Homework #9
Due Date:
Reading Assignment: 9.1–9.4
Problems:

1. 8–31

2. 9–2

3. 9–13

4. 9–14

5. 9–34

6. Let Y_n be the random variable describing the number of the probes for unsuccessful search on a binary search tree that contains n keys. The purpose of this problem is to compute $E(Y_n)$ and $Var(Y_n)$ using generating functions. The generating function for Y_n is defined by

$$G_n(Z) = \sum_k p_{nk} z^k,$$

where $p_{nk} = \Pr(Y_n = k)$, the probability that k probes are needed to do an unsuccessful search in an n-key tree. The sample space for Y_n is

$$S = \{(x_1, \ldots, x_{n-1}, x_n; y)\},$$

where each $(x_1, \ldots, x_{n-1}, x_n; y)$ is one of the $(n+1)!$ permutations of 1, 2, …, $n+1$. Here $x_1, \ldots, x_n$ represents the n keys already inserted, and y is the key that will be searched for unsuccessfully.

We want to develop a recurrence for p_{nk} by deriving a relationship between Y_n and Y_{n-1}. To do this, we use the model described in class: Another way to think of $s = (x_1, \ldots, x_{n-1}, x_n; y)$ is to regard $s' = (x_1, \ldots, x_{n-1}; y)$ as one of the $n!$ permutations of $1, 2, \ldots, n$, and to regard x_n as one of the $n+1$ fractions $\frac{1}{2}, \frac{3}{2}, \ldots, (n + \frac{1}{2})$.

(a) Show that

$$Y_n(s) = \begin{cases} Y_{n-1}(s') + 1 & \text{when } x_n = y \pm \frac{1}{2}, \\ Y_{n-1}(s') & \text{when } x_n \text{ is one of the other } n-1 \text{ values.} \end{cases}$$

(This isn't hard if you think about it.) Using this key fact, we have

$$\begin{aligned} p_{nk} = {} & \Pr\Big(Y_{n-1} = k-1 \text{ and } \big(x_n = y - \frac{1}{2} \text{ or } x_n = y + \frac{1}{2}\big)\Big) \\ & + \Pr\Big(Y_{n-1} = k \text{ and } x_n \text{ is one of the other } n-1 \text{ values}\Big). \end{aligned}$$

(b) Using independence, derive a recurrence for p_{nk} in terms of $p_{n-1,k}$ and $p_{n-1,k-1}$. Add the appropriate δ term to make the recurrence hold for all values of n and k and substitute the recurrence into the definition of $G_n(z)$. Compute $E(Y_n)$ and $Var(Y_n)$. As a check, the variance is

$$Var(Y_n) = 2H_{n+1} - 4H_{n+1}^{(2)} + 2.$$

(c) Express $E(Y_n)$ and $Var(Y_n)$ in the form $f(n)+c+g(n)+O(1/n)$, where c is a constant and $f(n)$ and $g(n)$ are some elementary functions. (e.g., n, e^n, $\log n$, $\log n/n$).