Stochastic Calculus: Defining the Integral

Thomas Norton with Prof. Wojbor Woyczynski
Department of Mathematics – Case Western Reserve University

Background: Stochastic Modeling

In stochastic modeling we seek to incorporate "noise" into the way we describe a system. By "noise" we mean factors that we cannot predict for certain, but with a certain probability. For example, a deterministic model for the temperature of the water in Wade Lagoon might just use average temperature data for Cleveland. A stochastic model would account for the nonzero probability of freezing temperatures in April. The deterministic model might predict that the temperature of the lagoon will be 60°F on April 18th; the stochastic model might say that there is an 85% chance that the temperature of the lagoon will be between 65°F and 70°F on April 18th. The stochastic model is mathematically trickier, but it's well worth it for the additional information it provides.

Stochastic Processes:
The solution returned by a stochastic model is known as a stochastic process. A stochastic process is a function of time whose values are not determined, but rather given by a probability distribution, as in the above example.

The most widely used stochastic process is the Brownian Motion process. Brownian Motion models the diffusion of a single particle through liquid, i.e. a continuous random walk. We understand Brownian Motion very well, so we usually try to express the "noise" in a stochastic model in terms of Brownian Motion.

Example: Population Growth Model

A stochastic model generally comes in the form of a Stochastic Differential Equation (SDE). Here's what this means in terms of the familiar population growth model (adapted from [1]):

A population has size $N(t)$ at time $t$ and individuals are born and die proportional to the size of the population. The difference equation governing this is $N(t+1) = N(t) + \alpha N(t)$. Or $\Delta N(t) = \alpha N(t)$, where $\Delta$ is the difference between the birth and death rates.

Note that in the usual model, we assume that we know what $\alpha$ is at all times. Of course, this is not very realistic. Birth and death rates fluctuate about the average based on the availability of food, spread of diseases, etc. Here is where a stochastic model improves on deterministic models: in a stochastic model, we replace $\alpha$ with $\alpha + \text{Noise}$. "Noise" is a stochastic process that we choose to have a probability distribution describing the likelihood of given deviations from the birth and death rates.

Adding this term, we transition from a deterministic to a stochastic model:

$$\Delta N(t) = \alpha N(t) + \text{Noise}$$

In a continuum limit, this stochastic difference equation becomes a stochastic differential equation:

$$dN(t) = \alpha N(t)dt + N(t)dB(t)$$

Stochastic Integration

Just as we need integrals to solve ordinary differential equations, we need stochastic integrals to solve stochastic differential equations:

$$dN(t) = \alpha N(t)dt + N(t)dB(t)$$

$$N(T) = N(0) + \int_0^T \alpha N(t)dt + \int_0^T N(t)dB_t$$

The focus of this project is on choosing a suitable definition of the stochastic integral to optimize numerical calculations. Recall the usual calculus definition of integrals:

$$\int_{x}^{y} f(x)dx = \lim_{\Delta x \to 0} \sum_{i=0}^{n} f(x_i)(y-x_i)$$

Defining the integral requires specifying the timepoint $t^*$ at which we evaluate the integrand. This choice was irrelevant in ordinary calculus, but proves important in the stochastic integral. In fact, each possible choice of $t^*$ leads to a distinct definition of the integral. In this project, we integrate numerically using many different definitions (i.e. many choices of $t^*$) to look for a definition that converges more quickly than the others.

Methods

Numerical integration is performed as the limit of a Riemann-Stieltjes sum:

$$\int_{a}^{b} f(x)dx = \lim_{\Delta x \to 0} \sum_{i=0}^{n} f(x_i)(y-x_i)$$

We pick an integrand $f$ and choose which definition of the stochastic integral to compute with (i.e. which choice of $t^*$). We then generate the Brownian Motion process and compute the above sum for a very large value of $N$ – this is what we use for the "true" value of the integral. We then compute the above sum for smaller values of $N$ and measure how much these deviate from the "true" large $N$ value. This tells us the convergence behavior of the chosen definition of the stochastic integral for the chosen integrand $f$.

These calculations are performed in Matlab according to the following schematic:

![Schematic Diagram]

Acknowledgements / References

I would like to first thank Professor Woyczynski for his guidance throughout this project and the graduate admissions process. I would also like to thank Professors Kowalski and Singer for their guidance in the senior project seminar.