

# Software Reliability for Multi-Type Defects: Application to Modern Bug Databases

Vignesh T Subrahmaniam  
Software Sciences and Analytics  
GE Global Research, Bangalore, 560066, India  
vignesh.ts@ge.com

Anup Dewanji\*  
Applied Statistics Unit  
Indian Statistical Institute, Kolkata, 700108, India  
dewanjia@isical.ac.in

and

Bimal K Roy  
Applied Statistics Unit  
Indian Statistical Institute, Kolkata, 700108, India  
bimal@isical.ac.in

**Abstract:** Software bug databases provide an important source of data for assessing the reliability of a software. Statistical analysis of these databases has been a challenge due to lack of information on usage and reporting rates of the defects. This work proposes a novel semi-parametric analysis that makes use of defect classifications into multiple types to enable estimation of a model without making strict assumptions about the underlying usage rate of the software. The proposed model can account for differences in reporting rates of different classes of defects and the self-exciting nature of the user-driven defect discovery process. Estimation of the proposed model can be done using generalized linear model procedures. New reliability metrics whose computation does not depend on the unknown usage rate of the software have been proposed and methods for estimating them have been developed. The proposed model has been applied to data retrieved from the bug database of a popular scripting language, named Python.

**Keywords:** Multivariate counting process, logarithmic Poisson model, partial likelihood, reliability metrics.