Stochastics and Statistics

Estimating the number of undetected software errors via the correlated capture–recapture model

Young H. Chun [*]

Department of Information Systems and Decision Sciences, E.J. Ourso College of Business Administration, Louisiana State University, Baton Rouge, LA 70803-6316, USA

Abstract

Sometimes a complex software system fails because of errors undiscovered in the design stage of the development process. Detecting these errors early in the process would eliminate many downstream problems. The so-called "capture–recapture" model, initially used by biologists to estimate the size of wildlife populations, has been widely used to estimate the number of software design errors. However, one simplifying assumption in capture–recapture models is that the inspections performed by various inspectors are statistically independent from each other. In the paper, we propose a novel method that is based on the correlation matrix of multiple inspectors. In a numerical analysis, we show that our method outperforms other traditional models that are based on the independence assumption.
© 2005 Elsevier B.V. All rights reserved.

1. Introduction

Sometimes a complex software system fails because of errors, bugs, or faults introduced in the design stage of the development process. The purpose of software inspection is to detect and correct those errors before they leak through subsequent development phases and the field. Design reviews can remove some of these errors, but often a few errors remain undetected until the software is developed further. It is important to eliminate these design errors at an early stage because they become much more expensive to fix as a software system proceeds through development.

[*] Tel.: +1 2255782506; fax: +1 2255782511.
E-mail address: chun@lsu.edu