

Effect of Truncation Based on Past Entropy

By

Farah Ghazi Alsakran

Supervisor

Ma'mon Abu Hammad

Al-Zaytoonah University of Jordan, 2024

Abstract

This thesis explores the historical interplay between truncation, entropy, and information measures, investigating how limiting or cutting off data, known as truncation, has influenced the understanding and analysis of information throughout history. Focusing specifically on entropy, a key concept in information theory, the research aims to unravel the intricate relationship between approaches to truncation and conceptions of entropy. The study unfolds in three main dimensions. First, it examines the historical applications of truncation, scrutinizing how past interpretations of entropy have shaped decisions related to data truncation, thus providing valuable insights into the historical development of information measures. Second, a critical analysis of mathematical expressions of entropy in various contexts is undertaken to assess their impact on determining optimal points for truncating data sets, offering a deeper understanding of the theoretical underpinnings. Lastly, the thesis proposes novel approaches for measuring the amount of information lost through truncation techniques, incorporating measures of entropy. By developing effective methodologies for quantifying information loss, the research aims to provide practical tools for present-day researchers and professionals, enabling them to make informed choices when employing data truncation strategies. In summary, this research contributes to our understanding of the interconnected nature of truncation, entropy, and information measures by offering a

comprehensive analysis of their historical evolution, intending to guide contemporary practitioners in strategic decision-making regarding data truncation, while drawing lessons from earlier entropy-influenced approaches. This thesis approximates the quantiles of the asymptotic distribution of this statistic. Furthermore, the thesis introduces an enhanced version of this statistical measure. Remarkably, even in cases involving small sample sizes, the proposed method provides a formula for determining quintiles based on the chi-square distribution.

Keywords: Shannon Entropy, Distributions, Past Entropy.