

Methods for Analyzing Health Care Claims Data

Veronica Montes De Oca

May, 2003

Abstract

In the paper *The Journal of Behavioral Services & Research*: “Utilization and cost of behavioral health services: Employee characteristics and workplace health promotion,” an analysis of behavioral health care utilization and costs was performed on health care claims data, including demographic and employment related factors. A two-part modeling method was used. The first modeling method consisted of logistic regression in order to identify the factors influencing the likelihood of utilizing behavioral health services. The second modeling method was linear regression in order to identify the factors influencing the cost of behavioral health claims.

In my thesis, I will use the linear regression method with a different set of health care claims data and focus on the factors influencing the cost of radiology claims, procedures and diagnosis. In addition, I will use bootstrapping as a nonparametric method of validating traditional linear regression. I will show the advantages and disadvantages of each method.