

MODELING DELAY BEHAVIOR IN COMMERCIAL AVIATION: A STATISTICAL DISTRIBUTION APPROACH

K.S.D. Ravihari* and S.P. Abeyesundara

Department of Statistics and Computer Science, University of Peradeniya, Peradeniya, Sri Lanka.

**s19497@sci.pdn.ac.lk*

Flight delays remain a persistent challenge for the aviation industry, impacting both passenger satisfaction and operational efficiency. This study aimed to identify suitable probability distributions for flight delay durations and to model their underlying patterns more effectively than conventional predictive methods. Data were obtained via web scraping from FlightAware.com, comprising approximately 26,000 flights operated by major international airlines between popular global airports from January 2023 to April 2025. The dataset included records from six major airlines and 11 high-traffic departure-arrival airport pairs. After cleaning and preprocessing, delays were calculated as the difference between actual and scheduled times, focusing on delays of 15 minutes or more, consistent with industry standards. Exploratory analysis revealed that delay durations exhibited multimodal and positively skewed behaviour. Several continuous probability distributions, including Lognormal, Weibull, Gamma, Pareto and Exponential, were fitted to the data, with goodness-of-fit assessed using the Anderson–Darling test. Model selection was guided by the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to balance fit and complexity. Recognising the limitations of single distributions, mixture models were estimated using the Expectation-Maximisation algorithm, with optimal components determined through likelihood ratio tests. Results showed that lognormal mixture models best captured the complex distribution of delays, although the number of mixture components varied, reflecting heterogeneity across flights and airlines. Distinct distribution peaks were observed at specific thresholds, such as 15 minutes for arrivals and 17 minutes for departures. These findings provide actionable insights for airlines and airport authorities, supporting improved scheduling and operational strategies. Future research could enhance and stratify these models by incorporating additional factors such as weather conditions, aircraft characteristics and air traffic congestion, leading to more robust statistical analysis.

Keywords: Aviation statistics, Flight delays, Mixture models, Operational efficiency, Probability distributions