Chapter 4

The quasi-Monte Carlo method

4.1 Motivations and definitions

Imagine we have a r.v. X which can be simulated by standard methods like the inverse distribution function, Box-Muller simulation method of Gaussian distributions... so that
$$\begin{cases} X = g(U) & \\ U = (U^1, \dots, U^d) \sim \mathcal{U}([0,1]^d) & \rightarrow d \in \mathbb{N}^* \end{cases}$$

So
$$\mathbb{E}(f(X)) = \mathbb{E}(f \circ g(U)) = \int_{[0,1]^d} f \circ g(u^1, \dots, u^d)\, du^1 \dots du^d$$

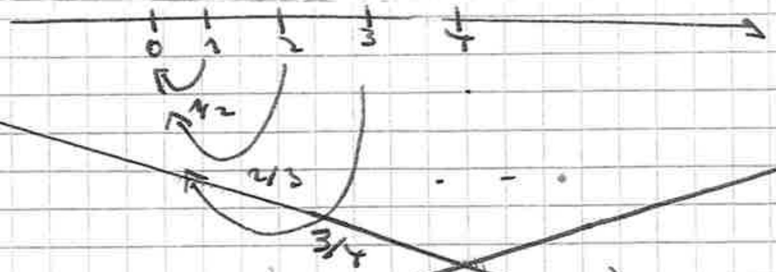
[illegible]

The sequence $(\mu_n)_{n \geq 1}$ weakly converges to μ (notation: $\mu_n \Rightarrow \mu$) if, $\forall f \in \mathcal{C}_b([0,1]^d, \mathbb{R})$
$$\int_{[0,1]^d} f \, d\mu_n \xrightarrow[(n \to \infty)]{} \int_{[0,1]^d} f \, d\mu \quad (w)$$

Proposition: If $\mu_n \Rightarrow \mu$ then (w) above holds for every bounded Borel function $f : ([0,1]^d, \lambda_d) \to \mathbb{R}$ such that $\lambda_d(\mathrm{Disc}(f)) = 0$ where $\mathrm{Disc}(f) = \{x : f \text{ discontinuous at } x\}$

Chapter 4

The quasi-Monte Carlo method



$$\sum_{n\geq 1} \mathbb{P}_0(\tau_0 \geq n) = \sum_{n\geq 1} \sum_{k=n}^{+\infty} \mathbb{P}_0(\tau_0 = k)$$

$$= \sum_{k=1}^{+\infty} \sum_{n=1}^{k} \mathbb{P}_0(\tau_0 = k)$$

Remember that for $(U_n)_{n\geq 1}$ an i.i.d. sequence of r.v. of law $\mathcal{U}([0,1]^d)$, then for every $f \in L^1_{\mathbb{R}}([0,1]^d, \lambda_d)$ Lebesgue

$$\frac{1}{n}\sum_{k=1}^{n} f(U_k(\omega)) \xrightarrow[n\to\infty]{a.s.} \mathbb{E}(f(U_1)) = \int_{[0,1]^d} f(u^1,\dots,u^d)\, du^1 \dots du^d$$

Def. (Weak Convergence)

Let (μ_n) be a sequence of prob. measures on $([0,1]^d, \lambda_d)$ and μ be a prob. measure on $([0,1]^d, \lambda_d)$.
The sequence $(\mu_n)_{n\geq 1}$ weakly converges to μ (notation: $\mu_n \Rightarrow \mu$) if, $\forall f \in \mathcal{C}_b([0,1]^d, \mathbb{R})$

$$\int_{[0,1]^d} f\, d\mu_n \xrightarrow[(n\to\infty)]{} \int_{[0,1]^d} f\, d\mu \quad (w)$$

Proposition: If $\mu_n \Rightarrow \mu$ then (w) above holds for every bounded Borel function $f : ([0,1]^d, \lambda_d) \to \mathbb{R}$ such that $\lambda_d(\mathrm{Disc}(f)) = 0$ where $\mathrm{Disc}(f) = \{x : f \text{ discontinuous at } x\}$

Theorem (Glivenko-Cantelli) If $(U_n)_{n\geq 1}$ is an i.i.d. sequence of law $\mathcal{U}([0,1]^d)$ then

$$\mathbb{P}(d\omega)\text{-a.s.} \quad \frac{1}{n}\sum_{k=1}^{n} \delta_{U_k(\omega)} \underset{(n\to\infty)}{\Rightarrow} \lambda_d$$

Def. A $[0,1]^d$-valued sequence $(\xi_n)_{n\geq 1}$ is uniformly distributed on $[0,1]^d$ if

$$\frac{1}{n}\sum_{k=1}^{n} \delta_{\xi_k} \Rightarrow \lambda_d \quad (n\to\infty)$$

Def. a) We define a partial order on $[0,1]^d$ (noted "$\leq$") by: $x=(x^1,\dots,x^d) \leq y=(y^1,\dots,y^d)$ if $x^i \leq y^i$ $(\forall i)$

b) The box $[\![x,y]\!]$ is defined by:

$$[\![x,y]\!] := \{\xi \in [0,1]^d,\ x \leq \xi \leq y\}$$

($[\![x,y]\!] \neq \emptyset$ iff $x \leq y$)

we can write $[\![x,y]\!] = \prod_{i=1}^{d} [x_i, y_i]$

Proposition (Portmanteau Theorem)

Let $(\xi_n)_{n\geq 1}$ be a $[0,1]^d$-valued sequence. The following assertions are equivalent

i) (ξ_n) is unif. distributed on $[0,1]^d$

ii) $\forall x \in [0,1]^d$, $\frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{[\![0,x]\!]}(\xi_k) \underset{(n\to\infty)}{\longrightarrow} \lambda_d([\![0,x]\!]) = \prod_{i=1}^{d} x_i$

iii) "discrepancy at the origin"

$$D_n^*(\xi) := \sup_{x\in[0,1]^d} \left|\frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{[\![0,x]\!]}(\xi_k) - \prod_{i=1}^{d} x_i\right| \underset{(n\to\infty)}{\longrightarrow} 0$$

Remark:

* If $(U_n)_{n\geq 1}$ is a sequence of i.i.d. r.v. $\sim \mathcal{U}([0,1]^d)$ then $(U_n(\omega))_{n\geq 1}$ is a.s. uniformly distributed.

The law of the iterated logarithm gives us:

$$\limsup_{n\to\infty} \sqrt{\frac{2n}{\log(\log n)}}\, D_n^*(U) \overset{a.s.}{=} 1$$

* We say that a sequence $(\xi_n)_{n\geq 1}$ has a low discrepancy if its discrepancy (at the origin) is asymptotically better than $D_n^*(U)$ ((U_n) above). We have a lower bound $D_n^*(\xi) \geq C_d \left(\frac{\log n}{n}\right)^{\max(\frac{d}{2},1)}$ for infinitely many n

(C_d = constant depending on d)

* We know sequences with a discrepancy of order $\frac{(\log n)^d}{n}$ (this is almost optimal).

Proposition (Koksma - Hlawka's inequality) If g is a finite variation function : $[0,1]^d \longrightarrow \mathbb{R}$ with variation $V(g)$ and (ξ_k) is a sequence with discrepancy $D_n^*(\xi)$ then $\forall n\geq 1$,

$$\left|\frac{1}{n}\sum_{k=1}^{n} g(\xi_k) - \int_{[0,1]^d} g \, d\lambda_d\right| \leq V(g)\times D_n^*(\xi)$$

Remark: We do not define what is a "finite variation function" (it is too complicated). If g is d times continuously differentiable then

$$V(g) = \sum_{k=1}^{d} \sum_{1\leq i_1<\dots<i_k\leq d} \int_{[0,1]^d} \left|\frac{\partial^k g}{\partial x_{i_1}\dots\partial x_{i_k}}(x)\right| dx$$

4.2 Low discrepancy sequences

Van Der Corput sequences

Let $p \in \mathbb{N}^*$ (↳ prime!). Let $n \in \mathbb{N}$, we can always write

$$n = a_0 + a_1 p + \dots + a_n p^n \quad (0 \leq a_i < p \ , 0 \leq i < n \ , \ 0 < a_n)$$

(This is the p-adique decomposition). The Van Der Corput sequence of base p is:

$$\phi_p(n) = \frac{a_0}{p} + \dots + \frac{a_n}{p^{n+1}} .$$

We can understand this definition in the following way: write $n = a_n a_{n-1} \dots a_0$ in base p and then $\phi_p(n) = 0, a_0 a_1 \dots a_n$ (in base p)

Halton sequences

Let $p_1, \dots, p_d$ be the first d prime numbers. The Halton sequence is: $\xi_n^d = (\phi_{p_1}(n), \dots, \phi_{p_d}(n))$ $(\forall n)$

We have $D_n^*(\xi^d) \leq \frac{1}{n} \prod_{i=1}^{d} \frac{p_i \log(p_i n)}{\log(p_i)}$

Faure sequences

In dimension d, we take r to be an odd prime number bigger than d. We define an operator T acting on x written $x = \sum_{k \geq 0} \frac{a_k}{r^{k+1}}$ by:

$$T(x) = \sum_{k \geq 0} \frac{b_k}{r^{k+1}} \quad \text{where } b_k = \sum_{i \geq 0} \binom{k}{i} a_i \mod r$$

We define the Faure sequence by

$$(\forall n) \quad u_n = (\phi_r(n-1), T(\phi_r(n-1)), \dots, T^{d-1}(\phi_r(n-1)))$$

We have: $D_n^*(u) \leq C \frac{(\log n)^d}{n}$

4.3 Pros and Cons of sequences with low discrepancy

The use of sequences with low discrepancy to compute integrals instead of the Monte Carlo method (based on pseudo-random numbers) is known as the Quasi-Monte Carlo method (QMC).

THE PROS

* Convergence is faster than with the MC method (but could be dim.-dependant as we have a $V(g)$ term in the Koksma-Hlawka's inequality).
* Numerical experiments : at least for d up to a few tens, QMC outperforms MC even if the integra-ted function does not have a finite variation

THE CONS

* All the non-asymptotic bounds for the discrepancy at the origin are very poor from a numerical point of view. These bounds cannot be relied on to provide (deterministic) error intervals (compare to MC which automatically provides a confidence interval).

* Finite variation functions are more and more difficult to identify as d increases.