Dennis Kuzminer
CSCI-UA 310-001 PS6

7. $E[X],\ X = X_1 + ... + X_n \rightarrow$ where X_i is the number of spins you make to get the ith distinct number after you have already gotten i - 1 distinct numbers.
$E[X_1 + ... + X_n] = E[X_1] + ... + E[X_n] \rightarrow$
Let's compare the values of the individual expectations for n = 10 and derive a trend.
Clearly, $E[X_1] = 1$
If 1 number is achieved, and each number has the same probability of being chosen, p of choosing a new number = 9/10; therefore, $E[X_2] = \frac{10}{9}$
If 2 number is achieved, and each number has the same probability of being chosen, p of choosing a new number = 8/10; therefore, $E[X_3] = \frac{10}{8}$
Generalizing, we can see that $E[X_i] = \frac{10}{10-i},\ i = [0,\ ...,\ 10) \rightarrow E[X_i] = \frac{n}{n-i},\ i = [0,\ ...,\ n) \rightarrow$
$E[X] = \sum_{i=0}^{n-1} \frac{n}{n-i} \rightarrow n \sum_{i=0}^{n-1} \frac{1}{n-i} \rightarrow$ Writing out the terms, we can rearrange this sum to $n \sum_{i=1}^{n} \frac{1}{i} \rightarrow$
We can estimate this with an integral, $\int_1^n \frac{1}{x}\ dx + M,\ M = max(1/1,\ 1/n) = 1 \rightarrow$
$ln(n) - ln(1) + 1 + c_1 \rightarrow ln(n) + c_2,\ c_2 = 1 + c_1 \rightarrow$ Putting back into the equation and generalizing c, we get $E[X] = n(ln(n) + c) \rightarrow nln(n) + nc$
$|E[X] - nln(n)|$ is bounded by nc. Additionally, we see that nc grows at most the rate of n.
Meaning, we can generalize be saying $E[X] = nln(n) + O(n)$