Statistics and Healthcare Fraud

by Don Edwards
University of South Carolina

Abstract. Healthcare fraud is a serious problem facing the United States. The National Health Care Anti-Fraud Association (www.nhcaa.org) states that in 2007 over 4 billion health insurance claims were processed in the United States and that fraud amounted to $68 billion. The problem has been featured in a number of NBC News Fleecing of America segments (most recently in January, 2010) and also a recent CBS News Sixty Minutes segment (October 2009).

Statistical methods are center-stage in the battle against healthcare billing abuse. The federal governments Centers for Medicare and Medicaid Services (CMS) guidelines (2011) stipulate that when a healthcare provider (physician, clinic, hospital, durable goods supplier, etc.) is selected for audit, a population of payments made to that provider over a specified timeframe is isolated and a probability sample selected. A certified investigator examines documentation and other evidence and assigns an overpayment amount for each sampled payment as the amount paid minus the amount justified by the evidence. These sample overpayments are then used to extrapolate to the universe: a 90% lower confidence bound for the total population overpayment is calculated and forwarded to the provider as a recoupment demand. The provider has the right to challenge the sampling and extrapolation methods in court, and often does so: statisticians working in this arena should be prepared to defend every decision they make under oath, alone, against an aggressive and well-funded team of provider lawyers and impressively educated statistical experts.

Until recently, the extrapolation methods in use for healthcare audits have almost exclusively been methods based on the central limit theorem (CLT). These cannot guarantee the prescribed 90% confidence level / under-recoupment rate and in fact often fail badly. The reasons for this are discussed, as well as alternative extrapolation methods based on counts and the hypergeometric distribution.

* Refreshments at 2:30 PM in RT 1517