

Low-discrepancy sequence

In mathematics, a **low-discrepancy sequence** is a sequence with the property that for all values of N , its subsequence $\mathbf{x}_1, \dots, \mathbf{x}_N$ has a low discrepancy.

Roughly speaking, the discrepancy of a sequence is low if the proportion of points in the sequence falling into an arbitrary set B is close to proportional to the measure of B , as would happen on average (but not for particular samples) in the case of an equidistributed sequence. Specific definitions of discrepancy differ regarding the choice of B (hyperspheres, hypercubes, etc.) and how the discrepancy for every B is computed (usually normalized) and combined (usually by taking the worst value).

Low-discrepancy sequences are also called **quasirandom** sequences, due to their common use as a replacement of uniformly distributed random numbers. The "quasi" modifier is used to denote more clearly that the values of a low-discrepancy sequence are neither random nor pseudorandom, but such sequences share some properties of random variables and in certain applications such as the quasi-Monte Carlo method their lower discrepancy is an important advantage.

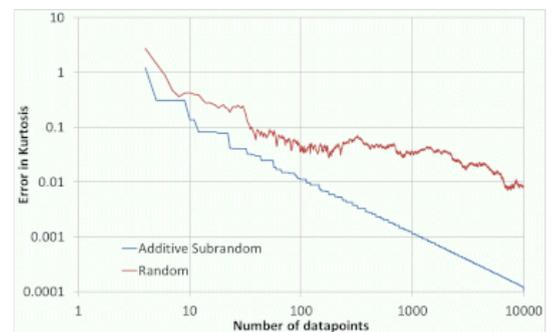
Applications

Quasirandom numbers have an advantage over pure random numbers in that they cover the domain of interest quickly and evenly.

Two useful applications are in finding the characteristic function of a probability density function, and in finding the derivative function of a deterministic function with a small amount of noise. Quasirandom numbers allow higher-order moments to be calculated to high accuracy very quickly.

Applications that don't involve sorting would be in finding the mean, standard deviation, skewness and kurtosis of a statistical distribution, and in finding the integral and global maxima and minima of difficult deterministic functions. Quasirandom numbers can also be used for providing starting points for deterministic algorithms that only work locally, such as Newton–Raphson iteration.

Quasirandom numbers can also be combined with search algorithms. With a search algorithm, quasirandom numbers can be used to find the mode, median, confidence intervals and cumulative distribution of a statistical distribution, and all local minima and all solutions of deterministic functions.



Error in estimated kurtosis as a function of number of datapoints. 'Additive quasirandom' gives the maximum error when $c = (\sqrt{5} - 1)/2$. 'Random' gives the average error over six runs of random numbers, where the average is taken to reduce the magnitude of the wild fluctuations

Low-discrepancy sequences in numerical integration

Various methods of numerical integration can be phrased as approximating the integral of a function f in some interval, e.g. $[0,1]$, as the average of the function evaluated at a set $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ in that interval:

$$\int_0^1 f(u) du \approx \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i).$$

If the points are chosen as $\mathbf{x}_i = i/N$, this is the *rectangle rule*. If the points are chosen to be randomly (or pseudorandomly) distributed, this is the *Monte Carlo method*. If the points are chosen as elements of a low-discrepancy sequence, this is the *quasi-Monte Carlo method*. A remarkable result, the **Koksma–Hlawka inequality** (stated below), shows that the error of such a method can be bounded by the product of two terms, one of which depends only on f , and the other one is the discrepancy of the set $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$.

It is convenient to construct the set $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ in such a way that if a set with $N + 1$ elements is constructed, the previous N elements need not be recomputed. The rectangle rule uses points set which have low discrepancy, but in general the elements must be recomputed if N is increased. Elements need not be recomputed in the random Monte Carlo method if N is increased, but the point sets do not have minimal discrepancy. By using low-discrepancy sequences we aim for low discrepancy and no need for recomputations, but actually low-discrepancy sequences can only be incrementally good on discrepancy if we allow no recomputation.

Definition of discrepancy

The *discrepancy* of a set $P = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ is defined, using Niederreiter's notation, as

$$D_N(P) = \sup_{B \in \mathcal{J}} \left| \frac{A(B; P)}{N} - \lambda_s(B) \right|$$

where λ_s is the s -dimensional Lebesgue measure, $A(\mathbf{B}; \mathbf{P})$ is the number of points in \mathbf{P} that fall into \mathbf{B} , and \mathbf{J} is the set of s -dimensional intervals or boxes of the form

$$\prod_{i=1}^s [a_i, b_i) = \{\mathbf{x} \in \mathbf{R}^s : a_i \leq x_i < b_i\}$$

where $0 \leq a_i < b_i \leq 1$.

The *star-discrepancy* $D_N^*(\mathbf{P})$ is defined similarly, except that the supremum is taken over the set \mathbf{J}^* of rectangular boxes of the form

$$\prod_{i=1}^s [0, u_i)$$

where u_i is in the half-open interval $[0, 1)$.

The two are related by

$$D_N^* \leq D_N \leq 2^s D_N^*.$$

Note: With these definitions, discrepancy represents the worst-case or maximum point density deviation of a uniform set. However, also other error measures are meaningful, leading to other definitions and variation measures. For instance, L^2 -discrepancy or modified centered L^2 -discrepancy are also used intensively to compare the quality of uniform point sets. Both are much easier to calculate for large N and s .

The Koksma–Hlawka inequality

Let \bar{I}^s be the s -dimensional unit cube, $\bar{I}^s = [0, 1] \times \cdots \times [0, 1]$. Let f have bounded variation $V(f)$ on \bar{I}^s in the sense of Hardy and Krause. Then for any $\mathbf{x}_1, \dots, \mathbf{x}_N$ in $I^s = [0, 1)^s = [0, 1) \times \cdots \times [0, 1)$,

$$\left| \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) - \int_{\bar{I}^s} f(u) du \right| \leq V(f) D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N).$$

The Koksma–Hlawka inequality is sharp in the following sense: For any point set $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ in I^s and any $\varepsilon > 0$, there is a function f with bounded variation and $V(f) = 1$ such that

$$\left| \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) - \int_{\bar{I}^s} f(u) du \right| > D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) - \varepsilon.$$

Therefore, the quality of a numerical integration rule depends only on the discrepancy $D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N)$.

The formula of Hlawka–Zaremba

Let $D = \{1, 2, \dots, d\}$. For $\emptyset \neq u \subseteq D$ we write

$$d\mathbf{x}_u := \prod_{j \in u} dx_j$$

and denote by $(\mathbf{x}_u, \mathbf{1})$ the point obtained from x by replacing the coordinates not in u by $\mathbf{1}$. Then

$$\frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) - \int_{\bar{I}^s} f(u) du = \sum_{\emptyset \neq u \subseteq D} (-1)^{|u|} \int_{[0,1]^{|u|}} \text{disc}(\mathbf{x}_u, \mathbf{1}) \frac{\partial^{|u|}}{\partial \mathbf{x}_u} f(\mathbf{x}_u, \mathbf{1}) d\mathbf{x}_u,$$

where $\text{disc}(z) = \frac{1}{N} \sum_{i=1}^N \prod_{j=1}^d 1_{[0, z_j)}(x_{i,j}) - \prod_{j=1}^d z_j$ is the discrepancy function.

The L^2 version of the Koksma–Hlawka inequality

Applying the Cauchy–Schwarz inequality for integrals and sums to the Hlawka–Zaremba identity, we obtain an L^2 version of the Koksma–Hlawka inequality:

$$\left| \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) - \int_{\bar{I}^s} f(u) du \right| \leq \|f\|_d \text{disc}_d(\{\mathbf{t}_i\}),$$

where

$$\text{disc}_d(\{t_i\}) = \left(\sum_{\emptyset \neq u \subseteq D} \int_{[0,1]^{|u|}} \text{disc}(\mathbf{x}_u, 1)^2 d\mathbf{x}_u \right)^{1/2}$$

and

$$\|f\|_d = \left(\sum_{u \subseteq D} \int_{[0,1]^{|u|}} \left| \frac{\partial^{|u|}}{\partial \mathbf{x}_u} f(\mathbf{x}_u, 1) \right|^2 d\mathbf{x}_u \right)^{1/2}.$$

L^2 discrepancy has a high practical importance because fast explicit calculations are possible for a given point set. This way it is easy to create point set optimizers using L^2 discrepancy as criteria.

The Erdős–Turán–Koksma inequality

It is computationally hard to find the exact value of the discrepancy of large point sets. The Erdős–Turán–Koksma inequality provides an upper bound.

Let $\mathbf{x}_1, \dots, \mathbf{x}_N$ be points in I^s and H be an arbitrary positive integer. Then

$$D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \leq \left(\frac{3}{2} \right)^s \left(\frac{2}{H+1} + \sum_{0 < \|h\|_\infty \leq H} \frac{1}{r(h)} \left| \frac{1}{N} \sum_{n=1}^N e^{2\pi i \langle h, \mathbf{x}_n \rangle} \right| \right)$$

where

$$r(h) = \prod_{i=1}^s \max\{1, |h_i|\} \quad \text{for } h = (h_1, \dots, h_s) \in \mathbb{Z}^s.$$

The main conjectures

Conjecture 1. There is a constant c_s depending only on the dimension s , such that

$$D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \geq c_s \frac{(\ln N)^{s-1}}{N}$$

for any finite point set $\mathbf{x}_1, \dots, \mathbf{x}_N$.

Conjecture 2. There is a constant c'_s depending only on s , such that: $D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \geq c'_s \frac{(\ln N)^s}{N}$

for infinite number of N for any infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots$.

These conjectures are equivalent. They have been proved for $s \leq 2$ by W. M. Schmidt. In higher dimensions, the corresponding problem is still open. The best-known lower bounds are due to Michael Lacey and collaborators.

Lower bounds

Let $s = 1$. Then

$$D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \geq \frac{1}{2N}$$

for any finite point set $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$.

Let $s = 2$. W. M. Schmidt proved that for any finite point set $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$,

$$D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \geq C \frac{\log N}{N}$$

where

$$C = \max_{a \geq 3} \frac{1}{16} \frac{a-2}{a \log a} = 0.023335 \dots$$

For arbitrary dimensions $s > 1$, K. F. Roth proved that

$$D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \geq \frac{1}{2^{4s}} \frac{1}{((s-1)\log 2)^{\frac{s-1}{2}}} \frac{\log^{\frac{s-1}{2}} N}{N}$$

for any finite point set $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$. Jozef Beck ^[1] established a double log improvement of this result in three dimensions. This was improved by D. Bilyk and M. T. Lacey to a power of a single logarithm. The best known bound for $s > 2$ is due D. Bilyk and M. T. Lacey and A. Vagharshakyan.^[2] For $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ there is a $t = 0$ so that

$$D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \geq t \frac{\log^{\frac{s-1}{2}+t} N}{N}$$

for any finite point set $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$.

Construction of low-discrepancy sequences

Because any distribution of random numbers can be mapped onto a uniform distribution, and quasirandom numbers are mapped in the same way, this article only concerns generation of quasirandom numbers on a multidimensional uniform distribution.

There are constructions of sequences known such that

$$D_N^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \leq C \frac{(\ln N)^s}{N}$$

where C is a certain constant, depending on the sequence. After Conjecture 2, these sequences are believed to have the best possible order of convergence. Examples below are the van der Corput sequence, the Halton sequences, and the Sobol' sequences. One general limitation is that construction methods can usually only guarantee the order of convergence. Practically, low discrepancy can be only achieved if N is large enough, and for large given s this minimum N can be very large. This means running a Monte-Carlo analysis with e.g. $s = 20$ variables and $N = 1000$ points from a low-discrepancy sequence generator may offer only a very minor accuracy improvement.

Random numbers

Sequences of quasirandom numbers can be generated from random numbers by imposing a negative correlation on those random numbers. One way to do this is to start with a set of random numbers r_i on $[0, 0.5)$ and construct quasirandom numbers s_i which are uniform on $[0, 1)$ using:

$$s_i = r_i \text{ for } i \text{ odd and } s_i = 0.5 + r_i \text{ for } i \text{ even.}$$

A second way to do it with the starting random numbers is to construct a random walk with offset 0.5 as in:

$$s_i = s_{i-1} + 0.5 + r_i \pmod{1}.$$

That is, take the previous quasirandom number, add 0.5 and the random number, and take the result modulo 1.

For more than one dimension, Latin squares of the appropriate dimension can be used to provide offsets to ensure that the whole domain is covered evenly.

Additive recurrence

For any irrational α , the sequence

$$s_n = \{s_0 + n\alpha\}$$

has discrepancy tending to $1/N$. Note that the sequence can be defined recursively by

$$s_{n+1} = (s_n + \alpha) \pmod{1}.$$

A good value of α gives lower discrepancy than a sequence of independent uniform random numbers.

The discrepancy can be bounded by the approximation exponent of α . If the approximation exponent is μ , then for any $\epsilon > 0$, the following bound holds:^[3]

$$D_N((s_n)) = O_\epsilon(N^{-1/(\mu-1)+\epsilon}).$$

By the Thue–Siegel–Roth theorem, the approximation exponent of any irrational algebraic number is 2, giving a bound of $N^{-1+\epsilon}$ above.

The recurrence relation above is similar to the recurrence relation used by a linear congruential generator, a poor-quality pseudorandom number generator.^[4]

$$r_i = (ar_{i-1} + c) \pmod{m}$$

For the low discrepancy additive recurrence above, a and m are chosen to be 1. Note, however, that this will not generate independent random numbers, so should not be used for purposes requiring independence.

The value of c with lowest discrepancy is the fractional part of the golden ratio:^[5]

$$c = \frac{\sqrt{5} - 1}{2} = \varphi - 1 \approx 0.618034.$$

Another value that is nearly as good is the fractional part of the silver ratio, which is the fractional part of the square root of 2:

$$c = \sqrt{2} - 1 \approx 0.414214.$$

In more than one dimension, separate quasirandom numbers are needed for each dimension. A convenient set of values that are used, is the square roots of primes from two up, all taken modulo 1:

$$c = \sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}, \sqrt{11}, \dots$$

However, a set of values based on the generalised golden ratio has been shown to produce more evenly distributed points.^[6]

The list of pseudorandom number generators lists methods for generating independent pseudorandom numbers. *Note:* In few dimensions, recursive recurrence leads to uniform sets of good quality, but for larger s (like $s > 8$) other point set generators can offer much lower discrepancies.

van der Corput sequence

Let

$$n = \sum_{k=0}^{L-1} d_k(n) b^k$$

be the b -ary representation of the positive integer $n \geq 1$, i.e. $0 \leq d_k(n) < b$. Set

$$g_b(n) = \sum_{k=0}^{L-1} d_k(n) b^{-k-1}.$$

Then there is a constant C depending only on b such that $(g_b(n))_{n \geq 1}$ satisfies

$$D_N^*(g_b(1), \dots, g_b(N)) \leq C \frac{\log N}{N},$$

where D_N^* is the star discrepancy.

Halton sequence

The Halton sequence is a natural generalization of the van der Corput sequence to higher dimensions. Let s be an arbitrary dimension and b_1, \dots, b_s be arbitrary coprime integers greater than 1. Define

$$\mathbf{x}(n) = (g_{b_1}(n), \dots, g_{b_s}(n)).$$

Then there is a constant C depending only on b_1, \dots, b_s , such that sequence $\{\mathbf{x}(n)\}_{n \geq 1}$ is a s -dimensional sequence with

$$D_N^*(\mathbf{x}(1), \dots, \mathbf{x}(N)) \leq C' \frac{(\log N)^s}{N}.$$

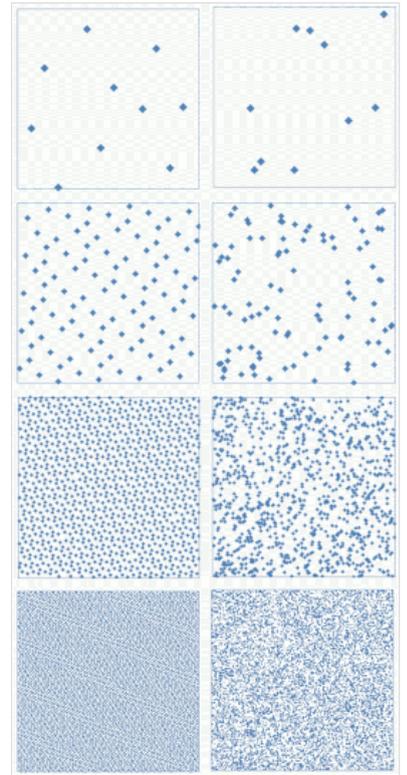
Hammersley set

Let b_1, \dots, b_{s-1} be coprime positive integers greater than 1. For given s and N , the s -dimensional Hammersley set of size N is defined by^[7]

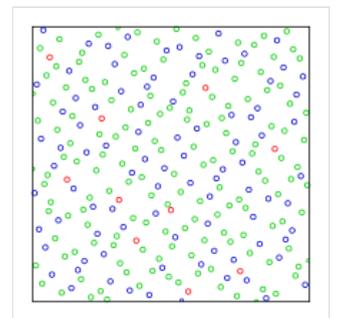
$$\mathbf{x}(n) = \left(g_{b_1}(n), \dots, g_{b_{s-1}}(n), \frac{n}{N} \right)$$

for $n = 1, \dots, N$. Then

$$D_N^*(\mathbf{x}(1), \dots, \mathbf{x}(N)) \leq C \frac{(\log N)^{s-1}}{N}$$



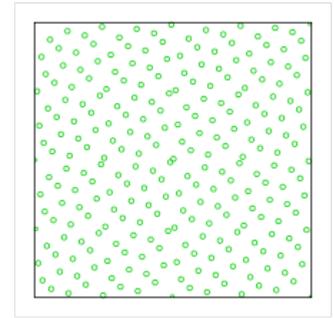
Coverage of the unit square. Left for additive quasirandom numbers with $c = 0.5545497\dots, 0.308517\dots$. Right for random numbers. From top to bottom, 10, 100, 1000, 10000 points.



First 256 points of the (2,3) Halton sequence

where C is a constant depending only on b_1, \dots, b_{s-1} .

Note: The formulas show that the Hammersley set is actually the Halton sequence, but we get one more dimension for free by adding a linear sweep. This is only possible if N is known upfront. A linear set is also the set with lowest possible one-dimensional discrepancy in general. Unfortunately, for higher dimensions, no such "discrepancy record sets" are known. For $s = 2$, most low-discrepancy point set generators deliver at least near-optimum discrepancies.



2D Hammersley set of size 256

Sobol sequence

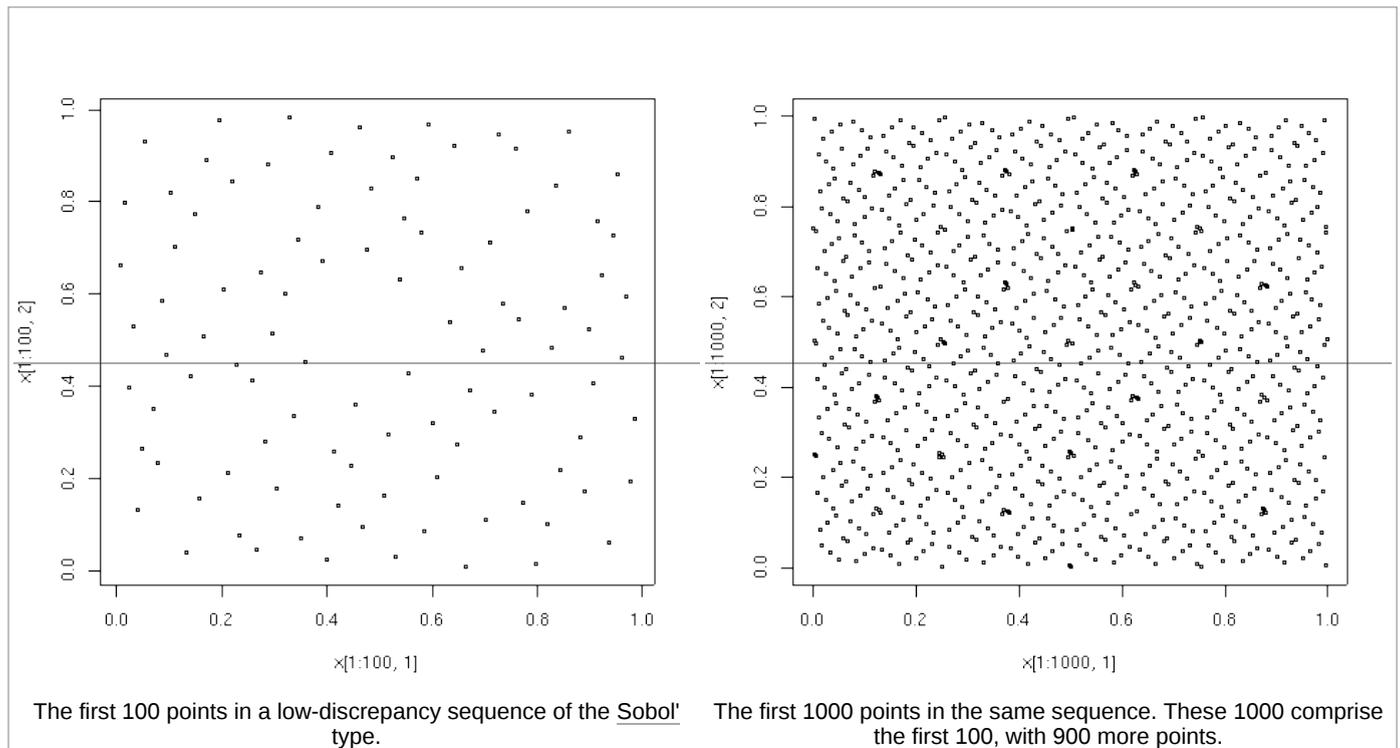
The Antonov–Saleev variant of the Sobol' sequence generates numbers between zero and one directly as binary fractions of length w , from a set of w special binary fractions, $V_i, i = 1, 2, \dots, w$ called direction numbers. The bits of the Gray code of i , $G(i)$, are used to select direction numbers. To get the Sobol' sequence value s_i take the exclusive or of the binary value of the Gray code of i with the appropriate direction number. The number of dimensions required affects the choice of V_i .

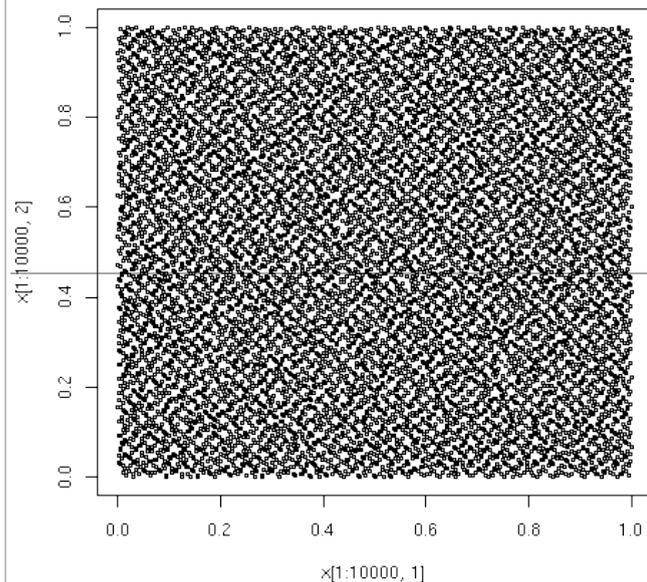
Poisson disk sampling

Poisson disk sampling is popular in video games to rapidly place objects in a way that appears random-looking but guarantees that every two points are separated by at least the specified minimum distance.^[8] This does not guarantee low discrepancy (as e. g. Sobol'), but at least a significantly lower discrepancy than pure random sampling. The goal of these sampling patterns is based on frequency analysis rather than discrepancy, a type of so-called "blue noise" patterns.

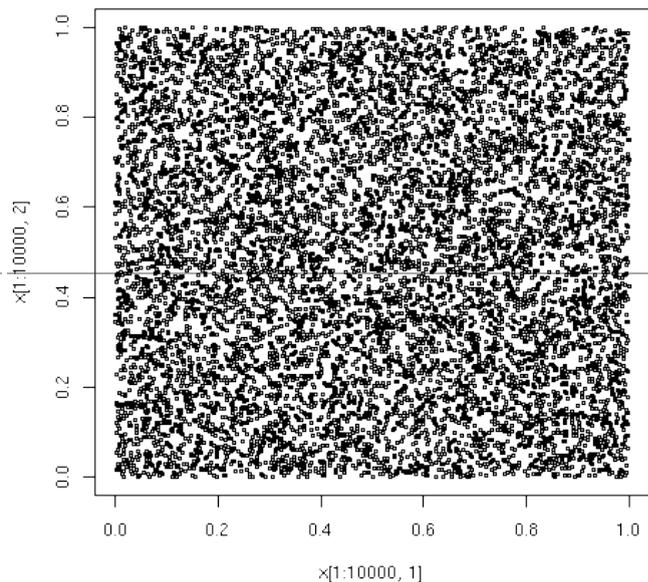
Graphical examples

The points plotted below are the first 100, 1000, and 10000 elements in a sequence of the Sobol' type. For comparison, 10000 elements of a sequence of pseudorandom points are also shown. The low-discrepancy sequence was generated by TOMS algorithm 659.^[9] An implementation of the algorithm in Fortran is available from Netlib.





The first 10000 points in the same sequence. These 10000 comprise the first 1000, with 9000 more points.



For comparison, here are the first 10000 points in a sequence of uniformly distributed pseudorandom numbers. Regions of higher and lower density are evident.

See also

- [Discrepancy theory](#)
- [Markov chain Monte Carlo](#)
- [Quasi-Monte Carlo method](#)
- [Sparse grid](#)
- [Systematic sampling](#)

Notes

1. Beck, József (1989). "A two-dimensional van Aardenne-Ehrenfest theorem in irregularities of distribution" (<https://eudml.org/doc/89992>). *Compositio Mathematica*. **72** (3): 269–339. MR 1032337 (<https://mathscinet.ams.org/mathscinet-getitem?mr=1032337>). S2CID 125940424 (<https://api.semanticscholar.org/CorpusID:125940424>). Zbl 0691.10041 (<https://zbmath.org/?format=complete&q=an:0691.10041>).
2. Bilyk, Dmitriy; Lacey, Michael T.; Vagharshakyan, Armen (2008). "On the Small Ball Inequality in all dimensions" (<https://doi.org/10.1016%2Fj.jfa.2007.09.010>). *Journal of Functional Analysis*. **254** (9): 2470–2502. arXiv:0705.4619 (<https://arxiv.org/abs/0705.4619>). doi:10.1016/j.jfa.2007.09.010 (<https://doi.org/10.1016%2Fj.jfa.2007.09.010>). S2CID 14234006 (<https://api.semanticscholar.org/CorpusID:14234006>).
3. Kuipers & Niederreiter 2005, p. 123
4. Knuth, Donald E. "Chapter 3 – Random Numbers". *The Art of Computer Programming*. Vol. 2.
5. Skarupke, Malte (16 June 2018). "Fibonacci Hashing: The Optimization that the World Forgot" (<https://probablydance.com/2018/06/16/fibonacci-hashing-the-optimization-that-the-world-forgot-or-a-better-alternative-to-integer-modulo/>). "One property of the Golden Ratio is that you can use it to subdivide any range roughly evenly ... if you don't know ahead of time how many steps you're going to take"
6. Roberts, Martin (2018). "The Unreasonable Effectiveness of Quasirandom Sequences" (<https://web.archive.org/web/20250301162105/https://extremelearning.com.au/unreasonable-effectiveness-of-quasirandom-sequences/>).
7. Hammersley, J. M.; Handscomb, D. C. (1964). *Monte Carlo Methods*. doi:10.1007/978-94-009-5819-7 (<https://doi.org/10.1007%2F978-94-009-5819-7>). ISBN 978-94-009-5821-0.
8. Herman Tulleken. Tulleken, Herman (March 2008). "Poisson Disk Sampling" (<http://devmag.org.za/2009/05/03/poisson-disk-sampling/>). *Dev.Mag.* No. 21. pp. 21–25.
9. Bratley, Paul; Fox, Bennett L. (1988). "Algorithm 659" (<https://doi.org/10.1145%2F42288.214372>). *ACM Transactions on Mathematical Software*. **14**: 88–100. doi:10.1145/42288.214372 (<https://doi.org/10.1145%2F42288.214372>). S2CID 17325779 (<https://api.semanticscholar.org/CorpusID:17325779>).

References

- Dick, Josef; Pillichshammer, Friedrich (2010). *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*.

Cambridge University Press. ISBN 978-0-521-19159-3.

- Kuipers, L.; Niederreiter, H. (2005), *Uniform distribution of sequences*, Dover Publications, ISBN 0-486-45019-8
- Harald Niederreiter (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. Society for Industrial and Applied Mathematics. ISBN 0-89871-295-5.
- Drmota, Michael; Tichy, Robert F. (1997). *Sequences, Discrepancies and Applications*. Lecture Notes in Math. Vol. 1651. Springer. ISBN 3-540-62606-9.
- Press, William H.; Flannery, Brian P.; Teukolsky, Saul A.; Vetterling, William T. (1992). *Numerical Recipes in C* (2nd ed.). Cambridge University Press. see Section 7.7 for a less technical discussion of low-discrepancy sequences. ISBN 0-521-43108-5.

External links

- Collected Algorithms of the ACM (<http://calgo.acm.org/>) (See algorithms 647, 659, and 738.)
 - Quasi-Random Sequences (<https://www.gnu.org/software/gsl/doc/html/qrng.html>) from the GNU Scientific Library
 - Quasi-random sampling subject to constraints (<http://finmathblog.blogspot.com/2013/09/quasi-random-sampling-subject-to-linear.html>) at FinancialMathematics.Com
 - C++ generator of Sobol' sequence (http://kirillsprograms.com/top_3Sobol.php)
 - SciPy QMC API Reference: [scipy.stats.qmc](https://scipy.github.io/devdocs/reference/stats.qmc.html) (<https://scipy.github.io/devdocs/reference/stats.qmc.html>)
-

Retrieved from "https://en.wikipedia.org/w/index.php?title=Low-discrepancy_sequence&oldid=1280247105"