MOD$_p$-tests, Almost Independence and Small Probability Spaces

(Extended Abstract)

Claudia Bertram-Kretzberg and Hanno Lefmann

Lehrstuhl Informatik II, Universität Dortmund, D-44221 Dortmund, Germany
bertram, lefmann@ls2.informatik.uni-dortmund.de

Abstract. We consider approximations of probability distributions over $\mathbb{Z}_p^n$. We present an approach to estimate the quality of approximations towards the construction of small probability spaces which are used to derandomize algorithms. In contrast to results by Even et al. [13], our methods are simple, and for reasonably small $p$, we get smaller sample spaces. Our considerations are motivated by a problem which was mentioned in recent work of Azar et al. [5], namely, how to construct in time polynomial in $n$ a good approximation to the joint probability distribution of i.i.d. random variables $X_1, \ldots, X_n$ where each $X_i$ has values in $\{0, 1\}$. Our considerations improve on results in [5].

1 Introduction

During the last years, techniques have been developed to minimize the number of random bits which are used by randomized algorithms. These methods are based on the replacement of independent random variables by some weakly dependent random variables which can be generated using fewer bits, therefore, dropping the running times of several algorithms. Alon, Babai and Itai [1] observed that it suffices for certain algorithms to use only pairwise independent bits instead of mutually independent ones. In general, to generate $k$-wise independent bits sample spaces of size only $O(n^k)$ can be used, cf. Karloff and Mansour [15]. However, for some algorithms a large amount of independence is desirable. In view of this, Berger and Rompel [8] showed that sometimes it suffices to consider only $(\log n)^{\epsilon}$-wise independence of random variables. Small probability spaces are very desirable for derandomizing randomized algorithms. The resulting sample space, which should reflect the behaviour of the considered random variables, can be investigated by exhaustive search or the method of conditional probabilities, cf. Alon and Spencer [4], and Motwani, Naor and Naor [18].

Instead of looking for small probability spaces, Naor and Naor [19] considered approximations to probability distributions. They used the notion of the bias of a distribution, cf. Vazirani [22]. Let $X_1, \ldots, X_n$ be random variables with values in $\{0, 1\}$. The bias of a subset $S \subseteq \{X_1, \ldots, X_n\}$ w.r.t. linear tests is defined by $|Pr[\sum_{i \in S} X_i \equiv 0 \mod 2] - Pr[\sum_{i \in S} X_i \equiv 1 \mod 2]|$. In an $\epsilon$-biased distribution, each subset $S$ of the random variables has bias at most $\epsilon$. Clearly, for uniform random variables, the bias is zero. Nao and Naor gave in [19] constructions of $\epsilon$-biased distributions where the sample space has size $\text{poly}(n, 1/\epsilon)$. A different construction based on Weil's theorem on quadratic residues was given by Peralta.
Alon, Goldreich, Håstad and Peralta [3] gave three constructions, including Peralta’s, for $\epsilon$-biased sample spaces $S \subseteq \mathbb{Z}_p^n$ w.r.t. linear tests in $\mathbb{Z}_2$ of size $O(n^2/(\epsilon^2(\log(n)/\epsilon))^{d}))$ where $\delta = 1, 0$ and 2, respectively. Azar, Motwani and Naor [5] generalized the work of [3] to random variables with values from arbitrary groups, in particular for $\mathbb{Z}_p = \{0, 1, \ldots, p - 1\}$, the set of residues modulo $p$, cf. [14] for applications. There, they used Weil’s theorem on character sums and Fourier transforms to measure the quality of approximations of the uniform distribution over $\mathbb{Z}_p^n$. Here, we use a more elementary way to achieve this, and we obtain sharper bounds.

Besides Weil’s theorem on quadratic residues, a similar behaviour of the underlying structures is given by Lindsey’s inequality [6] or by the corresponding inequalities for Expander- respectively Ramanujan graphs [17]. These phenomena can be summarized under the term Quasirandomness, see [11], namely, the structures behave approximately like random, that is, show small discrepancies. From that point of view, it is natural that the combinatorial notion of discrepancy was taken into account with the work of Even, Goldreich, Luby, Nisan and Veličković [13]. Indeed, Alon, Bruck, Naor, Naor and Roth [2] used Ramanujan graphs to construct good error-correcting codes which also yield small sample spaces for approximating the joint distribution of random variables.

Azar, Motwani and Naor stated in [5] the problem of finding good approximations for the joint distribution of i.i.d. random variables $X_1, \ldots, X_n$ with values in $\{0, 1\}$, where $Pr[X_1 = 0] = 1 - Pr[X_1 = 1] = \sigma \neq \frac{1}{2}$. Even, Goldreich, Luby, Nisan and Veličković [13] considered this problem in a general setting, namely, for independent random variables $X_1, \ldots, X_n$ with values in $\{1, \ldots, m\}$ where $Pr[X_i = j] = p_{i,j}, 1 \leq i \leq n$ and $1 \leq j \leq m$. In [13], constructions of small sample spaces were given which approximate the joint distribution of $X_1, \ldots, X_n$. To do so, they used the combinatorial notion of discrepancy, cf. [7]. Let $R_n$ be the set of all axis aligned rectangles of the $n$-dimensional cube $[0, 1)^n$. For any finite set $S \subseteq [0, 1)^n$ and any rectangle $R \in R_n$ with volume $vol(R)$, the discrepancy of $S$ on $R_n$ is defined by $disc_S(R_n) = \sup_{R \in R_n} |vol(R) - |S \cap R||/|S|.$

A sample space $S \subseteq \{1, \ldots, m\}^n$ is $(\epsilon, k)$-independent w.r.t. the joint distribution of the independent random variables $X_1, \ldots, X_n$ with values in $\{1, \ldots, m\}$ if for any sequence $(\alpha_{i_1}, \ldots, \alpha_{i_k}) \in \{1, \ldots, m\}^k$ it holds $|Pr[(X_{i_1}, \ldots, X_{i_k}) = (\alpha_{i_1}, \ldots, \alpha_{i_k})] - \prod_{j=1}^{k} p_{i_j, \alpha_{i_j}}| \leq \epsilon$. In [13], Even et al. showed that sets $S$ with small discrepancy, i.e. $disc_S(R_n) \leq \epsilon$, yield sample spaces which are $(\epsilon, k)$-independent w.r.t. the joint distribution of random variables. Their construction has the advantage to be universal. One construction in [13] yields an $(\epsilon, k)$-independent sample space $S \subseteq \{1, \ldots, m\}^n$ of size $poly(\log n, 2^k, 1/\epsilon)$, while the other two constructions yield $(\epsilon, k)$-independent spaces $S \subseteq \{1, \ldots, m\}^n$ of size $poly((n/\epsilon)^{\log(1/\epsilon)})$ and $poly((n/\epsilon)^{\log n})$, respectively. The results of [13] were extended and applied by Chari, Rohatgi and Srinivasan [10]. Again using the notion of discrepancy and projections, they constructed an $(\epsilon, k)$-independent sample space $S$ of size $poly(\log n, 1/\epsilon, \min\{2^k, k^{\log(1/\epsilon)}\})$.

Our considerations here are motivated by the problem from Azar, Motwani and Naor [5]. In contrast to the work of [13] and [10] where the discrepancy of axis