Formalization of Normal Random Variables in HOL

Muhammad Qasim[1(✉)], Osman Hasan[1], Maissa Elleuch[2,3], and Sofiène Tahar[1]

[1] Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada
{m_qasi,o_hasan,tahar}@ece.concordia.ca
[2] CES Laboratory, Sfax University, Sfax, Tunisia
maissa.elleuch@ceslab.org
[3] Digital Research Center of Sfax, Sfax, Tunisia

Abstract. Many components of engineering systems exhibit random and uncertain behaviors that are normally distributed. In order to conduct the analysis of such systems within the trusted kernel of a higher-order-logic theorem prover, in this paper, we provide a higher-order-logic formalization of Lebesgue measure and Normal random variables along with the proof of their classical properties. To illustrate the usefulness of our formalization, we present a formal analysis of the probabilistic clock synchronization in wireless sensor networks.

1 Introduction

Many engineering systems exhibit *normally distributed* elements of randomness. Some notable examples include noise in communication channels, lengths and weights of manufactured goods, message arrival times in communication networks, blood pressure readings of a general population, lifetimes of an electric bulb and maximum speed of a car. The importance of normal distribution is also evident from its relationship with the central limit theorem [2], which states that, given certain conditions, the arithmetic mean of a sufficiently large number of iterations of independent random variables, each with a well-defined expected value and variance, is approximately normally distributed, regardless of the underlying distribution [20]. Therefore, if the sample size is large enough, the sample mean of other distributions may also be treated as normal.

Traditionally, paper-and-pencil based approaches are used for carrying out probabilistic analysis. This method, however, is prone to human error and is not scalable to deal with large systems. Similarly, simulation cannot provide accurate results due to approximations in numerical computations and its incompleteness, which is an outcome of enormous processing time requirements.

Given the safety-critical nature of present age engineering systems, these inaccuracies cannot be tolerated. Higher-order-logic theorem proving, which provides computerized mathematical proofs, can overcome the above-mentioned limitations and has been used to formalize probability theory [16], Markov

M. Kohlhase et al. (Eds.): CICM 2016, LNAI 9791, pp. 44–59, 2016.
DOI: 10.1007/978-3-319-42547-4_4