Journal de Théorie des Nombres
de Bordeaux **14** (2002), 415–424

Uniform distribution modulo one and binary search trees

par Michel DEKKING et Peter van der WAL

On the occasion of the 65th birthday of Michel Mendès France

Résumé. On peut construire à partir d'une suite $x = (x_k)_{k=1}^{\infty}$ de nombres distincts de l'intervalle $[0,1]$ un arbre binaire en plaçant successivement ces nombres sur les nœuds selon un algorithme "gauche-droite" (cela revient à classer les nombres selon l'algorithme Quicksort). On note $H_n(x)$ la hauteur de l'arbre obtenu à partir des nombres $x_1, \dots, x_n$. Il est évident que

$$\frac{\log n}{\log 2} - 1 \leq H_n(x) \leq n-1.$$

Si la suite x est obtenue comme valeurs de variables aléatoires indépendantes uniformes sur $[0,1]$, alors on sait qu'il existe $c > 0$ tel que $H_n(x) \sim c \log n$, $(n \to \infty)$, presque-sûrement Récemment, Devroye et Goudjil ont montré que si les x sont les suites de Weyl, i.e., $x_k = \{\alpha k\}$, $k = 1, 2, \dots$, où α est une variable aléatoire uniforme sur $[0,1]$, alors

$$H_n(x) \sim (12/\pi^2) \log n \log\log n, \quad n \to \infty,$$

en probabilité.
Dans ce papier nous considérons la classe de *toutes* les suites x uniformément réparties pour lesquelles nous montrons que l'on a nécessairement $H_n(x) = o(n)$ quand $n \to \infty$.

Abstract. Any sequence $x = (x_k)_{k=1}^{\infty}$ of distinct numbers from $[0,1]$ generates a binary tree by storing the numbers consecutively at the nodes according to a left-right algorithm (or equivalently by sorting the numbers according to the Quicksort algorithm).
Let $H_n(x)$ be the height of the tree generated by $x_1, \dots, x_n$. Obviously

$$\frac{\log n}{\log 2} - 1 \leq H_n(x) \leq n-1.$$

If the sequences x are generated by independent random variables having the uniform distribution on $[0,1]$, then it is well known that there exists $c > 0$ such that $H_n(x) \sim c \log n$ as $n \to \infty$ for almost all sequences x. Recently Devroye and Goudjil have shown that

Manuscrit reçu le 27 décembre 2000.

if the sequences x are Weyl sequences, i.e., defined by $x_k = \{\alpha k\}$, $k = 1, 2, \ldots$, and α is distributed uniformly at random on $[0,1]$ then

$$H_n(x) \sim (12/\pi^2) \log n \log\log n$$

as $n \to \infty$ in probability.

In this paper we consider the class of *all* uniformly distributed sequences x, and we show that the only behaviour that is excluded by the equidistribution of x is that of the worst case, i.e., for these x we necessarily have $H_n(x) = o(n)$ as $n \to \infty$.

Michel DEKKING, Peter VAN DER WAL
Thomas Stieltjes Institute for Mathematics
and Delft University of Technology
Faculty ITS, Department CROSS
Mekelweg 4, 2628 CD Delft
The Netherlands
E-mail : F.M.Dekking@its.tudelft.nl, vanderwal@eurandom.tue.nl