Evaluation of the CLDP Sampling Methodology for POD Estimation Applied to Hit Miss Analysis with Dependent Data

The major goal of a probability of detection (POD) study is to be able to determine the size of the flaw that can be detected by the nondestructive inspection (NDI) system with reasonable certainty. Generally, this process begins with the use of experimental data from which a POD curve is generated in order to compute important estimators such as a90 (the flaw size detected with 0.90 probability) and a90|95 (the largest flaw size detectable with 0.90 probability at 95% confidence). If binary data is provided, a hit-miss analysis approach is used in which the POD is estimated as a function of flaw size. Traditional statistical methods assume data structures in which experimental data is independent, however, advanced statistical methods are emerging for other data structures. Often these methods focus on the modeling aspect more so than on sampling methods that may enhance and enable the modeling approach.

This work considers the hit-miss modeling approach for POD data and begins by examining dependent experimental data. Dependent data is common in both structural health monitoring applications and experimental designs in which test specimens are expensive and, therefore, multiple measurements are collected on each specimen, usually until flaws reach a relatively large size. The logistic model used for hit-miss analysis can be extended for dependent data, but may struggle with convergence and estimation issues, often exacerbated by excessive measurements when the flaw is too small or too large. To overcome these issues, this work presents a new sampling strategy called the Covariance Limiting Defining Pair (CLDP) that uses data associated with the largest flaw such that all smaller flaws were not detected and the smallest flaw such that all larger flaws were detected for a given specimen. These points are termed “limit defining” because they directly contribute to defining the most valuable portion of the POD curve with respect to flaw size estimation. A POD curve can be efficiently built for hit-miss data using the CLDP sampling method, limiting the amount of required data, thus reducing model complexity while maintaining the ability to estimate values such as a90 and a90|95. Simulation and experimental data are used to demonstrate both the robustness of the CLDP sampling method for estimation and its efficiency to reliably estimate POD statistics while testing fewer specimens.