Constructive Sampling and Spherical-Bessel Expansion of Densities

Abstract

We present a constructive method for generating samples from any continuous one-dimensional distribution using a binary tree of conditional Bernoulli choices, and we expand the resulting intermediate density on $[0,1]$ in a spherical-Bessel basis. We correct the independence assumption often made in Bernoulli-bit models, state the correct inversion formula for transforming densities, and give precise orthogonality relations and normalizations for spherical-Bessel functions. We also provide convergence conditions and practical coefficient estimation formulas, along with positivity enforcement strategies.

1 Introduction

Given a target continuous density f_X on $\mathbb{R}$ with CDF F_X, a common method to simulate X is to transform uniform random variables $U \in [0,1]$ via $X = F_X^{-1}(U)$. Here, we propose a constructive sampling scheme that builds U via a binary expansion with prefix-dependent branching probabilities so that U has any desired distribution on $[0,1]$. We then expand f_U in a spherical-Bessel basis, estimate the coefficients from samples, and invert the mapping to obtain f_X.

2 Constructive sampling on $[0,1]$

Let μ be a Borel probability measure on $[0,1]$ with density f_U. We define the dyadic partition at depth k as

$$I_{j,k} = \left[\frac{j}{2^k}, \frac{j+1}{2^k}\right), \quad j = 0, \ldots, 2^k - 1.$$

Let $m_{j,k} = \mu(I_{j,k})$.

2.1 Prefix-dependent branching probabilities

We generate U as

$$U_N = \sum_{k=1}^{N} B_k 2^{-k},$$

where at each step k we condition on the previously chosen bits (prefix) and set the probability

$$p_k(\text{prefix}) = \frac{m_{j1,k}}{m_{j0,k} + m_{j1,k}},$$

where j indexes the current prefixs dyadic interval at level $k-1$ and $j0$, $j1$ denote the left/right children intervals at level k. We then draw $B_k \sim \text{Bernoulli}\big(p_k(\text{prefix})\big)$.

Theorem 1. *For any μ on $[0,1]$, the above construction yields $\mathbb{P}(U_N \in I_{j,N}) = m_{j,N}$ for all j. If f_U is L-Lipschitz, then*

$$\|\mathcal{L}(U_N) - \mu\|_{\mathrm{TV}} \leq \frac{L}{2}\, 2^{-N}.$$

Proof. The exact dyadic mass property follows by induction on the binary tree. The total variation bound follows from approximating f_U by its average on each dyadic cell, which incurs L^1 error at most $L2^{-N}$; TV is half the L^1 error. □

2.2 Dyadic rational uniqueness

Numbers of the form $j/2^N$ have two binary expansions. We fix the convention of using the terminating expansion (ending in zeros). This affects a set of measure zero and has no impact for absolutely continuous μ.

3 Mapping to the target variable X

Let $\Phi : \mathbb{R} \to [0,1]$ be a differentiable, strictly monotone mapping. For $u = \Phi(x)$, the change-of-variables formula gives

$$f_X(x) = f_U(\Phi(x))\, \Phi'(x).$$

If $\Phi = F_X$, then $f_U \equiv 1$ and $f_X(x) = F_X'(x)$ as expected.

4 Spherical-Bessel expansion on $[0,1]$

We expand f_U in spherical-Bessel functions:

$$f_U(u) \approx \sum_{n=1}^{N} a_n\, j_\ell(\alpha_{\ell,n} u),$$

where j_ℓ is the spherical-Bessel function of order ℓ and $\alpha_{\ell,n}$ is its n-th positive zero.

4.1 Orthogonality and normalization

The set $\{j_\ell(\alpha_{\ell,n} u)\}_{n\geq 1}$ is orthogonal in the weighted space $L^2([0,1], u^2\,du)$:

$$\int_0^1 u^2 j_\ell(\alpha_{\ell,n} u) j_\ell(\alpha_{\ell,m} u)\, du = \frac{1}{2}[j_{\ell+1}(\alpha_{\ell,n})]^2 \delta_{nm}.$$

Thus the coefficients are

$$a_n = \frac{2}{[j_{\ell+1}(\alpha_{\ell,n})]^2} \int_0^1 f_U(u) j_\ell(\alpha_{\ell,n} u) u^2\, du.$$

4.2 Convergence

If $f_U \in L^2([0,1], u^2\,du)$, then the truncated series converges to f_U in the weighted L^2 norm. Convergence in unweighted $L^2([0,1])$ requires additional regularity near $u = 0$.

5 Estimating coefficients from samples

Given i.i.d. samples $\{U_i\}$ from f_U, an unbiased Monte Carlo estimator is

$$\widehat{a}_n = \frac{2}{[j_{\ell+1}(\alpha_{\ell,n})]^2} \cdot \frac{1}{M} \sum_{i=1}^{M} j_\ell(\alpha_{\ell,n} U_i) U_i^2.$$

Alternatively, estimate f_U via a kernel density estimator on $[0,1]$ and compute the projection integrals by quadrature.

6 Positivity and renormalization

Spectral truncation can yield negative regions in the reconstructed density. Remedies include:

- Nonnegative least squares fitting in the chosen basis.
- Expanding $\log f_U$ and exponentiating.
- Perturbative correction and renormalization.

7 Conclusion

We corrected the binary-expansion-based sampling model to use prefix-dependent branching, fixed the inversion formula, clarified the spherical-Bessel projection, and provided practical algorithms. The method yields exact finite-level matching of dyadic masses and spectral convergence under smoothness assumptions.

A Exact binary tree sampler pseudocode

```
Input: Target CDF F_U on [0,1], depth N
prefix = 0
for k = 1..N:
    left  = prefix / 2^(k-1)
    right = (prefix+1) / 2^(k-1)
    mid   = (2*prefix + 1) / 2^k
    mass_interval = F_U(right) - F_U(left)
    if mass_interval == 0:
        prob1 = 0
    else:
        prob1 = (F_U(right) - F_U(mid)) / mass_interval
    sample B_k ~ Bernoulli(prob1)
    if B_k == 1:
        prefix = 2*prefix + 1
    else:
        prefix = 2*prefix
U_N = prefix / 2^N
```

B Monte Carlo coefficient estimation

Given samples U_i from f_U:

$$\widehat{a}_n = \frac{2}{[j_{\ell+1}(\alpha_{\ell,n})]^2} \cdot \frac{1}{M} \sum_{i=1}^{M} j_\ell(\alpha_{\ell,n} U_i) U_i^2.$$

This is unbiased for a_n and converges at rate $O(M^{-1/2})$ in RMS.