

A Bayesian Two-Part Predictive Model for Health Care Costs Using Individual-level Data

Margie Rosenberg

University of Wisconsin-Madison

Abstract: The incidence rates and costs of inpatient health care are each known to have long-tailed distributions. A Bayesian two-part model that considers the probability of hospitalization in the first part and the costs of the inpatient stay in the second part is developed using individual-level data collected over time. Predictive distributions for each individual resulting from the modeling process are used to estimate the number of hospitalizations and costs by individual, and by calendar year. One advantage of this approach is that differences by individuals, such as disease severity, are automatically reflected. In addition, these predictive distributions can be used to illustrate the extreme variability in hospital utilization. Knowledge of this potential uneven and unpredictable occurrence of utilization, and potential cost, would be beneficial in the design of insurance programs or for disease management programs.