Discussion (10/15/2013) Statistical Methods in Geographic Variations by Thérèse A. Stukel

James O'Malley, Ph.D.

The Dartmouth Institute for Health Policy and Clinical Practice

Geisel School of Medicine

Dartmouth College

Important Events in 2013

- Jack E. Wennberg Conference
- International Year of Statistics ("Statistics2013")
- 250 years since Bayes Theorem first presented publically
- 300 years since Jacob Bernoulli's Ars Conjectandi, seminal work on probability (including law of large numbers)

Why Model Provider Variations?

- Enable equitable comparisons between health units (hospitals, health plans, regions …)
  - Account for differences in patient populations
- Identify sources/causes of variations
- Understand patterns of variation
  - Report rates with maximum precision

Adjusting for Patient Case-mix

1. Model outcomes, then make predictions for target population
2. Model selection, then weight observations to equate to standard population
3. Combine the above two approaches

- First approach is most common; latter two seen more in survey analysis and causal inference

Controversies

- Type of Adjustment
  - Indirect standardization
  - Fixed effects model
  - Random effects models
- Inclusion of unit-level predictors (e.g., hospital volume) in model
- Borrowing strength from prior years
- Use of Bayesian modeling

An Experimental-design

- Randomly assign patients to hospitals (H)
- Measure outcome (Y) for each patient in each hospital
- Gain precision by adjusting for patient characteristics (X)
- Specification of H as random or fixed depends on target population of interest
  - These hospitals or all hospitals
  - Systematic and stochastic components of variation (McPherson, Wennberg et al 1982)

Observational Data: Patients not randomly assigned to hospitals

- Model as for randomized design

$$Y_{jk} = \beta_0 + \theta_j + \beta_1 X_{jk} + e_{jk}$$

- θ_j is effect of jth hospital
- Assume X_{jk} independent of e_{jk} (and for FE)
- Assume θ_j independent of e_{jk}
- Assume X_{jk} independent of θ_j ?
- Identification: θ_j has normal distn

Subtle Generalization of RE Model

- Repeated measures allow distinction between individual and group level effects

$$Y_{jk} = \beta_0 + \beta_b \bar{X}_j + \theta_j + \beta_w (\mathrm{X}_{jk} - \bar{X}_j) + e_{jk}$$

- Now θ_j independent of $\%X_{jk}^{0} = X_{jk} - \bar{X}_j$
- Perceived weakness of random effect modeling (Kalbfleisch and Wolfe, 2013) lessened/removed?
- Clarifies distinction between RE and FE!!!
- Ideally end up with $\mathrm{var}(\theta_j) = 0$

Shrinkage is natural?

- If the model is correct you must win!
- The problem is really that the mean is not an adequate summary. Why not report:
  - 5% and 95% quintiles of posterior distribution too
  - Summarizing information multiple ways is not multiple testing!
- If concerned about outliers use a model that allows for outliers (Jones and Spiegelhalter, 2011; Ohlssen et al 2006; West 1986)

Reduce shrinkage by allowing small hospitals to borrow from prior years

- Bidirectional smoothing (Jones and Spiegelhalter 2011)
- Use hierarchical model with time series component to yield better predictions of health care performance indicators

Summary

- Modeling and reporting are distinct tasks
  - Mixing them up leads to confusion
  - Output from fitted model can be translated in multiple ways to effectively communicate to stakeholders
  - Only one model!
- Hierarchical models are flexible
  - Embrace and allow for heterogeneity!

Another Generalization

- As in a randomized block design, allow interactions with block
- Random-intercept, random-slope model:

$$Y_{jk} = \beta_0 + \theta_{0j} + (\beta_1 + \theta_{1j})x_{jk} + e_{jk}$$

- Model identification: $\theta_j = (\theta_{j0}, \theta_{j1})$ has bivariate normal distn

Spatial statistics methods

- Geographic variations literature seldom uses models for spatial statistics
- If unit-level correlations dissipate smoothly with distance account, model spatial correlation and obtain more precise results
- Dilemma: Do you want to explain away spatial correlations?