2 Typical sequences etc

2a Typical sequences

We introduce a shift operator

$$T : \{0,1\}^\infty \to \{0,1\}^\infty , \quad (Tx)(k) = x(k+1) .$$

It is continuous, onto, and not one-to-one (in fact, two-to-one). If f is a continuous function $\{0,1\}^\infty \to \mathbb{R}$ then $f \circ T$ is also a continuous function $\{0,1\}^\infty \to \mathbb{R}$, but in addition it is insensitive to the first coordinate $x(1)$ of $x \in \{0,1\}^\infty$. Likewise, $f \circ T^n$ is insensitive to $x(1), \dots, x(n)$. If F is a closed subset of $\{0,1\}^\infty$ then its inverse image $T^{-1}(F)$ is also a closed subset of $\{0,1\}^\infty$, but insensitive to the first coordinate. Likewise, $T^{-n}(F)$ is insensitive to $x(1), \dots, x(n)$. The same holds for open sets.

Similarly, for every measurable $A \subset \{0,1\}^\infty$ the set $T^{-1}(A)$ is also measurable, and $\mu(T^{-1}(A)) = \mu(A)$. (Hint: $T^{-1}(A) = \{0,1\} \times A$.) Thus, $\mu(T^{-n}(A)) = \mu(A)$.

2a1 Exercise. If $U \subset \{0,1\}^\infty$ is a nonempty open set then the set $\limsup_n T^{-n}(U)$ is comeager.

Prove it.

That is, $T^n x \in U$ infinitely often, for quasi all x.

It follows easily that the set $\{T^n x : n = 1,2,\dots\}$ is dense in $\{0,1\}^\infty$ for quasi all x.

Similarly, if $A \subset \{0,1\}^\infty$ is a measurable set of positive measure then the set $\limsup_n T^{-n}(U)$ is of full measure.[1] That is, $T^n x \in A$ infinitely often, for almost all x. The two approaches agree here. But…

[1]This fact follows from Kolmogorov's 0–1 law. Moreover, $\frac{1}{n}\sum_{k=1}^n \mathbb{1}_{T^{-n}(U)} \to \mu(U)$ almost surely by the ergodic theorem.

2a2 Exercise. If $U_n \subset \{0,1\}^\infty$ are nonempty open sets then the set $\limsup_n T^{-n}(U_n)$ is comeager.

Prove it.

In contrast, if $A_n \subset \{0,1\}^\infty$ are measurable sets of positive measure such that $\sum_n \mu(A_n) < \infty$ then $\sum_n \mu(T^{-n}(A_n)) < \infty$, and by the Borel-Cantelli lemma, $\limsup_n T^{-n}(A_n)$ is a null set.

A wonder: many sets of the form $\limsup_n T^{-n}(U_n)$ are null sets, and nevertheless, two such sets are never disjoint. Moreover, countably many such sets always have nonempty intersection.

2a3 Exercise. If $f : \{0,1\}^\infty \to \mathbb{R}$ is a continuous function then $\limsup_n f \circ T^n = \max f$ quasi-everywhere. (That is, the function $\limsup_n f \circ T^n$ is quasi-everywhere equal to the number $\max f = \max_{x \in \{0,1\}^\infty} f(x)$.)

Prove it.

Similarly, $\limsup_n f \circ T^n = \operatorname{ess\,sup} f$ almost everywhere for every measurable $f : \{0,1\}^\infty \to \mathbb{R}$. But...

2a4 Exercise. If $f_n : \{0,1\}^\infty \to \mathbb{R}$ are continuous functions then $\limsup_n f_n \circ T^n = \limsup_n (\max f_n)$ quasi-everywhere.

Prove it.

2a5 Exercise. Assume that $f_n : \{0,1\}^\infty \to \mathbb{R}$ are measurable functions, and $p_n \in (1,\infty)$ satisfy $\frac{p_n}{\log n} \to \infty$. Then

$$\limsup_n f_n \circ T^n \le \limsup_n \|f_n\|_{p_n} \quad \text{almost everywhere.}$$

Here $\|f_n\|_{p_n} = \big(\int |f_n|^{p_n} \big)^{1/p_n}$.

Prove it.

It can happen that $\limsup_n \|f_n\|_{p_n} < \limsup_n (\max f_n)$ for continuous f_n. (Try indicators of small closed-and-open sets.)

2a6 Exercise. If $f_n : \{0,1\}^\infty \to \mathbb{R}$ and $g_n : \{0,1\}^\infty \to \mathbb{R}$ are continuous functions such that $f_n \circ T^n - g_n \to 0$ uniformly then $\limsup_n g_n = \limsup_n (\max g_n)$ quasi-everywhere.

Prove it.

That holds also for $f_n \circ T^{k_n}$ provided that $k_n \to \infty$. In particular, recall 1e1(b):

$$g_n(x) = \frac{x(1) + \dots + x(n)}{n} \, ;$$

choose $k_n \to \infty$ such that $\frac{k_n}{n} \to 0$; note that $\max |g_n - g_n \circ T^{k_n}| \le \frac{2k_n}{n} \to 0$; now 1e1(b) follows.

Given a set $A \subset \{0,1\}^\infty$, we define its projection to $\{0,1\}^n$:

$$A[1:n] = \{x[1:n] : x \in A\}$$

where

$$x[1:n] = (x(1), \dots, x(n)) .$$

2a7 Exercise. If $U_n \subset \{0,1\}^\infty$ are open sets such that $U_n[1:n] = \{0,1\}^n$ then the set $\limsup_n U_n$ is comeager.

Prove it. Deduce 1e2(b) and 1e4(a) as special cases.

Consider the set $\{0,1\}_\infty = \bigcup_{n=1}^\infty \{0,1\}^n$ of all finite sequences. The concatenation $x \cdot y \in \{0,1\}_\infty$ of two finite sequences $x, y \in \{0,1\}_\infty$ is $(x_1, \dots, x_m, y_1, \dots, y_n)$ for $x = (x_1, \dots, x_m)$ and $y = (y_1, \dots, y_n)$. Given a map $f : \{0,1\}_\infty \to \{0,1\}_\infty$, we introduce for every n a set $U_n \subset \{0,1\}^\infty$ of all infinite sequences x that begin with the concatenation $x[1:n] \cdot f(x[1:n])$. Clearly, U_n is open (and closed), $U_n[1:n] = \{0,1\}^n$. By 2a7, the set $A_f = \limsup_n U_n$ is comeager. Note that $f(x[1:n])$ may be much longer than $x[1:n]$ (and the length of $f(x[1:n])$ may depend on $x[1:n]$).

The intersection of A_f over *all* f is of course empty (think, why). However, for countably many functions f the intersection is still comeager. In particular, all computable f are a countable set. Thus, a generic $x \in \{0,1\}^\infty$ satisfies the following:

for every computable $f : \{0,1\}_\infty \to \{0,1\}_\infty$,
for infinitely many n,
x begins with $x[1:n] \cdot f(x[1:n])$.

(Of course, the infinite set of n depends not only on x but also on f.)

We turn to products. Given $n_1, n_2, \dots \in \{1, 2, 3, \dots\}$ we have

$$\begin{gathered} \{0,1\}^\infty = \{0,1\}^{n_1} \times \{0,1\}^{n_2} \times \dots , \\ x = \big(x[1:n_1], x[n_1+1 : n_1+n_2], \dots \big) , \\ x[n : n+k] = \big(x(n), \dots, x(n+k) \big) . \end{gathered}$$

Accordingly, given nonempty $A_1 \subset \{0,1\}^{n_1}$, $A_2 \subset \{0,1\}^{n_2}$, $\dots$ we have $A_1 \times A_2 \times \dots \subset \{0,1\}^\infty$, and

$$\mu(A_1 \times A_2 \times \dots) = \mu_{n_1}(A_1) \mu_{n_2}(A_2) \dots$$

be the infinite product convergent (to a positive number) or divergent (to zero). Thus

$$\mu(A_1 \times A_2 \times \dots) > 0 \quad \Longleftrightarrow \quad \sum_{k=1}^{\infty} \bigl(1 - \mu_{n_k}(A_k)\bigr) < \infty \,,$$

which is closely related to the two Borel-Cantelli lemmas.

The product set $A_1 \times A_2 \times \dots$ is closed; it is also nowhere dense provided that $A_k \ne \{0,1\}^{n_k}$ for infinitely many k (otherwise it is closed-and-open). Interestingly, every nowhere dense set is contained in some nowhere dense product set (see below).

Consider

$$A[n:\infty] = \{x[n:\infty] : x \in A\} \,, \qquad x[n:\infty] = (x(n), x(n+1), \dots) \,.$$

2a8 Exercise. If A is nowhere dense then $A[n:\infty]$ is nowhere dense.

Prove it.

Given a nowhere dense $A \subset \{0,1\}^\infty$, we take n_1 and $x_1 \in \{0,1\}^{n_1}$ such that $x[1:n_1] \ne x_1$ for all $x \in A$. Then we take n_2 and $x_2 \in \{0,1\}^{n_2}$ such that $x[1:n_2] \ne x_2$ for all $x \in A[n_1+1:\infty]$, that is, $x[n_1+1:n_1+n_2] \ne x_2$ for all $x \in A$. And so on. We get $A \subset A_1 \times A_2 \times \dots$ where $A_k = \{0,1\}^{n_k} \setminus \{x_k\}$.

What if A is a closed null set? Then it is nowhere dense and therefore contained in some nowhere dense product set. However, what about the measure of this product set? Can we make it zero? Or at least, small? I do not know.

2b Random walk and conditioning

Functions $S_n : \{0,1\}^\infty \to \mathbb{Z}$,

$$S_n(x) = \sum_{k=1}^{n} (2x(k) - 1)$$

are a random walk, — a random element of the set of all sequences $(s_0, s_1, s_2, \dots)$ such that $s_0 = 0$ and $s_{n+1} - s_n = \pm 1$. By 1e1,

$$\lim_n \frac{S_n(x)}{n} = 0 \quad \text{for almost all } x,$$

but

$$\liminf_n \frac{S_n(x)}{n} = -1 \,, \quad \limsup_n \frac{S_n(x)}{n} = +1 \quad \text{for quasi all } x.$$

The set

$$A = \{x : \forall n \; S_n(x) \ge 0\}$$

is null and meager. A challenge: what happens to the random walk under the condition $x \in A$?

The topological approach

The set A is closed in $\{0,1\}^\infty$, therefore compact, and may be treated as a compact metrizable space. Accordingly, meager and comeager subsets of A are well-defined (even though they all are meager in $\{0,1\}^\infty$).

The technique of Sect. 2a does not help, since the shift T fails to map A to A. We turn to a more general technique.

We leave $\{0,1\}^\infty$ and turn to the set X of all sequences $(x_0, x_1, x_2, \dots)$ such that $x_0 = 0$ and $x_{n+1} - x_n = \pm 1$ for $n = 0, 1, \dots$. We transfer from $\{0,1\}^\infty$ to X the metrizable topology. The pointwise convergence in $\{0,1\}^\infty$ (recall 1d2(b)) turns into the pointwise convergence in X (think, why):

$$x_n \xrightarrow[n\to\infty]{} x \quad \Longleftrightarrow \quad \forall k \, \Big(x_n(k) \xrightarrow[n\to\infty]{} x(k) \Big) \, .$$

Still, a neighborhood of x may be takes as $\{y : y(1) = x(1), \dots, y(n) = x(n)\}$; and 1d5 still applies. The same holds for $X^+ = \{x \in X : \forall k \; x(k) \ge 0\}$.

2b1 Lemma. Let X be a compact metrizable space, $f_n : X \to \mathbb{R}$ continuous functions, and

$$c = \inf_{U,n} \sup_{x \in U, k} f_{n+k}(x)$$

(be it finite or infinite) where U runs over all nonempty open sets in X. Then

$$\limsup_n f_n(x) \ge c \quad \text{for quasi all } x \in X \, .$$

Proof. It is sufficient to prove that for all $\varepsilon > 0$ and n the set

$$A_{\varepsilon,n} = \{x : \sup_k f_{n+k}(x) \le c - \varepsilon\}$$

is nowhere dense. Given a nonempty open set U we note that $\sup_{x \in U,k} f_{n+k}(x) \ge c$, take $x \in U$ and k such that $f_{n+k}(x) > c - \varepsilon$ and observe that the nonempty open subset $V = \{y \in U : f_{n+k}(y) > c - \varepsilon\}$ of U does not intersect $A_{\varepsilon,n}$. □

2b2 Exercise. Deduce 2a4 from 2b1 as a special case.

We apply 2b1 to functions f_n on X^+, $f_n(x) = x(n)/n$, and get

$$\limsup_n \frac{x(n)}{n} = 1 \quad \text{for quasi all } x \in X^+ .$$

(It is not the same as the similar fact for X.) We also apply 2b1 to $x \mapsto -x(n)$ and get

$$\liminf_n x(n) = 0 \quad \text{for quasi all } x \in X^+ .$$

On the other hand,

$$n - x(n) \xrightarrow[n\to\infty]{} \infty \quad \text{for quasi all } x \in X^+$$

(since it is increasing and cannot be bounded).

2b3 Exercise. Prove that

$$\liminf_n \frac{n - x(n)}{\log\log\log n} = 0 \quad \text{for quasi all } x \in X^+ .$$

The probabilistic approach

Regretfully, no general definition of a conditional distribution on a null set is available. Not even on a closed null set. However, reasonable ad hoc definitions are available for many special cases, including our $X^+ \subset X$.

We approximate the null set X^+ with sets X_n^+ of positive probability,

$$X_n^+ = \{x \in X : \min(x(0), \dots, x(n)) \ge 0\} ;$$
$$X_n^+ \downarrow X^+ \quad \text{as } n \to \infty ,$$

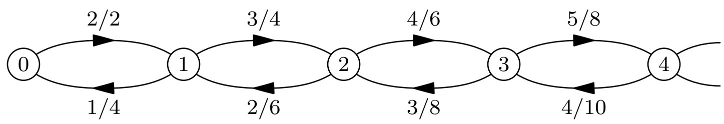
and define conditional probabilities by the formula

$$\text{(2b4)} \qquad \mathbb{P}\big(A \big| X^+\big) = \lim_n \mathbb{P}\big(A \big| X_n^+\big) = \lim_n \frac{\mathbb{P}\big(A \cap X_n^+\big)}{\mathbb{P}\big(X_n^+\big)}$$

not for all measurable A (otherwise we would get $\mathbb{P}\big(X^+ \big| X^+\big) = 0$) but for all "elementary sets" $A \in \mathcal{F}_1 \cup \mathcal{F}_2 \cup \dots$ where $\mathcal{F}_n$ consists of all sets of the form $\{x \in X : (x(0), \dots, x(n)) \in B\}$ for arbitrary $B \subset \mathbb{Z}^{n+1}$. It appears that the limit (2b4) exists for all these A, and extends uniquely to a probability measure on X^+. In particular,[1]

$$\mathbb{P}\big(x(0)=a_0,\dots,x(n)=a_n \,\big|\, X^+\big) = \frac{a_n+1}{2^n} = \\ = \frac{a_1+1}{2(a_0+1)}\cdot\frac{a_2+1}{2(a_1+1)}\cdots\frac{a_n+1}{2(a_{n-1}+1)}$$

whenever $a_0=0$, $a_k-a_{k-1}=\pm1$ and $a_k\ge 0$ for $k=1,\dots,n$. Thus the conditional random walk is a Markov chain:



Its asymptotic behavior is well-known[2] and far from being trivial. Almost all $x\in X^+$ satisfy

$$\liminf_n \frac{\log\frac{x(n)}{\sqrt n}}{\log\log n} = -1\,, \quad \limsup_n \frac{x(n)}{\sqrt n \log\log n} = \sqrt2\,.$$

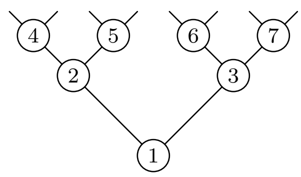
Thus, for every $\varepsilon>0$ they satisfy

$$\frac{1}{\log^{1+\varepsilon} n} \le \frac{x(n)}{\sqrt n} \le (\sqrt2+\varepsilon)\log\log n \quad \text{for all } n > N(\varepsilon,x)\,.$$

2c Trees

The set $\{1,2,3,\dots\}$ turns into the binary tree T_2, being endowed with the binary relation

(2c1) $$n\in\{2m,2m+1\}$$



interpreted as "n is a child of m" (that is, "m is the parent of n"). Thus, $\{0,1\}^\infty$ may be thought of as $\{0,1\}^{T_2}$. Every (infinite) branch of the tree leads to a map $\{0,1\}^{T_2}\to\{0,1\}^\infty$. It means, we choose a subsequence. For example, the leftmost branch corresponds

[1] $\mathbb{P}\big(x(0)=a_0,\dots,x(n)=a_n, x(n+1)\ge 0,\dots,x(n+k)\ge 0\big) = 2^{-n}\mathbb{P}\big(\max(x(0),\dots,x(k))\le a_n\big) = 2^{-n}\mathbb{P}\big(-a_n-1\le x(k)\le a_n\big)$ (using reflection); $\mathbb{P}\big(-a_n-1\le x(k)\le a_n\big)\sim\frac{2}{\sqrt{2\pi k}}(a_n+1)$ as $k\to\infty$ (using the normal approximation); in particular (for $n=0$), $\mathbb{P}\big(X_k^+\big)\sim\frac{2}{\sqrt{2\pi k}}$; thus $\mathbb{P}\big(x(0)=a_0,\dots,x(n)=a_n\,\big|\,X^+_{n+k}\big)\to 2^{-n}(a_n+1)$ as $k\to\infty$.

[2] B.M. Hambly, G. Kersting, A.E. Kyprianou (2003), "Law of the iterated logarithm for oscillating random walks conditioned to stay non-negative", Stochastic Processes and their Applications **108** 327–343.

to the subsequence $\big(x(1), x(2), x(4), \dots\big) = \big(x(2^{k-1})\big)_k$, while the rightmost branch to $\big(x(1), x(3), x(7), \dots\big) = \big(x(2^k - 1)\big)_k$.

2c2 Exercise. Let $S_1 \subset S_2$ be countable sets. Prove that the restriction map $\{0,1\}^{S_2} \to \{0,1\}^{S_1}$ is genericity preserving (in the sense of 1f3).

Informally, if a sequence is generic then its subsequence is also generic. (This is about $\big(x(n)\big)_n$ and $\big(x(n_k)\big)_k$ provided that $(n_k)_k$ is not dependent on x, of course.) We may apply it to countably many subsequences. However, the binary tree has uncountably many branches. What about existence of an atypical branch, say, a branch with $\sum_k x(n_k) < \infty$, or even $\sum_k x(n_k) = 0$?

The probabilistic approach

We want to find the probability

$$\mathbb{P}\big(\exists (n_k)_k \; \forall k \;\; x(n_k) = 0\big) \tag{2c3}$$

(where $(n_k)_k$ runs over all branches). This is a reformulation of a well-known question about the simple branching (or Galton-Watson) process. The probability (2c3) is the non-extinction probability. The extinction probability is the least root of the equation

$$\frac{\theta^2 + 1}{2} = \theta \,, \quad 0 \le \theta \le 1 \,,$$

and is equal to 1. Thus, the probability (2c3) is 0.

It happens because the branching process is critical. Consider now the ternary tree T_3. Here the branching process is supercritical; the equation becomes

$$\frac{\theta^3 + 1}{2} = \theta \,; \quad \theta = \frac{\sqrt{5} - 1}{2} \,;$$

the probability (2c3) is now $1 - \theta = \frac{3 - \sqrt{5}}{2}$. Of course, for a given branch $(n_k)_k$ the event $\sum_k x(n_k) = 0$ is of zero probability; however, existence of such "atypical" branch is of positive probability.

The topological approach

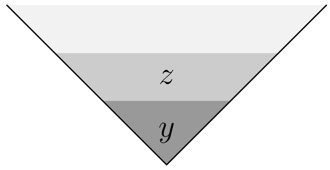
Given $A \subset \{0,1\}^\infty$, we introduce $\tilde{A} \subset \{0,1\}^\infty$ by

$$\tilde{A} = \{x : \exists (n_k)_k \;\; \big(x(n_k)\big)_k \in A\}$$

(where $(n_k)_k$ runs over all branches).

2c4 Proposition. If A is meager then $\tilde{A}$ is meager.

Proof (sketch). Clearly, if $A = A_1 \cup A_2 \cup \dots$ then $\tilde{A} = \tilde{A}_1 \cup \tilde{A}_2 \cup \dots$; thus we assume that A is nowhere dense and prove that $\tilde{A}$ is nowhere dense. We use 1d5. Given an initial segment y of a function $T_2 \to \{0,1\}$, we seek its continuation z incompatible with $\tilde{A}$. For every (finite, maximal) branch of y we choose the corresponding portion of z to be a function of the level number only. This is possible since A is nowhere dense. □



Clearly, the argument applies not only to T_2 but also to T_3. The topological approach is quite pessimistic: it claims that extinction is inevitable in all cases! Likewise, percolation to infinity is impossible in all dimensions (and even all locally finite graphs). Curiously enough, on the plane we get infinitely many white and black contours around the origin that are exactly square! The probabilistic theory of percolation is *much* more deep and complicated.

Likewise, the topological approach claims that the random walk is recurrent in all dimensions; but probabilistically, it is recurrent in dimensions 1 and 2 but transient in dimensions $3, 4, \dots$ (Polya).

2d Graphs

A point $x \in \{0,1\}^\infty$ may also be treated as a graph. To this end we fix a countable set $\{v_1, v_2, \dots\}$ of vertices and connect v_n with v_{n+k} by an edge if and only if $x\big((2n-1)2^{k-1}\big) = 1$; here $1 \le n < n+k < \infty$. Alternatively we may deal with $\{0,1\}^S$ where S is the set of all unordered pairs of (different) vertices. Anyway, we get a random element of the set of all graphs (undirected, with no loops and multiple edges) on the given countable set of vertices. For each pair of vertices we decide whether they are connected by edge or not, independently of other choices.

Is the random graph connected? Yes, it is, in both approaches (topological and probabilistic). Moreover,

(2d1) the distance between two vertices never exceeds 2.

For example, $\operatorname{dist}(v_1, v_2) > 2$ when

$$x(2)x(3) = x(4)x(6) = x(8)x(12) = x(16)x(24) = \dots = 0\,;$$

the set of such x is both null and meager, since it is the product with infinitely many factors of probability 3/4.

For a similar reason

(2d2) no vertex is on distance 1 from all other vertices.

Two graphs are called isomorphic if some permutation (bijection to itself) of $\{v_1, v_2, \dots\}$ transforms one graph to another.

The two properties (2d1), (2d2) fail to ensure isomorphism. For example, here are two nonisomorphic graphs satisfying these properties:

(a) v_n is connected by an edge with v_{n+k} if and only if $k > 1$;

(b) v_n is connected by an edge with v_{n+k} if and only if $n+1$ divides $n+k+1$.

They are not isomorphic; in (a), in contrast to (b), each vertex is connected by an edge with all but finitely many vertices.

A challenge: are all random graphs isomorphic? It means, (1) is there a comeager equivalence class? (2) is there an equivalence class of full measure? And if (1) and (2) hold, then we ask (3) is it the same equivalence class in both cases?

Here is a far-reaching strengthening of (2d1), (2d2):

(2d3) For every pair (V_1, V_2) of disjoint finite sets of vertices
there exists a vertex outside $V_1 \cup V_2$
connected by an edge with every vertex of V_1 but no vertex of V_2.

This property is satisfied almost everywhere and quasi-everywhere, since (as before) it is violated only on a product set with infinitely many factors of the same probability less than 1.

Therefore such graphs exist! (Do you see an example?)

2d4 Lemma. If two graphs satisfy (2d3) then they are isomorphic.

Proof (sketch). Given an isomorphism between their finite subgraphs, we can extend it to an isomorphism between larger finite subgraphs. Moreover, we can add to the first finite subgraph any point we want; and the same for the second subgraph. □

We see that the set of all graphs satisfying (2d3) is an equivalence class, comeager and of full measure.[1]

Amazingly, we can take a different product measure on $\{0,1\}^\infty$, $\mathbb{P}\big(x(n)=1\big) = p \in (0,1)$, and get different random graphs in the same equivalence class!

[1]See also Wikipedia, "Rado graph".

Hints to exercises

2a1: $\cap_k T^{-(n+k)}(F)$ is nowhere dense (here F is the complement of U).

2a2: similar to 2a1.

2a3: $U = \{x : f(x) > \max f - \varepsilon\}$.

2a4: similar to 2a3.

2a5: $\mathbb{P}\big(f \ge a\big) \le \frac{1}{a^p} \int |f|^p$.

2a6: use 2a4.

2a7: $\cap_k F_{n+k}$ is nowhere dense (here F_n is the complement of U_n).

2a8: $A[2:\infty]$ is the union of two nowhere dense sets.

2b3: $f_n(x) = -\frac{n - x(n)}{\log\log\log n}$.

Index