Limitations of Realistic Monte-Carlo Techniques

A. M. Ponwuk, **O. M. Kosheleva**, and **V. Kreinovich**

Computational Science Program, University of Texas at El Paso,
El Paso, TX 79968, USA,{ampownuk, olgak, vladik}@utep.edu

Keywords: *Monte-Carlo Techniques; Interval Uncertainty.*

Abstract

Data processing means applying some algorithm $f(x_1, \ldots, x_n)$ to the values of the quantities $x_1, \ldots, x_n$, resulting in a value $y = f(x_1, \ldots, x_n)$. Values x_i usually come from measurements. Measurement are never absolutely accurate; the measurement result $\widetilde{x}_i$ is, in general, different from the actual (unknown) value x_i of the corresponding quantity: $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i \neq 0$. Because of the this, the computed value $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ is, in general, different from the ideal value $y = f(x_1, \ldots, x_n)$. It is therefore desirable to estimate the accuracy $\Delta y \stackrel{\text{def}}{=} \widetilde{y} - y$.

In many practical situations, the measurement errors are relatively small. In such cases, we can safely ignore terms which are quadratic or higher order in Δx_i, and conclude that $\Delta y = \sum\limits_{i=1}^{n} c_i \cdot \Delta x_i$, where $c_i \stackrel{\text{def}}{=} \dfrac{\partial f}{\partial x_i}$.

When we know the probability distributions for all Δx_i (and we know that they are independent), then we can use Monte-Carlo techniques: several times $k = 1, \ldots, N$, we simulate $\Delta x_i^{(k)}$, then the differences $\widetilde{y} - f(\widetilde{x}_1 - \Delta x_1^{(k)}, \ldots, x_n - \Delta x_n^{(k)})$ have the same distribution as Δy.

Alternatively, we can use numerical differentiation to estimate all the derivatives c_i, and then use the above formula, but this would require $n+1$ calls to the algorithm f, which for large n can be too long. In contrast, the Monte-Carlo method needs $N+1$ calls, where N is determined only by the accuracy with which we want Δ (and does not depend on n).

In many practical situations, we only know the upper bound Δ_i on each measurement error Δx_i: $|\Delta x_i| \leq \Delta_i$. In this case, the only information that we have about the unknown (actual) value x_i is that it is in the interval $[\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$. In this case, the value Δy is bounded by $\Delta = \sum\limits_{i=1}^{n} |c_i| \cdot \Delta_i$. A straightforward computation of Δ requires $n+1$ calls to f, but there is a faster method based on using Cauchy distribution: we simulate $\Delta x_i^{(k)}$ based on Cauchy with parameter Δ_i, then the differences $\widetilde{y} - f(\widetilde{x}_1 - \Delta x_1^{(k)}, \ldots, x_n - \Delta x_n^{(k)})$ are Cauchy distributed with the desired parameter Δ.

This method works, but its simulation is not realistic: we know that $|\Delta x_i| \leq \Delta_i$, but a Cauchy distribution goes beyond this bound. It has been known that if we consider simulations in which all $\Delta x_i^{(k)}$ are independent, then no realistic Monte-Carlo technique can always compute Δ. In this paper, we prove that this result holds even without the independence assumption, i.e., that the simulated values $\Delta x_i^{(k)}$ *have to* go beyond $[-\Delta_i, \Delta_i]$.