Abstract

Name of the student: **Jai Bishnu Agrawal** Roll No: **231030024**

Degree for which submitted: **M.Tech.** Department: **Department of Civil Engineering**

Thesis title: **Incorporating Time-varying Covariates in Crash Risk Modeling using Survival Analysis and Count Models**

Thesis supervisor: **Dr. Aditya Medury**

Month and year of thesis submission: **May 2026**

This thesis examines how crash risk estimation changes with the choice of modeling framework and the temporal resolution of crash data. Traditional crash prediction approaches, such as Poisson and Negative Binomial models, estimate crash frequency by aggregating crashes over fixed time intervals, typically monthly or yearly, and relating these counts to roadway and traffic characteristics. Although these models are widely used, they do not account for the exact timing of crash events within the observation period. As a result, short-term variations in crash risk and the effects of variables that change over time, such as traffic flow and rainfall, may not be adequately represented.

To address these limitations, this study applies survival analysis, specifically the Andersen–Gill extension of the Cox proportional hazards model, as an alternative framework for crash risk modeling. Instead of working with aggregated crash counts, the survival-based approach models the time intervals between consecutive crashes. This allows the analysis to incorporate time-varying factors more naturally and to represent repeated crashes occurring at the same roadway segment as separate events over time.

The study begins with a simulation analysis in which crash events are generated using NHPP. Since the underlying relationships between variables are predefined, the simulation provides a controlled setting for comparing the performance of survival and count-based models. Multiple scenarios are developed to evaluate how rainfall and different levels of temporal aggregation influence model behavior and predictive performance.

The methodology is then applied to real-world crash and traffic data collected from 54 signalised intersections in Melbourne, Australia. Models are estimated using both monthly and yearly data in order to investigate how temporal aggregation affects parameter estimation, prediction accuracy, and the identification of significant risk factors.

The findings highlight important differences between survival-based and traditional count-based approaches in the way crash occurrence is represented and interpreted. Overall, the study demonstrates that incorporating temporal detail and time-varying covariates can improve the understanding of crash dynamics and provide a more flexible framework for crash risk assessment.