Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

PARAMETRIZATION OF CUMULATIVE MEAN BEHAVIOR OF SIMULATION OUTPUT DATA

Dashi I. Singham
Michael P. Atkinson

Naval Postgraduate School
1411 Cunningham Road
Monterey, CA 93943, USA

ABSTRACT

We develop a new measure of reliability for the mean behavior of a process by calculating the probability that the cumulative sample mean will ever deviate from its long-term mean, and its true mean, over a period of time. This measure can be used as an alternative to estimating system performance using confidence intervals. We derive the tradeoffs between four critical parameters for this measure: the underlying variance of the data, the starting sample size of a procedure, and the precision and confidence in the result.

1 INTRODUCTION

We propose a new metric for evaluating system performance that is stronger than the traditional confidence interval. We present a measure of reliability for the cumulative mean behavior of a process, by calculating the probability that the sample mean of a time series stays within some fixed distance from its long-term mean after a given initial sample size. The long-term mean could be the true mean, or the sample mean after a long period of time. The underlying time series is assumed to meet the conditions for a functional central limit theorem (FCLT), an assumption used in many simulation output analysis methods.

We calculate this measure by structuring simulation output data as a standardized time series (Schruben 1983), which under the FCLT assumption is a Brownian bridge in the limit. Manipulating the standardization allows us to evaluate the difference between the cumulative mean and the long-term mean as a function of a Brownian bridge. We derive a lower bound for the probability that this difference between the means is always less than a specified amount after a specified initial sample size by calculating boundary crossing properties of Brownian bridges. This measure provides more information than a traditional confidence interval, which only evaluates the cumulative mean once.

In addition to the implications for confidence interval procedures, this measure is useful in experimental settings. Examples include evaluating a production system over a year, where cumulative average performance each month converges to average performance over the year. We may be interested in knowing the likelihood that the cumulative performance early in the year will deviate from the end of year results.

2 A MEASURE OF RELIABILITY FOR MEAN BEHAVIOR

We calculate the probability that the cumulative sample mean of simulation output data Y_i after k initial samples, which is $\overline{Y}_i, i = k, \ldots$ ever deviates from its long term mean $\overline{Y}_m$, or its true mean μ, by more than δ. Let σ^2 be the variance of the Y_i data points, δ be the allowable deviation, and m be the total sample size. We write this measure, standardized by time, as PB (probability in bounds):

$$PB \equiv P\left(\bigcap_{t\in[k/m,1]} \left|\frac{1}{m}\sum_{i=1}^{m} Y_i - \frac{1}{\lfloor mt \rfloor}\sum_{i=1}^{\lfloor mt \rfloor} Y_i\right| \leq \delta\right).$$

Our result gives a lower bound for PB.

Theorem 2.1. Let Φ be the cumulative distribution function of the standard normal distribution. Under the FCLT assumption, the probability that the cumulative sample mean $\overline{Y}_i$ stays within distance δ from the long-term mean $\overline{Y}_m$ over the range $i = k, \ldots, m$ has a lower bound

$$PB_L(\delta,\sigma,k,m) = 1 - 4\sum_{j=1}^{\infty}\left(\Phi\left(\frac{\delta\sqrt{k}}{\sigma\sqrt{1-\frac{k}{m}}}(4j-1)\right) - \Phi\left(\frac{\delta\sqrt{k}}{\sigma\sqrt{1-\frac{k}{m}}}(4j-3)\right)\right).$$

The lower bound for the probability that the sample mean $\overline{Y}_i$ stays within distance δ from μ over the range $i = k, \ldots, \infty$ as $m \to \infty$ is

$$PB_L(\delta,\sigma,k) = 1 - 4\sum_{j=1}^{\infty}\left(\Phi\left(\frac{\delta\sqrt{k}}{\sigma}(4j-1)\right) - \Phi\left(\frac{\delta\sqrt{k}}{\sigma}(4j-3)\right)\right). \quad (1)$$

3 RESULTS AND CONCLUSIONS

Figure 1 shows PB_L in the limit as $m \to \infty$ for a variety of combinations of k and δ/σ using (1). Because δ and σ only appear in (1) as δ/σ we condense them to one term on the y-axis. On the x-axis, as k increases we see that PB_L increases, because the sample mean is more likely to deviate from the true mean at smaller sample sizes. On the y-axis, when δ is high relative to σ, PB_L is higher because the bounds are loose relative to the variance of the procedure. However, when σ is large, this ratio decreases and the probability of staying within the bounds decreases. This shows the importance of having δ/σ relatively large, for any value of k, or alternatively, if a small δ is required, to use a larger k.

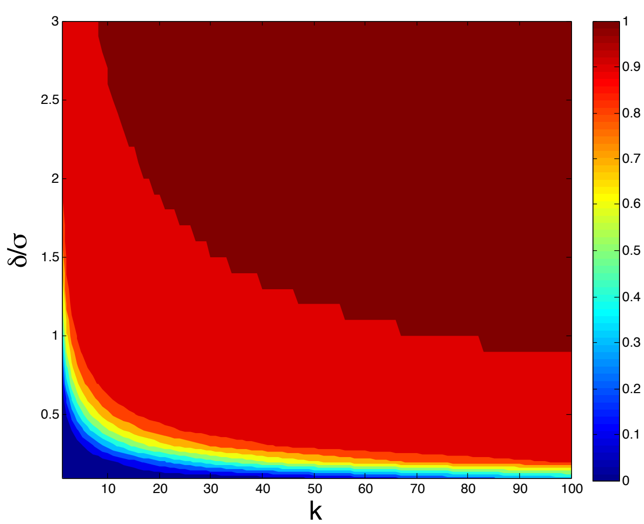

Figure 1: Calculation of (1) showing PB_L for a range of values of k and δ/σ, as $m \to \infty$.

REFERENCES

Schruben, L. 1983. "Confidence Interval Estimation Using Standardized Time Series". *Operations Research* 31 (6): 1090–1108.