

Car racing simulations: Monte Carlo Method vs Latin Hypercube Sampling.

Gianluca Rosso

gianluca.rosso@sis-statistica.org

Published on LinkedIn Pulse in July 7, 2016

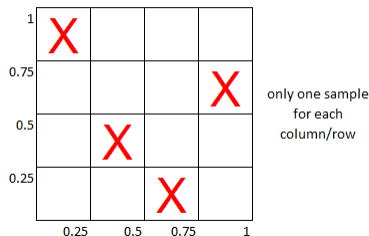
Due to the fact that during pit-stops no refuel is done but only tyres change, fresh tyres could be a winning option, but we must consider that pit-stops have a time particularly high (due to pit-in and pit-out essentially) with some risk variables: mistakes during tyres change (due to technical failure or human errors), getting stuck in traffic, pit-in mistakes (speed exceeding), pit-out mistakes (pit exit line crossing), traffic into the pitlane, probability of a Safety Car.

These risks must be considered. Each element is normally associated with a particular probability of occurrence. This probability is related to data collection from other races. The number of permutations is vast and the need is to know in advance.

So in 2000 Monte Carlo method (MC) came in. This technique uses randomly generated numbers in order to approximate outcome. The result of a probabilistic simulation (Monte Carlo simulation) is a quantified probability.

But just looking at different computational fields a better method could be considered. We are talking about the Latin Hypercube Sampling (LHS) method, which is the answer just coming from banking and finance risk management tools. With LHS we need less samples to achieve a needed threshold (in races every performance is measured as a threshold). With the same number of samples LHS achieve a better simulation performance in relation with MC. This because MC is memoryless: sample points are generated without taking into account the previously generated sample point).

SQUARE GRID



LHS has a memory, so is more efficient. The key of LHS method is the stratification of the data distribution and sample is more representative (in MC sampling you might end up with some points clustered closely, while other intervals within the space get no samples).

In LHS there are two methods:

1. RANDOM (RLHS) – selecting random points in each interval;
2. MEDIAN (MLHS) – using the median value of each equiprobable interval.

The number of "dimensions" is arbitrary, whereby each sample is the only one in each axis-aligned hyperplane containing it.

LHS can achieve the same threshold ϵ using less number of samples.

$$\max_{i,j} \Delta_{ij} \leq \epsilon \quad \text{where} \quad \Delta_{ij} = \frac{\|x_i - x_j\|_1}{\|x_i + x_j\|_1}$$

