Probability & Statistics for EECS: Homework #06

Due on Mar 26, 2023 at 23:59

Name: **Penghao Wang**
Student ID: 2021533138

Problem 1

Firstly we denote an indicator I_j, which is 1 if the j-th type toy is selected, and 0 otherwise. Then we denote that the total number of distinct toy types is X, we then have

$$X = \sum_{j=1}^{n} I_j.$$

We then find the $E(X)$, which is

$$E(X) = E(\sum_{j=1}^{n} I_j) = \sum_{j=1}^{n} E(I_j).$$

We can denote that the probability of the j-th type toy is head is p_j, then we have

$$E(I_j) = p_j.$$

So we have

$$E(X) = \sum_{j=1}^{n} p_j.$$

As $p_j = (1 - (1 - \frac{1}{n})^t)$, where t is the number that we totally collected toys. So we have

$$E(X) = \sum_{j=1}^{n} (1 - (1 - \frac{1}{n})^t) = n(1 - (1 - \frac{1}{n})^t)$$

Problem 2

We denote an indicator A_i, which is 1 if the i-th block not equal the $(i+1)$-th block, and 0 otherwise, and A is the total number of runs, then we have that

$$A = \sum_{i=1}^{n-1} A_i + 1.$$

So we have that

$$E(A) = E(\sum_{i=1}^{n-1} A_i + 1) = 1 + \sum_{i=1}^{n-1} E(A_i).$$

Then we denote event B_j is the jth block is different from $j+1th$, then $P(B_j) = p(1-p)+(1-p)p = 2p(1-p)$. So we have that $E(A) = 1 + \sum_{i=1}^{n-1} P(B_i) = 1 + 2(n-1)p(1-p)$.

Problem 3

1. Fristly we need to consider let the last one be the tagged elk, then we need to make k-1 captures. We have

$$P(X=k) = \frac{\binom{n}{m-1}\binom{N-n}{k}}{\binom{N}{m+k-1}} * \frac{n-(m-1)}{N-(m-1)-k}$$
$$= \frac{\binom{m+k-1}{m-1}\binom{N-m-k}{n-m}}{\binom{N}{n}}$$

.

Then the total number of elk in the new sample, as $Y = X + m$, then we have that

$$P(Y=y) = P(X=y-m) = \frac{\binom{n}{m-1}\binom{N-n}{y-m}}{\binom{N}{y-1}} * \frac{n-m+1}{N-y+1}$$

2. Define that untaggd elk are labeled 1, 2, ..., N - n, then we define that $X_1, X_2, ..., X_m$ are the number of untaggd elk before the first tagged elk, then number between first and second tagged elk, ..., then we have that $X_1 = I_1 + ... + I_{N-n}$, I_j is the indicator of untagged elk j that is captured before tagged elk. Then we have that $E(I_j) = \frac{1}{n+1}$, so we have $E(X_1) = \frac{N-n}{n+1}$, then we have that $E(X_j) = \frac{N-n}{n+1}$ for all $j = 1, 2, ..., m$. So we have

$$E(X) = \frac{m(N-n)}{n+1}$$

, as $Y = X + m$, we have

$$E(Y) = E(X+m) = \frac{m(N+1)}{n+1}.$$

3. Suppose that $E[Y]$ is an integer, and sample with fixed size, we have that the number of tagged elk is distribute to hypergeometric distribution, and we deonte that tagged elk in the sample is Z, we have that $E[Z] = \frac{m(N+1)}{n+1} * \frac{n}{N} = m * \frac{1+\frac{1}{N}}{1+\frac{1}{n}} < m$ So is less than m.

Problem 4

1.

Problem 5

To distribute the 14 balls into 5 boxes, we firstly denote event A be all the cases to put 14 balls into 5 boxes, B be the event all the cases to put 14 balls into boxes, with striction that one can at most put balls. Event C_i be that the $i-th$ box has balls more than 6. Then we have

$$B^c = \cup_{i=1}^{5} C_i.$$

Then we have with inclusion-exclusion,

$$P(B^c) = \sum_i P(C_i) - \sum_{i<j} P(C_i \cap C_j) + ... + (-1)^{5+1} P(C_1 \cap ... \cap C_5)$$

As for $P(C_i)$, we have that firstly take 7 balls into the i th box, then distribute the left 7 balls into 5 boxes, that is $P(C_i) = \binom{11}{4}$. Then as for $P(C_i \cap C_j)$, we have that distribute 7 balls into a box and the left 7 balls into another box, that is only 1 cases.
So we have that

$$\begin{aligned} P(B^c) &= \sum_i P(C_i) - \sum_{i<j} P(C_i \cap C_j) + ... + (-1)^{5+1} P(C_1 \cap ... \cap C_5) \\ &= \frac{5 * \binom{11}{4} - 10 * 1 + 0 - 0 + 0}{\binom{18}{4}} \\ &= \frac{1640}{3060}. \end{aligned}$$

So we have that ways of B is $A - B^c = 3060 - 1640 = 1420$.
So there are totally 1420 ways to distribute 14 balls into 5 boxes, with the restriction that one box can at most put 6 balls.