A

On Simulating a Class of Bernstein Polynomials

VINEET GOYAL, Columbia University
KARL SIGMAN, Columbia University

Given a black box that generates independent Bernoulli samples with an unknown bias p, we consider the problem of simulating a Bernoulli random variable with bias $f(p)$ (where f is a given function) using a finite (computable in advance) number of independent Bernoulli samples from the black box. We show that this is possible if and only if f is a Bernstein polynomial with coefficients between 0 and 1, and we explicitly give the algorithm. Our results differ from Keane and O'Brien (1994) in that our goal is more modest/stringent, since we are considering algorithms that use a *finite* number of samples as opposed to allowing a random number (such as in acceptance rejection algorithms).

General Terms: Simulation; Bernstein polynomials

1. INTRODUCTION

In this paper, we consider the problem of simulating a Bernoulli random variable with success probability $f(p)$ from independent Bernoulli samples of an unknown bias $p \in \mathcal{P} \subset (0,1)$ where f is a function such that $0 < f(p) < 1$ for all $p \in \mathcal{P}$. In [Keane and O'Brien 1994], it is shown that this is possible if and only if

(1) f is continuous on $\mathcal{P}$, and
(2) either f is constant, or there exists an integer $n \geq 1$ such that $\min(f(p), 1-f(p)) \geq \min(p, 1-p)^n,\ \forall p \in \mathcal{P}$.

(Thus, for example, when $\mathcal{P} = (0,1)$ the function $f(p) = p/2$ works, but when $\mathcal{P} = (0, 0.5)$ the function $f(p) = 2p$ does not work.) The authors also give a procedure to simulate a Bernoulli $f(p)$-variable if the above conditions are satisfied. However, the number N of Bernoulli-p samples required by the procedure is random and unbounded even though the expected number of samples is finite. In [Nacu and Peres 2005] the class of functions, f, that admit *fast simulation*, i.e., the tail probabilities of the number of bernoulli samples N is exponentially bounded is characterized. In particular, it is shown that a function $f : \mathcal{P} \rightarrow (0,1)$ has a fast simulation if f is real analytic on the closed interval $\mathcal{P} \subset (0,1)$. Conversely, if f has a fast simulation, then it is real analytic on any open subset of $\mathcal{P}$. Their method of proof involves approximating f by a sequence of Bernstein polynomials.[1]

In the present paper as a modest contribution to the field, we are interested in characterizing the family of functions f that can be simulated using a *finite* number of bernoulli samples with unknown bias p. We restrict our attention to the class of simu-

[1] In general, given a function $f : [0,1] \rightarrow [0,1]$, the n^{th} Bernstein approximation is given by $Q_n(x) = \sum_{k=0}^{n} f(n/k)\binom{n}{k}x^k(1-x)^{n-k}$.

Authors' addresses: V. Goyal and K. Sigman, Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027 ({vg2277, ks20}@columbia.edu).

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

lation algorithms where the number of required samples, N, is a constant that depends only on the function f. It is useful to consider such a class of algorithms for computational efficiency of the simulation procedure. The need to estimate the probability of an event where the probability of basic events is unknown arises in many important applications. For instance, in the network reliability problem, often we need to estimate the probability that the network will be connected when each link can fail independently with an unknown probability. An insurance company needs to estimate that the probability that there are more than k claims in a given year to manage risk when each claim happens independently with an unknown probability. Therefore, given a black-box that generates independent bernoulli samples of basic events (such as link failure or claim), a fast simulation procedure is important to efficiently estimate the probability of the required event. This has been addressed in the literature. In [Peres 1992], for example, the problem of generating unbiased bernoulli samples from independent bernoulli samples with unknown bias is considered and and a recursive procedure is derived that is more efficient than the von Neuman's procedure [Neumann 1951]. In particular, the number of unbiased samples generated by the procedure from n biased samples is close to the entropy bound. Even in [Nacu and Peres 2005], the basic motivation is the efficiency of the simulation algorithm as the authors are interested in characterizing the class of functions that can be simulated by a random number of samples whose distribution has an exponentially decaying tail. Finally, as a fundamental example, the fact that a Bernoulli $2p$ rv can't be simulated by using Bernoulli (p) rvs with unknown $p \in (0, 0.5)$ (proved in [Keane and O'Brien 1994]) led to a very important result showing that in general, stationary versions of regenerative processes can't be simulated (see Section 8, Page 420 in [Asmussen and Glynn 2008]; in particular Example 8.6 on Page 422.)

Interestingly, we can completely characterize the family of functions that can be simulated using a finite number of bernoulli samples. Furthermore, our characterization also gives an explicit algorithm for simulating $f(p)$ if f can be simulated using a finite number of samples. The characterization is based on ideas similar to approximating f using Bernstein polynomials in [Nacu and Peres 2005]. To describe the characterization, let us define the family of *Bernstein polynomials*.

Definition 1.1. A *Bernstein polynomial* of degree n is a linear combination of the following $n+1$ *Bernstein basis polynomials*,

$$h_{j,n}(x) = \binom{n}{j} x^j (1-x)^{n-j}, j = 0, \ldots, n.$$

Therefore, a bernstein polynomial, f of degree n is defined as

$$f(x) = \sum_{j=0}^{n} \alpha_j h_{j,n},$$

where $\alpha_j \in \mathbb{R}, j = 0, \ldots, n$ are referred to as the *Bernstein coefficients*.

It is easy to observe that every polynomial is a Bernstein polynomial with the appropriate Bernstein coefficients. Moreover, the Bernstein coefficients for any polynomial can be computed efficiently by solving a simple (lower triangular) system of linear equations. By expressing the polynomial as a Bernstein polynomial, we can often use some nice properties of the Bernstein basis polynomial.

We show that a Bernoulli $f(p)$-random variable can be simulated using a bounded number of samples of a Bernoulli p-random variable ($p \in (0, 1)$ unknown) if and only if f is a *Bernstein polynomial* with *Bernstein coefficients* in the interval $[0, 1]$. Since for any polynomial f, the Bernstein coefficients can be computed efficiently, it is easy to check whether a particular function f can be simulated using a finite number of Bernoulli p-samples. Furthermore, as we will see, the proof of the above characterization gives an explicit algorithm for the simulation.

2. MAIN RESULT

In this section, we prove our main result. We first show that if f is a polynomial with Bernstein coefficients between 0 and 1, then a Bernoulli $f(p)$-variable can be simulated using a finite number of samples of Bernoulli p-variables. We prove this by giving an algorithm that uses a finite number of Bernoulli p-samples and an independent auxiliary variable with known distribution. We then prove the converse, i.e., if a function f is simulated using a finite number of independent Bernoulli p-samples and an independent auxiliary variable with known distribution, then f must be a Bernstein polynomial with Bernstein coefficients between 0 and 1.

The Simulation Algorithm $\mathcal{A}$. Suppose f is a Bernstein polynomial of degree n with Bernstein coefficients between 0 and 1:

$$f(p) = \sum_{j=0}^{n} \alpha_j \binom{n}{j} p^j (1-p)^{n-j},$$

where $\alpha_j \in [0,1]$ for all $j = 0, \ldots, n$.

(1) Let $X_1, \ldots, X_n$ be n independent samples from the Bernoulli distribution with bias p and let U be an independent sample from the uniform distribution over $(0,1)$.
(2) Compute

$$Y = \sum_{j=0}^{n} \mathbb{I}(U \le \alpha_j) \cdot \mathbb{I}\left(\sum_{k=1}^{n} X_k = j\right),$$

where $\mathbb{I}(\cdot)$ denotes the indicator function that outputs 1 if the event is true and 0 otherwise.
(3) Return Y.

Note that the algorithm uses only a finite number (equal to the degree of f) of samples of Bernoulli p-variable and one independent sample from a uniform distribution.

PROPOSITION 2.1. *Algorithm $\mathcal{A}$ simulates a Bernoulli $f(p)$-variable.*

PROOF. Let us compute the probability that Algorithm $\mathcal{A}$ returns 1, i.e., $P(Y=1)$. Let

$$E_j = \{U \le \alpha_j\} \cap \left\{\sum_{k=1}^{n} X_k = j\right\}.$$

Therefore, $Y = \mathbb{I}(E_0) + \ldots + \mathbb{I}(E_n)$. Note that events $E_j, j = 0, \ldots, n$ are mutually exclusive. Therefore,

$$\begin{aligned}
P(Y=1) &= P\left(\sum_{j=0}^{n} \mathbb{I}(E_j) = 1\right) \\
&= \sum_{j=0}^{n} P(\mathbb{I}(E_j) = 1) && (1) \\
&= \sum_{j=0}^{n} P(U \le \alpha_j) \cdot P\left(\sum_{k=0}^{n} X_k = 1\right) && (2) \\
&= \sum_{j=0}^{n} \alpha_j \binom{n}{k} p^k (1-p)^{n-k} \\
&= f(p),
\end{aligned}$$

where (1) follows as $E_j, j = 0, \ldots, n$ are mutually exclusive, Equation (2) follows as U is an independent sample from $\mathsf{Uniform}(0,1)$ and the second to last equation follows as $X_1 + \ldots + X_n$ has a binomial (n, p) distribution. □

Next we prove the converse.

PROPOSITION 2.2. *Suppose a function f can be simulated using a finite number of independent Bernoulli p-samples and an independent auxiliary variable from a known distribution, then f must be a Bernstein polynomial with Bernstein coefficients between 0 and 1.*

PROOF. Consider an algorithm $\hat{\mathcal{A}}$ that uses M independent samples of Bernoulli p-variable (p unknown) for some $M \in \mathbb{Z}_+$ and one independent sample of an auxiliary variable with a known distribution and generates a Bernoulli $f(p)$-variable. We can assume wlog that the auxiliary variable is $\mathsf{Uniform}(0,1)$. Let $X_1, \ldots, X_M$ denote the independent Bernoulli p-samples and U be an independent sample from $\mathsf{Uniform}(0,1)$. Let $\hat{\mathcal{A}}(X_1, \ldots, X_M, U)$ denote the output of Algorithm $\hat{\mathcal{A}}$.

Note that there are 2^M possible values of the M-tuple $(X_1, \ldots, X_M)$. For any set of samples $(X_1, \ldots, X_M)$ and value of U, $\hat{\mathcal{A}}$ either returns 0 or 1. For any $\boldsymbol{x} = (x_1, \ldots, x_M) \in \{0,1\}^M$, let $n(\boldsymbol{x}) = x_1 + \ldots + x_M$, the number of 1's in $\boldsymbol{x}$. Also, for any $\boldsymbol{x} \in \{0,1\}^M$, let

$$T(\boldsymbol{x}) = \{u \in [0,1] \mid \hat{\mathcal{A}}(\boldsymbol{x}, u) = 1\}.$$

Therefore,

$$\begin{aligned}
P(\hat{\mathcal{A}}(X_1,\ldots,X_M,U)=1) &= \sum_{\boldsymbol{x}\in\{0,1\}^M} P((X_1,\ldots,X_M)=\boldsymbol{x})\cdot P(U\in T(\boldsymbol{x})) \\
&= \sum_{\boldsymbol{x}\in\{0,1\}^M} \left(p^{n(\boldsymbol{x})}(1-p)^{M-n(\boldsymbol{x})}\cdot P(U\in T(\boldsymbol{x}))\right) \\
&= \sum_{j=0}^{M}\sum_{\boldsymbol{x}:n(\boldsymbol{x})=j} \left(p^{j}(1-p)^{M-j}\cdot P(U\in T(\boldsymbol{x}))\right) \\
&= \sum_{j=0}^{M} p^{j}(1-p)^{M-j}\left(\sum_{\boldsymbol{x}:n(\boldsymbol{x})=j} P(U\in T(\boldsymbol{x}))\right). \qquad (3)
\end{aligned}$$

Note that for any $j=0,\ldots,M$,

$$\sum_{\boldsymbol{x}:n(\boldsymbol{x})=j} P(U\in T(\boldsymbol{x}) \le \binom{M}{j} = \alpha_j\binom{M}{j},\ \text{for some } \alpha_j\in[0,1].$$

Therefore, from (3), we have

$$P(\hat{\mathcal{A}}(X_1,\ldots,X_M,U)=1) = \sum_{j=0}^{M}\alpha_j\binom{M}{j}p^j(1-p)^{M-j},$$

for some $\alpha_j\in[0,1]$ for all $j=0,\ldots,M$. Since Algorithm $\hat{\mathcal{A}}$ simulates $f(p)$, then

$$f(p) = \sum_{j=0}^{M}\alpha_j\binom{M}{j}p^j(1-p)^{M-j},$$

which is a Bernstein polynomial with Bernstein coefficients $\alpha_j\in[0,1], j=0,\ldots,M$. □

From Propositions 2.1 and 2.2, we have the following theorem.

THEOREM 2.3. *A function $f:\mathcal{P}\to(0,1)$ can be simulated using a finite number of independent Bernoulli p-samples for all $p\in\mathcal{P}$ and an independent auxiliary variable from a known distribution if and only if f is a Bernstein polynomial with Bernstein coefficients between 0 and 1.*

REFERENCES

ASMUSSEN, S. AND GLYNN, P. 2008. *Stochastic Simulation: Algorithms and Analysis*. Springer-Verlag, New York, NY, USA.

KEANE, M. AND O'BRIEN, G. 1994. A Bernoulli factory. *ACM Transactions on Modeling and Computer Simulation (TOMACS) 4,* 2, 213–219.

NACU, Ş. AND PERES, Y. 2005. Fast simulation of new coins from old. *The Annals of Applied Probability 15,* 1A, 93–115.

NEUMANN, J. 1951. Various techniques used in connection with random digits. *National Bureau of Standards, Applied Math Series 12*, 36–38.

PERES, Y. 1992. Iterating von Neumann's procedure for extracting random bits. *The Annals of Statistics 20,* 1, 590–597.