CS302 - Problem Set 5

If you'd like a refresher on sample spaces, and calculating the probability of a sequence of random events, see this text, pages 767-768, Steps 1-3.

This looks like a lot of questions, but it is because I am giving you a step-by-step guide.

1. Suppose you are searching an array A of length n for an element with value t. You may assume A has no repeated elements. The strategy *sampling without replacement* works as follows: Let $T = \{1, 2, \ldots, n\}$. Choose an element $i \in T$ at random, and check if $A[i] = t$. If it is, return i. If it is not, remove i from T (i.e. set T to equal $T - \{i\}$). Repeat by choosing an element of the updated set T at random. Repeat until you sample an index i such that $A[i] = t$ and t is found. The strategy *sampling with replacement* is similar except there is no update to T, so throughout the algorithm, T is always equal to $\{1, 2, \ldots, n\}$. (Replacement refers to whether the guessed index is placed back ("replaced") into the set T or not after it is guessed.)

   For the following problems you should assume that $n = 3$ and $A[1] = t$ (the index of the element of interest is 1.) Problems (a)-(h) deal with sampling without replacement, and problems (i)-(l) deal with sampling with replacement.

   (a) What is the sample space S for sampling without replacement? (Please list all elements of the sample space.)

   (b) Let $p : S \to \mathbb{R}$ be the function that gives the probability of each element of S occurring. List $p(s)$ for each $s \in S$.

   (c) Let $R : S \to \mathbb{R}$ be the function that gives the number of rounds that occur for each element of the sample space. (If you make g guesses before finding the item you are looking for, the number of rounds is g.) What is R for each element of S?

   (d) Using your answers to the previous parts, calculate (directly, without using indicator random variables) the average number of rounds for the sampling without replacement strategy for an array of length 3 if the element you are looking for is in the 1st position.

   (e) How would your answer to (d) change if the element you were looking for were not in the 1st position?

   (f) Challenge (Wait to turn page until attempted for challenge): Write R as a weighted sum of indicator random variables.

   Or: turn to next page.

Consider the indicator random variables $X_r : S \to \{0, 1\}$ where

$$X_r(s) = \begin{cases} 1 & \text{if } s \text{ has at least } r \text{ rounds} \\ 0 & \text{if } s \text{ less than } r \text{ rounds.} \end{cases} \tag{1}$$

We can write R as a weighted sum of these indicator random variables:

$$R(s) = \sum_{r=1}^{3} \alpha_r X_r(s) \tag{2}$$

where $\alpha_r \in \mathbb{R}$. What values of $\alpha_1, \alpha_2, \alpha_3$ make Eq. (2) true?

(g) What is the probability that at least r rounds occur?

(h) Using linearity of expectation, properties of indicator random variables, and your answers to the previous two questions, recalculate $\mathbb{E}[R]$ and check that your answer is the same as before.

(i) What is the sample space S' for search with replacement? Please describe the set in words or using mathematical notation.

(j) Let $R : \mathbb{S}' \to \mathbb{R}$ be the function that gives the number of rounds that occur for each element of the sample space. Write R as a weighted sum of indicator random variables X_r (where X_r is defined above).

(k) With search with replacement, what is the probability that at least r rounds occur?

(l) Using linearity of expectation and properties of indicator random variables, what is the average number of rounds in search with replacement in an array of size 3?

(m) Please comment on the advantages or disadvantages of using sampling with or without replacement.

2. Approximately how long did you spend on this assignment (round to the nearest hour)?