§3.8. Assignment 2, Fall 2020

Exercise 3.39. THE AMOUNT OF REGULARITY IN A RANDOM PERMUTATION. Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ be a uniform random permutation of $(1, \ldots, n)$. The amount of regularity is defined by

$$A_n = \sum_{i<j} 1_{[\sigma_i < \sigma_j]} .$$

Determine exact expressions for $\mathbb{E}\{A_n\}$ and $\mathbb{E}\{A_n^2\}$. Furthermore, show that $\mathbb{V}\{A_n\} \sim cn^3$ for some constant c, and that $A_n/\mathbb{E}\{A_n\} \to 1$ in probability.

Exercise 3.40. A STRATIFIED RANDOM BINARY SEARCH TREE. We are given data $U_1, \ldots, U_n$ that are i.i.d. uniform $[0,1]$. Let $n = 2^h$ for some integer h. We artificially add the data points $i/2^h$, $1 \le i \le 2^h - 1$, and construct a complete binary tree with these artificial data. The keys $U_1, \ldots, U_n$ are then added, one by one, by standard insertion. Call the resulting tree with $2n-1$ nodes the stratified random binary search tree. Let H_n be its height. Show that this tree is nearly perfectly balanced by proving that $\mathbb{E}\{H_n\} \sim \log_2 n$.

Exercise 3.41. THE EXACT DISTRIBUTION OF THE DEPTH D_n. Let D_n be the depth of the node with the largest time stamp in a random binary search tree on n nodes. Convince yourself that $Q_{n,k} \stackrel{\text{def}}{=} n!\mathbb{P}\{D_n = k\}$ is integer-valued. Using integer-valued computations only, derive a recursion for $Q_{n,k}$ that will permit you to compute all values $Q_{n,k}, 0 \le k \le n-1$, and thus the entire distribution, in time $O(n^2)$.

Exercise 3.42. THE LARGEST HIDDEN COMPLETE BINARY TREE. A complete binary tree of (C_h) height h contains $2^{h+1} - 1$ nodes. A random binary search tree on n nodes can contain none or several terminal occurrences of C_h, where "terminal" refers to the fact that all leaves of C_h are also leaves in the random binary search tree. For example, every leaf represents a terminal occurrence of C_0. Let H_n be the largest height h such that there is a terminal instance of C_h in the random binary search tree. It is known that $H_n/\log_2 \log n \to 1$ in probability. We are asking you to "almost" show this. Let N_h be the number of occurrences of C_h. Then show the following: for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{E}\{N_h\} = \begin{cases} 0 & \text{if } h \ge (1+\epsilon)\log_2 \log n \ ; \\ \infty & \text{if } h \le (1-\epsilon)\log_2 \log n \ . \end{cases}$$