

**week7b: Kapitel 4: Monte Carlo Integration**  
**4.3: Additive und multiplikative Zufallsvariablen**

Wir hatten letztes Mal gesehen, dass wir ein Integral der Form

$$I := \int_{\mathbb{R}^n} F(\phi) p(\phi) d^n \phi \quad (1)$$

mit  $p \geq 0$  und  $\int p(x) d^n x = 1$  durch die folgende Monte Carlo Summe approximieren können:

$$I \approx S_N(F) := \frac{1}{N} \sum_{i=1}^N F(\phi_i) \quad (2)$$

wenn die  $\{\phi_i\}_{i=1}^N$  eine Folge von  $N$  unabhängigen  $p$ -verteilten Zufallsvektoren (oder Zahlen, falls  $n = 1$ ) sind. Aus dem zentralen Grenzwertsatz hatten wir dann den folgenden Ausdruck für den Monte Carlo Fehler hergeleitet:

$$\text{Monte Carlo Fehler} \sim \sqrt{\frac{\mathbf{V}[F]}{N}} \quad (3)$$

In typischen Anwendungen ist die Varianz  $\mathbf{V}[F]$  von derselben Grössenordnung wie das  $\mathbf{E}[F]$ , vielleicht 4 mal so gross oder auch nur halb so gross oder auch mal 10 mal so gross, aber auf jeden Fall unabhängig von  $N$ , der Anzahl der Monte Carlo Simulationen. Da das  $F$  ja gegeben ist, da kann man zunächst mal nichts variieren, betrachtet man das  $\mathbf{V}[F]$  in (3) typischerweise als Konstante und sagt dann, der Monte Carlo Fehler geht wie Eins durch Wurzel  $N$ , also

$$\text{Monte Carlo Fehler} \sim \frac{1}{\sqrt{N}} \quad (4)$$

In den meisten Fällen, etwa bei der Berechnung von Optionspreisen mit Monte Carlo, das ist die Standardmethode etwa in den Derivate-Abteilungen von Grossbanken, liefert das auch genau das richtige Verhalten. Es gibt allerdings Beispiele, wo die Varianz  $\mathbf{V}[F]$  einen geradezu dramatischen Einfluss auf die Rechengenauigkeit oder more basic auf die prinzipielle Monte Carlo Berechenbarkeit hat, und das wollen wir uns hier jetzt anschauen.

Wir betrachten die folgenden Grössen: Es seien  $\phi_1, \phi_2, \dots, \phi_n$  unabhängige, standard-normalverteilte Zufallszahlen und definieren dann

$$s_n := \frac{1}{n} \sum_{i=1}^n \phi_i^2 \quad (5)$$

$$\pi_n := \prod_{i=1}^n \phi_i^2 \quad (6)$$

Offensichtlich gilt dann

$$\mathbb{E}[s_n] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\phi_i^2] = \frac{1}{n} \sum_{i=1}^n 1 = 1 \quad (7)$$

$$\mathbb{E}[\pi_n] = \prod_{i=1}^n \mathbb{E}[\phi_i^2] = \prod_{i=1}^n 1 = 1 \quad (8)$$

Wir werden gleich sehen, dass, wenn wir versuchen, diese Erwartungswerte mit Monte Carlo zu berechnen, wir ein massiv unterschiedliches Verhalten bei den  $s_n$ 's und den  $\pi_n$ 's haben. Das liegt an den Varianzen, berechnen wir die: Bei einer Summe von unabhängigen Zufallsvariablen addieren sich einfach die Varianzen, diese Rechnung hatten wir ja auch schon beim zentralen Grenzwertsatz im Kapitel 2 gemacht:

$$\begin{aligned} \mathbb{V}[s_n] &= \mathbb{V}\left[\frac{1}{n} \sum_{i=1}^n \phi_i^2\right] \\ &= \frac{1}{n^2} \mathbb{V}\left[\sum_{i=1}^n \phi_i^2\right] \\ &= \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}[\phi_i^2] \\ &= \frac{1}{n^2} \sum_{i=1}^n (3 - 1) = \frac{2}{n} \end{aligned} \quad (9)$$

Berechnen wir jetzt die Varianz von den  $\pi_n$ 's. Wir haben

$$\mathbb{V}[\pi_n] = \mathbb{E}[\pi_n^2] - (\mathbb{E}[\pi_n])^2 = \mathbb{E}[\pi_n^2] - 1$$

mit

$$\begin{aligned} \mathbb{E}[\pi_n^2] &= \mathbb{E}\left[\prod_{i=1}^n \phi_i^4\right] \\ &= \prod_{i=1}^n \mathbb{E}[\phi_i^4] \\ &= \prod_{i=1}^n 3 = 3^n \end{aligned}$$

also insgesamt

$$\mathbb{V}[\pi_n] = 3^n - 1 \quad (10)$$

Für  $n = 1$  stimmen  $\mathbb{V}[s_n]$  und  $\mathbb{V}[\pi_n]$  noch überein, aber dann haben wir massiv unterschiedliches Verhalten, die  $\mathbb{V}[s_n]$  konvergieren nach 0, das ist das Gesetz der grossen Zahlen, aber die  $\mathbb{V}[\pi_n]$  gehen nach unendlich, und das auch noch sehr schnell, nämlich exponentiell. Woran liegt das? Für grosse  $n$  ist das Verhalten der  $\pi_n$ 's im wesentlichen dasselbe wie das Verhalten der folgenden  $\xi_n$ 's,

$$\xi_n := \prod_{i=1}^n x_i$$

wobei die  $x_i$  eine Folge von unabhängigen Zufallszahlen sind mit, sagen wir, folgender Definition:

$$x_i := \begin{cases} 4 & \text{mit W'keit } \frac{1}{4} \\ 0 & \text{mit W'keit } \frac{3}{4} \end{cases}$$

Offensichtlich ist

$$\mathbb{E}[x_i] = 4 \times \frac{1}{4} + 0 \times \frac{3}{4} = 1$$

und damit

$$\mathbb{E}[\xi_n] = \prod_{i=1}^n \mathbb{E}[x_i] = 1 \quad (11)$$

Betrachten wir einmal, welche Werte für die  $\xi_n$ 's überhaupt möglich sind. Ist auch nur ein  $x_i$  gleich 0, dann ist auch sofort das gesamte Produkt und damit das  $\xi_n$  gleich 0. Also, einen Wert ungleich 0 bekommen wir nur dann, wenn alle  $x_i$  ungleich 0 sind, und nach Definition der  $x_i$  kommt dann nur die 4 in Frage, jeweils mit W'keit  $\frac{1}{4}$ . Also haben wir:

$$\xi_n = \begin{cases} 4^n & \text{mit W'keit } \frac{1}{4^n} \\ 0 & \text{mit W'keit } 1 - \frac{1}{4^n} \end{cases} \quad (12)$$

Offensichtlich können wir  $\mathbb{E}[\xi_n] = 1$  auch direkt aus Gleichung (12) berechnen, ebenso die Varianz:

$$\begin{aligned} \mathbb{V}[\xi_n] &= \mathbb{E}[(\xi_n - \mathbb{E}[\xi_n])^2] \\ &= \mathbb{E}[(\xi_n - 1)^2] \\ &= (4^n - 1)^2 \times \frac{1}{4^n} + (0 - 1)^2 \times \left(1 - \frac{1}{4^n}\right) \\ &= (4^{2n} - 2 \times 4^n + 1) \times \frac{1}{4^n} + \left(1 - \frac{1}{4^n}\right) \\ &= 4^n - 2 + \frac{1}{4^n} + 1 - \frac{1}{4^n} \\ &= 4^n - 1 \end{aligned} \quad (13)$$

Die Verteilung der  $\xi_n$ 's ist also massiv unsymmetrisch, alle sind 0 bis auf ein einziges, das tritt aber nur mit sehr kleiner W'keit auf, hat dann aber einen sehr grossen Wert, so dass der Erwartungswert also noch eine 1 ist. Dasselbe Verhalten werden wir dann auch bei den  $\pi_n$ 's sehen, fast alle  $\pi_n$ 's sind ziemlich nullig, nur einige sehr wenige sind sehr, sehr gross. Wenn man die aber bei einer Monte Carlo Simulation mit "erwischen" möchte, darf die W'keit für diese grossen  $\pi_n$ 's aber nicht zu klein sein. Bei einer W'keit von  $10^{-9}$  etwa muss man dann um die 1 Milliarde Simulationen machen, damit man einen solchen Wert mal erwischt.

Bei Produkten mit einer hohen Anzahl an Faktoren bietet es sich immer an, den Logarithmus zu betrachten, der macht aus Produkten ja Summen. Also betrachten wir einmal

$$\log(\pi_n) = \log\left(\prod_{i=1}^n \phi_i^2\right) = \sum_{i=1}^n \log(\phi_i^2) =: n \times \ell_n$$

mit der Zufallsvariablen

$$\ell_n := \frac{1}{n} \sum_{i=1}^n \log(\phi_i^2)$$

Das  $\ell_n$  ist dann wieder ein Mittelwert von unabhängigen, identisch verteilten Zufallsvariablen und muss, da seine Varianz gegen 0 konvergiert, gegen seinen Erwartungswert konvergieren,

$$\ell_n \stackrel{n \rightarrow \infty}{\approx} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\log(\phi_i^2)] = \mathbb{E}[\log(\phi^2)]$$

mit

$$\mathbb{E}[\log(\phi^2)] = \int_{-\infty}^{\infty} \log(\phi^2) e^{-\phi^2/2} \frac{d\phi}{\sqrt{2\pi}} \approx -1.27$$

Die Zahl  $-1.27$  wurde dabei mit einer Monte Carlo Simulation berechnet. Wir können also schreiben

$$\pi_n = e^{n \times \ell_n} \stackrel{n \rightarrow \infty}{\approx} e^{n \times \mathbb{E}[\ell_n]} \approx e^{-1.27 \times n} \quad (14)$$

und das ist dann etwa für  $n = 100$  eine sehr kleine Zahl. Plottet man  $\log \pi_n \approx -1.27 \times n$  als Funktion von  $n$ , sollte man also in etwa eine Gerade mit Steigung  $-1.27$  bekommen.

Okay, schauen wir uns die Sachen jetzt in R an.