Ranking Methods that Address Specific Goals

Thomas A. Louis, PhD, Research & Methodology Directorate
U. S. Census Bureau, thomas.arthur.louis@census.gov
Biostatistics, Johns Hopkins Bloomberg School of Public Heath
Baltimore MD USA, tlouis@jhsph.edu

Allocation of funds to high poverty regions, development of language-specific ballots, prioritization of public health interventions and environmental remediations in small areas, profiling health service providers, comparative evaluation of school effectiveness, gene and SNP identification studies; all depend on the relative position (ranks) of unit-specific parameters. Invalid or sub-optimal ranks can have serious scientific, policy and financial consequences. Therefore, societal equity, effective and efficient deployment of resources, and quality improvement require appropriate ranking with specific focus on identifying the highest and lowest ranking units. Ranking is challenging because intuitive approaches tend to perform poorly. For example, when estimation uncertainties vary over units to be ranked, those for which the variance of the direct (MLE) estimate is relatively high tend to be at the extremes when ranks are based on the them; on the other hand, units with relatively low variance tend to be at the extremes when ranks are produced by Z-scores testing that a unit's parameter is equal to all others. In essence, effective ranking requires finding the "sweet spot" between these two approaches, and Bayesian structuring coupled with a ranking-specific loss function has proven very effective in this regard. We outline the Bayesian approach and relevant loss functions, present simulation evaluations and a data illustration. We compare MLE-based, Z-score based, target parameter posterior mean-based, and Bayes optimal ranks. We use estimated Standardized Mortality Ratios from the United States Renal Data System to identify domain-specific considerations, illustrate approaches, and generalize issues to other contexts. Results show that optimal ranks substantially outperform the other contenders, underscoring the importance of basing estimates on a formal structure. At least as important, results show that in most realistic situations even optimal ranks are estimated with considerable uncertainty that should be taken into account when implementing policies. Along with ranks, unit-specific MLE or posterior mean estimates of the target parameters generally are reported, and ranks based on these are different from the optimal ranks. We address this face-validity issue by presenting "triple-goal" estimates, those for which the ranks of the estimates are optimal, the histogram of the estimates is optimal, and that retain much of the Bayes advantage in estimating unit-specific parameters.

Key Words: Bayesian Modeling, Ranking Targeted Loss Functions, Triple-goal Estimates