The 123 Theorem and its extensions

Noga Alon *
and Raphael Yuster
Department of Mathematics
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University, Tel Aviv, Israel

Abstract

It is shown that for every $b > a > 0$ and for every two independent identically distributed real random variables X and Y

$$Prob[|X - Y| \leq b] < (2\lceil b/a \rceil - 1)Prob[|X - Y| \leq a].$$

This is tight for all admissible pairs a, b. Higher dimensional extensions are also considered.

1 Introduction

Our first result in this note is the following theorem, which we name after the three constants in its statement.

Theorem 1.1 (The 123 Theorem) *Let X and Y be two independent, identically distributed real random variables. Then*

$$Prob[|X - Y| \leq 2] < 3Prob[|X - Y| \leq 1].$$

The problem of determining the smallest possible constant C so that for every two independent, identically distributed (=i.i.d.) real random variables the inequality

$$Prob[|X - Y| \leq 2] \leq CProb[|X - Y| \leq 1]$$

holds was suggested by G. A. Margulis and communicated to us by Y. Peres. Several researchers, including Peres, observed that the smallest possible C satisfies $3 \leq C \leq 5$, where the lower bound

*Research supported in part by a United States Israel BSF Grant

follows by considering two i.i.d. real random variables X and Y distributed uniformly on the discrete set $\{2, 4, 6, \ldots, 2n\}$. Here $Prob[|X - Y| \leq 1] = 1/n$ whereas $Prob[|X - Y| \leq 2] = 3/n - 2/n^2$, showing that $C \geq 3 - 2/n$ for every n. Theorem 1.1 thus shows that in fact $C = 3$. As we learned from Peres, a slightly weaker version of this theorem (without the strict inequality) has also been proved, independently of our work (and before us) by A. Kotlov in a different way. Our proof, presented in the next section, is shorter and has the additional advantage that it actually gives the following more general result.

Theorem 1.2 *Let $b > a > 0$ be two reals. Then for every two i.i.d. real random variables X and Y*

$$Prob[|X - Y| \leq b] < (2\lceil b/a \rceil - 1)Prob[|X - Y| \leq a].$$

Moreover, the constant $2\lceil b/a \rceil - 1$ cannot be improved.

The above questions can be considered for vector-valued random variables as well. Let $V = R^d$ be a finite dimensional Euclidean space, and let $b \geq a > 0$ be two reals. Let $C(V, a, b)$ denote the smallest possible C such that for every two i.i.d. random variables X and Y taking values in V

$$Prob[||X - Y|| \leq b] \leq CProb[||X - Y|| \leq a].$$

Note that $C(V, a, b) = C(V, a', b')$ whenever $b/a = b'/a'$ and hence it suffices to consider the case $a = 1$. Our method supplies rather tight estimates for the function $C(R^d, 1, b)$. However, the problem of determining this function precisely seems difficult, even for the Euclidean plane R^2. Yet, the technique does enable us to determine this function precisely for infinitely many values of b.

Theorem 1.3 *Let $V = R^d$, $n \geq 2$ and suppose that there is no set F of $n + 1$ points in a ball of radius b so that the center belongs to F and the distance between any two points of F exceeds 1. If there is a lattice in V with minimum distance 1 so that there are n points of it in a ball of radius smaller than b centered at a lattice point, then $C(V, 1, b) = n$.*

This theorem, together with the main result of [1], implies that there is an absolute constant $\epsilon > 0$ such that in the Euclidean plane R^2, $C(R^2, 1, b) = 7$ for all $1 < b < 1 + \epsilon$ and $C(R^2, 1, b) = 19$ for all $2 < b < 2 + \epsilon$. Similarly, the Leech Lattice and the known bounds on kissing numbers (see [2]) imply that there exists an $\epsilon > 0$ such that in dimension 24, $C(R^{24}, 1, b) = 196,561$ for all $1 < b < 1 + \epsilon$. Also, for $1 < b < 1 + \epsilon$, $C(R^3, 1, b) = 13$ and $C(R^8, 1, b) = 241$, where $\epsilon > 0$ is an appropriate absolute constant.

The rest of this note is organized as follows. In Section 2 we consider the real case and present the proofs of Theorems 1.1 and 1.2. In Section 3 we consider higher dimensions.

2 Real random variables

Our basic approach is combinatorial, and (vaguely) resembles the method of Katona in [3]. Let $T = (x_1, x_2, \ldots, x_m)$ be a sequence of not necessarily distinct reals. For any positive b, define

$$T_b = \{(x_i, x_j) : 1 \leq i, j \leq m,\ |x_i - x_j| \leq b\}.$$

We need the following simple combinatorial lemma.

Lemma 2.1 *For any sequence T as above and for every integer $r > 1$,*

$$|T_r| < (2r-1)|T_1|.$$

Proof. We apply induction on $|T| = m$. The result is trivial for $m = 1$. Assuming it holds for $m-1$, we prove it for m (> 1). Given a sequence $T = (x_1, \ldots, x_m)$ let $t+1$ be the maximum number of points of T in a closed interval of length 2 centered at a member of T. Let x_i be any rightmost point of T so that there are $t+1$ members of T in the interval $[x_i - 1, x_i + 1]$ and define $T' = T \setminus \{x_i\}$. The number of members of T' in the interval $[x_i - 1, x_i + 1]$ is clearly t and hence x_i appears in precisely $2t+1$ ordered pairs of T_1. Thus

$$|T_1| = 2t + 1 + |T'_1|.$$

The interval $[x_i - r, x_i + r]$ is the union of the $2r-1$ smaller intervals

$$[x_i - r, x_i - r + 1), \ldots, [x_i - 2, x_i - 1), [x_i - 1, x_i + 1], (x_i + 1, x_i + 2], \ldots, (x_i + r - 1, x_i + r] \quad (1)$$

By the choice of x_i, each of these smaller intervals can contain at most $t+1$ members of T, and each of the last $r-1$ ones, which lie to the right of x_i, can contain at most t members of T. Altogether there are thus at most $(r-1)(t+1) + rt$ members of T' in $[x_i - r, x_i + r]$ and hence

$$|T_r| \leq 2(r-1)(t+1) + 2rt + 1 + |T'_r| = (2r-1)(2t+1) + |T'_r|.$$

By the induction hypothesis $|T'_r| < (2r-1)|T'_1|$ and hence $|T_r| < (2r-1)|T_1|$, completing the proof. □

Corollary 2.2 *Let X and Y be two i.i.d real random variables. For a positive b, define $p_b = Prob[|X - Y| \leq b]$. Then for every integer r, $p_r \leq (2r-1)p_1$.*

Proof. Fix an integer m, and let $S = (x_1, \ldots, x_m)$ be a random sequence of m elements, where each x_i is chosen, randomly and independently, according to the distribution of X. By Lemma 2.1

$$|S_r| < (2r-1)|S_1|.$$

Therefore, the expectation of $|S_r|$ is smaller than that of $(2r-1)|S_1|$. However, by the linearity of expectation it follows that the expectation of $|S_b|$ is precisely $m + m(m-1)p_b$ for every positive b. Therefore,

$$m + m(m-1)p_r < (2r-1)(m + m(m-1)p_1),$$

implying that for every integer m,

$$p_r < (2r-1)p_1 + \frac{2r-2}{m-1}.$$

The desired result $p_r \leq (2r-1)p_1$ follows, by letting m tend to infinity. □

The last corollary suffices to prove the assertions of Theorems 1.1 and 1.2, without the strict inequality. To prove the strict inequality we need an additional argument, which follows. Let X and Y be two i.i.d. real random variables, suppose $r > 1$ is an integer, and suppose that $p_r = (2r-1)p_1$, where p_r and p_1 are defined as before. For two reals a and b, define $\mu_b(a) = Prob[|X - a| \leq b]$.

Claim: there exists a real a so that $\mu_r(a) > (2r-1)\mu_1(a)$.

Proof. Otherwise, $\mu_r(a) \leq (2r-1)\mu_1(a)$ for each a, and since $p_r = (2r-1)p_1$ and p_b is simply the expectation of $\mu_b(a)$ when a is chosen according to the distribution of X, it follows that $\mu_r(a) = (2r-1)\mu_1(a)$ with probability 1 (when a is chosen according to the distribution of X). Let A be the set of all real values a for which $\mu_r(a) = (2r-1)\mu_1(a)$, and define

$$\delta = Sup\{\mu_1(a) : a \in A\}.$$

Clearly $\delta > 0$. Let ϵ be a small positive constant such that

$$k(\delta - (2r)^k\epsilon) > 1, \quad where \quad k = \lceil 2/\delta \rceil. \tag{2}$$

Pick $a_0 \in A$ so that $\mu_1(a_0) > \delta - \epsilon$. We next define a sequence of k pairwise disjoint unit intervals $I_1, \ldots, I_k$ in the line, so that for each i, $Prob[X \in I_i] > \delta - (2r)^i\epsilon$. Since in view of (2) this is impossible, the assertion of the claim will follow. The first interval I_1 is simply the interval $(a_0 + r - 1, a_0 + r]$. Observe that by the choice of a_0,

$$Prob[a_0 - r \leq X \leq a_0 + r] \geq (2r-1)(\delta - \epsilon).$$

Split the interval $[a_0 - r, a_0 + r]$ into $2r-1$ smaller intervals as in (1). Note that the definition of δ implies that the probability that X lies in any one of these intervals is at most δ. Therefore, for each of these smaller intervals, (and in particular for the last one- $I_1 = (a_0 - r + 1, a_0 + r]$) the probability that X lies in the interval is at least $(2r-1)(\delta - \epsilon) - (2r-2)\delta > \delta - 2r\epsilon$, as needed. Suppose, now, that the pairwise disjoint unit intervals $I_1, \ldots, I_j$ have already been defined, where I_j is the rightmost interval, and $Prob[X \in I_i] > \delta - (2r)^i\epsilon$ for all $1 \leq i \leq j$ $(< k)$. We can now

define I_{j+1} as follows. Since X attains values in A with probability 1, and it lies in I_j with positive probability, there is an $a_j \in I_j \cap A$. Obviously, $\mu_1(a_j) \geq Prob[X \in I_j] > \delta - (2r)^j\epsilon$. Therefore, as $a_j \in A$,

$$Prob[a_j - r \leq X \leq a_j + r] \geq (2r-1)Prob[X \in I_j] > (2r-1)(\delta - (2r)^j\epsilon).$$

We can thus define $I_{j+1} = (a_j + r - 1, a_j + r]$ and conclude, as before, that

$$Prob[X \in I_{j+1}] > \delta - (2r-1)(2r)^j\epsilon > \delta - (2r)^{j+1}\epsilon,$$

as required. This supplies the desired contradiction and completes the proof of the claim. □

Returning to our two i.i.d. real random variables X and Y for which $p_r = (2r-1)p_1$, observe that by the claim there is real a such that

$$\mu_r(a) = (2r-1)\mu_1(a) + \beta, \tag{3}$$

where $\beta > 0$. Let $\alpha > 0$ be a small constant satisfying $\alpha/(1-\alpha) < \beta/(r-1)$. Define X' as the random variable which has the distribution of X with probability $(1-\alpha)$ and with probability α it gets the value a. For any real b, let $p'_b = Prob[|X' - Y'| \leq b]$, where X', Y' are i.i.d. random variables with the distribution of the above X'. By the definition of X', for every positive b,

$$p'_b = (1-\alpha)^2 p_b + 2\alpha(1-\alpha)\mu_b(a) + \alpha^2.$$

By Corollary 2.2 applied to X', $p'_r \leq (2r-1)p'_1$. In view of the last equality and (3) this implies that

$$(1-\alpha)^2 p_r + 2\alpha(1-\alpha)((2r-1)\mu_1(a) + \beta) + \alpha^2 \leq (2r-1)[(1-\alpha)^2 p_1 + 2\alpha(1-\alpha)\mu_1(a) + \alpha^2].$$

Therefore,

$$p_r \leq (2r-1)p_1 - 2\frac{\alpha}{1-\alpha}\beta + (2r-2)\frac{\alpha^2}{(1-\alpha)^2} < (2r-1)p_1,$$

where the last inequality follows from the choice of α. This shows that equality is impossible in Corollary 2.2. We have thus proved the following

Proposition 2.3 *In the notation of Corollary 2.2, for every integer $r > 1$, $p_r < (2r-1)p_1$.* □

We can now complete the proof of Theorem 1.2 (which implies, of course, Theorem 1.1).

Proof of Theorem 1.2. Applying the last proposition to $X' = X/a$, $Y' = Y/a$ and $r = \lceil b/a \rceil$ we conclude that

$$Prob[|X - Y| \leq b] = Prob[|X' - Y'| \leq b/a] \leq Prob[|X' - Y'| \leq r]$$

$$< (2r-1)Prob[|X'-Y'| \leq 1] = (2\lceil b/a \rceil - 1)Prob[|X-Y| \leq a],$$

as required. To see that the constant $2\lceil b/a \rceil - 1$ cannot be improved, let γ be a real satisfying $a < \gamma$ and $\gamma(\lceil b/a \rceil - 1) < b$. Let m be a large integer and let the two i.i.d. random variables X and Y be distributed uniformly on the discrete set $\gamma, 2\gamma, \ldots, m\gamma$. One can easily check that the ratio between $Prob[|X-Y| \leq b]$ and $Prob[|X-Y| \leq a$ approaches $2\lceil b/a \rceil - 1$ as m tends to infinity. □

3 Vector valued random variables

The basic method in the previous section can be modified and extended to higher dimensions. Let $V = R^d$ be the d-dimensional Euclidean space. We start with the following simple observation.

Lemma 3.1 *Suppose there exists a lattice in V with minimum distance 1 so that there are n points of it in a ball of radius smaller than b centered at a lattice point, then $C(V, 1, b) \geq n$.*

Proof. Let L be the above lattice, and let $\gamma > 1$ be close enough to 1 so that γL contains n points in a ball of radius b centered at a lattice point. Let R be a large real, and let X_R and Y_R be two i.i.d. random variables, each uniformly distributed on the points of γL whose norm is at most R. It is easy to check that when R tends to infinity the ratio between $Prob[||X_R - Y_R|| \leq b]$ and $Prob[||X_R - Y_R|| \leq 1]$ approaches the number of points of γL in a ball of radius b centered at a point of γL, which is at least n. □

Lemma 3.2 *Suppose $n \geq 2$, and suppose there is no set F of $n+1$ points in a ball of radius b in V, such that the center is in F and the distance between any two members of F exceeds 1. Let $T = (x_1, \ldots, x_m)$ be a sequence of points in V. For any positive c and s, define*

$$T_{c,s} = sm + |\{(x_i, x_j) : 1 \leq i \neq j \leq m, ||x_i - x_j|| \leq c\}|.$$

If $ns > 2n + s$, then

$$T_{b,s} < nT_{1,s}.$$

Proof. We apply induction on m. The result is trivial for $m = 1$, since in this case $T_{b,s} = T_{1,s} = s$. Assuming it holds for $m-1$ we prove it for m ($m > 1$). Given a sequence T of cardinality m as above, let $t+1$ be the maximum number of members of T in a ball of radius 1 centered at a point of T. Let x be a point of T with $t+1$ members of T in the radius-1 ball centered at x, and define $T' = T \setminus \{x\}$. Clearly

$$T_{1,s} = T'_{1,s} + s + 2t.$$

Let F be a subset of maximum cardinality of T in the ball of radius b centered at x, so that $x \in F$ and the distance between any two members of F is strictly bigger than 1. By the assumption, $|F| \leq n$. Moreover, any point of T in the ball of radius b centered at x lies in a radius-1 ball centered at a point of F. Since, by the maximality in the choice of x, no such ball can contain more than $t+1$ points, it follows that there are at most $n(t+1)$ members of T in the ball of radius b centered at x (including x itself). Therefore,

$$T_{b,s} \leq T'_{b,s} + s + 2(nt + n - 1).$$

By the induction hypothesis $T'_{b,s} < nT'_{1,s}$ provided $ns > 2n + s$ and hence, for such s,

$$T_{b,s} < nT'_{1,s} + s + 2(nt + n - 1) < nT'_{1,s} + ns + 2nt = nT_{1,s},$$

completing the proof. □

Corollary 3.3 *Suppose $n \geq 2$, and suppose there is no set F of $n+1$ points in a ball of radius b in V, such that the center is in F and the distance between any two members of F exceeds 1. Then $C(V, 1, b) \leq n$. That is, for any two i.i.d. V-valued random variables X and Y*

$$Prob[||X - Y|| \leq b] \leq n \; Prob[||X - Y|| \leq 1]. \tag{4}$$

Proof. For X, Y as above and for a positive c, define $p_c = Prob[||X - Y|| \leq c]$. Fix a positive s satisfying $ns > 2n + s$. For a fixed, large integer m, let $T = (x_1, \ldots, x_m)$ be a random sequence of m elements, where each x_i is chosen, randomly and independently, according to the distribution of X. By Lemma 3.2

$$T_{b,s} < nT_{1,s}.$$

Therefore, the expectation of $T_{b,s}$ is smaller than that of $nT_{1,s}$. By linearity of expectation the expectation of $T_{c,s}$ is precisely $sm + m(m-1)p_c$ for every positive c. Therefore,

$$sm + m(m-1)p_b < n(sm + m(m-1)p_1),$$

implying that for every integer m,

$$p_b < np_1 + \frac{(n-1)s}{m-1}.$$

The desired result (4) follows, by letting m tend to infinity. □

Theorem 1.3 follows from Lemma 3.1 and Corollary 3.3. We next describe some consequences of this theorem.

In [1] it is shown that the minimum radius of a two dimensional ball containing 8 points one of which is at the center, so that all mutual distances are at least 1, is $\frac{1}{2}cosec(\pi/7) = 1.15....$ It is also shown that the minimum radius of a two dimensional ball containing 20 points one of which is its center such that all mutual distances are at least 1 is strictly bigger than 2. This, together with Theorem 1.3 (and the existence of the hexagonal lattice) implies the following.

Proposition 3.4 *(i) If* $1 < b < \frac{1}{2}cosec(\pi/7)$ *then* $C(R^2, 1, b) = 7$.
(ii) There exists an $\epsilon > 0$ *so that for all* $2 < b < 2 + \epsilon$, $C(R^2, 1, b) = 19$. □

Similarly, one can determine the asymptotic behaviour of $C(R^2, 1, b)$ as b tends to infinity. It is well known (see [1]), that the maximum number of points that can be placed in a radius-b two dimensional ball so that one of the points is at the center, and all mutual distances are at least 1, is $(1 + o(1))\frac{2\pi}{\sqrt{3}}b^2$. This is realized by the hexagonal lattice and hence, by Theorem 1.3 the following statement holds.

Proposition 3.5 *As* b *tends to infinity,*

$$C(R^2, 1, b) = (1 + o(1))\frac{2\pi}{\sqrt{3}}b^2. \quad \square$$

The *kissing number* τ_d is the maximum number of points that can be placed on the boundary of a unit ball in R^d, so that the distance between any two of the points is at least 1. By compactness it follows that the minimum radius of a ball in R^d containing $\tau_d + 2$ points one of which is at the center, so that all mutual distances are at least 1, is strictly bigger than 1. The exact values of τ_d are known only for $d = 1, 2, 3, 8$ and 24. Trivially $\tau_1 = 2$ and $\tau_2 = 6$. The value of τ_3 was the subject of a discussion between Isaac Newton and David Gregory in 1694. Newton believed that $\tau_3 = 12$, and as shown by various researchers in the nineteenth century, this is indeed the case. A very short proof of this fact appears in [4]. In [6], [5] it is shown that $\tau_8 = 240$ and $\tau_{24} = 196,560$. In all the above cases the highest possible kissing numbers are attainable by lattices. (It is known that this is *not* the case in dimension 9.) The relevant lattices are the trivial one in dimension 1, the hexagonal lattice in dimension 2, the face-centered cubic lattice in dimension 3, the lattice E_8, (sometimes called the 8-dimensional diamond lattice), in dimension 8, and the well known Leech lattice in dimension 24. See [2] for more details. Theorem 1.3 thus gives the following.

Proposition 3.6 *For every* $d \in \{1, 2, 3, 8, 24\}$ *there exists an* $\epsilon_d > 0$ *such that for all* $1 < b < 1+\epsilon_d$, $C(R^d, 1, b) = \tau_d + 1$. □

Note that the above result for $d = 1, 2$ has already been proved (in a stronger form) in Theorem 1.2 and in Proposition 3.4, part (i).

The basic results in this section can be extended to other norms in finite dimensional real spaces. Some norms, like the l_∞-norm, are simpler than the Euclidean one for this purpose, and it is not too difficult to determine the function $C(l_\infty^d, b, 1)$ (defined in the obvious way) precisely for many values of b. However, for any fixed $d \geq 2$ we are unable to determine any of the functions $C(R^d, b, 1)$ or $C(l_\infty^d, b, 1)$ for all b and this problem remains open.

Acknowledgement We would like to thank Yuval Peres for helpful discussions and comments.

References

[1] P. Bateman and P. Erdős, *Geometrical extrema suggested by a lemma of Besicovitch*, Amer. Math. Monthly 58 (1951), 306-314.

[2] J. H. Conway and N. J. A. Sloane, **Sphere packings, lattices and groups**, Springer Verlag, 1988.

[3] G. Katona, *Graphs, vectors, and inequalities in probability theory*, Mat. Lapok 20 (1969), 123-127.

[4] J. Leech, *The problem of the thirteen spheres*, Math. Gazette 40 (1956), 22-23.

[5] V. I. Levenshtein, *On bounds for packing in n-dimensional Euclidean space*, Doklady Akad. Nauk SSR 245 (1979), 1299-1303. English translation in: Soviet Math. Doklady 20 (1979), 417-421.

[6] A. M. Odlyzko and N. J. A. Sloane, *New bounds on the number of unit spheres that can touch a unit sphere in n dimensions*, J. Combinatorial Theory A 26 (1979), 210-214.