Approximating Any Continuous PDF Using Infinite Bernoulli Expansions with Constructive Parameterization

Abstract

This document presents a mathematical and computational framework for approximating any continuous probability density function (PDF) on the real line by using an infinite sum of Bernoulli distributions. In addition to the theoretical result, we detail a constructive method for choosing the Bernoulli success probabilities $\{p_k\}$, enabling implementation of this approximation algorithm in practice.

1 Theorem

Let $f_X(x)$ be a continuous PDF on $\mathbb{R}$, with cumulative distribution function $F_X(x)$, assumed to be strictly increasing and continuous. Then:

(1) Define $U = F_X(X)$. Then $U \sim \text{Uniform}(0, 1)$.

(2) Any target distribution on $[0, 1]$ (e.g., U) can be approximated in total variation by a binary expansion:

$$U = \sum_{k=1}^{\infty} B_k \cdot 2^{-k}, \quad B_k \sim \text{Bern}(p_k)$$

(3) Applying the inverse transform $X = F_X^{-1}(U)$ yields samples approximately distributed according to $f_X(x)$.

2 Proof Sketch

(1) **Transform to [0,1]**: Define $U = F_X(X)$. Since F_X is continuous and strictly increasing, this maps $x \in \mathbb{R}$ to $u \in [0, 1]$.

(2) **Binary Expansion Representation**: Any $u \in [0, 1]$ can be represented in binary as $u = \sum_{k=1}^{\infty} b_k 2^{-k}$, where $b_k \in \{0, 1\}$. If $b_k \sim \text{Bern}(0.5)$, the result is uniform. More generally, a non-uniform distribution can be induced by modifying the bias p_k of each Bernoulli trial.

(3) **Pushforward Distribution**: Choose $\{p_k\}$ such that the binary expansion $U = \sum_{k=1}^{\infty} B_k 2^{-k}$ approximates a target distribution on $[0, 1]$ in total variation. Apply $X = F_X^{-1}(U)$ to recover the desired f_X.

3 Constructing $\{p_k\}$ for Arbitrary Distributions

We now present methods to construct the sequence $\{p_k\}$ for approximating a known target PDF $f_U(u)$ on $[0, 1]$.

Method 1: Moment Matching (approximate)

- Choose a small number of terms N.
- Solve for $p_k \in [0,1]$ such that the first N moments of the distribution $U = \sum_{k=1}^{N} B_k 2^{-k}$ match the corresponding moments of $f_U(u)$.
- Use numerical optimization (e.g., least-squares fit of moments).

Method 2: Quantile Matching (preferred)

- Let $q_j = P(U \leq j/2^N)$ be the CDF of the target at dyadic fractions.
- Define a probability tree corresponding to all 2^N binary strings of length N.
- Optimize $\{p_k\}$ such that the output distribution of $\sum_{k=1}^{N} B_k 2^{-k}$ matches $\{q_j\}$.
- This reduces to fitting a binary tree with branching probabilities to a desired distribution (similar to Huffman-style coding).

4 Algorithm (with Constructive p_k Selection)

Given a continuous PDF $f_X(x)$ and its invertible CDF F_X:

Step 1: Transform target PDF to unit interval: define $f_U(u)$ as the desired distribution on $[0,1]$. If unknown, take $f_U(u) = 1$.

Step 2: Choose bit resolution: choose number of bits N for binary approximation.

Step 3: Compute $\{p_k\}$: use moment or quantile matching to compute $p_1, \ldots, p_N$.

Step 4: Sample binary digits: for each trial, sample $B_k \sim \text{Bern}(p_k)$.

Step 5: Construct U: $U = \sum_{k=1}^{N} B_k \cdot 2^{-k}$.

Step 6: Transform to X: $X = F_X^{-1}(U)$.

Step 7: Repeat to generate samples from $f_X(x)$.

5 Notes

- If $p_k = 0.5$ for all k, then $U \sim \text{Uniform}(0,1)$.
- For non-uniform f_U, numerical computation of p_k is required.
- The approximation improves exponentially in N, subject to smoothness of f_U.

6 Convergence and Error Bounds

Let $U_N := \sum_{k=1}^{N} B_k \cdot 2^{-k}$ denote the finite sum approximation. If f_U is Lipschitz continuous, then:

$$\|\mathbb{P}_{U_N} - f_U\|_{\text{TV}} \leq C \cdot 2^{-N}$$

for some constant $C > 0$, implying **exponential convergence** in N. This result holds under regularity assumptions on f_U such as bounded variation or Lipschitz continuity.

7 Discussion on Assumptions

This construction assumes:

- F_X is continuous and strictly increasing ensuring invertibility and valid transformation to U.
- The target density f_U is square-integrable and admits a binary expansion approximation.
- The method is limited to 1D; higher-dimensional generalizations are nontrivial.

8 Extensions

- Alternative bases (e.g., ternary, Legendre) could be used in place of binary expansions.
- Adaptive bit-length strategies can improve efficiency for sharply peaked distributions.
- Extensions to multivariate distributions may involve coupled binary expansions or tree-based methods.

9 References

- Knuth, D. E., and Yao, A. C. (1976). *The Complexity of Nonuniform Random Number Generation.*
- Devroye, L. (1986). *Non-Uniform Random Variate Generation.*
- Li, M., and Vitnyi, P. (2008). *An Introduction to Kolmogorov Complexity and Its Applications.*
- Peres, Y. (1992). *Iterating von Neumann's Procedure for Extracting Random Bits.*
- Billingsley, P. (1995). *Probability and Measure.*