{ij}) é observada para a unidade 'i' (que pode ser o indivíduo, por exemplo), enquanto 'j' representa o grupo, que, no nosso caso, corresponde a cada combinação única de ano e país.

Os coeficientes ' β_{0j} ', ' β_{1j} ' e ' β_{2j} ' representam, respectivamente, o intercepto, o coeficiente para o preditor ' X_{11ij} ' e o coeficiente para o preditor ' X_{12ij} ' no grupo 'j'. Por sua vez, os preditores de nível 1 (' X_{11ij} ' e ' X_{12ij} ') para a unidade 'i' no grupo 'j' podem, por exemplo, ser variáveis como gênero, classe social ou adesão partidária. Por fim, o termo ' ϵ_{ij} ' refere-se ao erro associado à unidade 'i' no grupo 'j'. No nível 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}X_{21j} + \gamma_{02}X_{22j} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

$$\beta_{2j} = \gamma_{20}$$

Onde:

- 'γ₀₀' é o intercepto médio;
- 'γ₀₁' e 'γ₀₂' são os coeficientes que explicam a variação do intercepto em relação aos preditores de nível 2;
- 'γ₁₀' e 'γ₂₀' representam os efeitos médios dos preditores de nível 1. Esses preditores podem incluir, por exemplo, IDH, porcentagem de Evangélicos na amostra analisada ou grau de polarização política estimada na amostra à qual o respondente pertence;
- u_{0j} é o erro associado ao intercepto do grupo 'j'.

Esta é uma descrição básica do modelo de regressão multinível, com os preditores de nível 1 e nível 2 retirados dos modelos realmente adotados. É importante notar que a complexidade do modelo pode aumentar, dependendo de várias considerações, como a inclusão de mais preditores, a especificação de covariâncias entre os níveis, entre outros fatores.