COMP90056 Stream Computing and Applications
Assignment A

# Theory question

Student Name: Zhuohan Xie
Student Number: 871089

After we buy $n \ln n$ boxes, the expected value of number of coupons that have been found at most once should approximately be $1 + \frac{n \ln n}{n-1}$ .

Considering a particular coupon, after we bought $n \ln n$ boxes, the probability we didn't find it is $\left(1 - \frac{1}{n}\right)^{n \ln n}$, and the probability we found once is $\frac{1}{n} \times \left(1 - \frac{1}{n}\right)^{n \ln n - 1} \times \binom{n \ln n}{1}$, since these two cases are independent, therefore, we can add them to get the probability that we found this coupon at most once after we bought $n \ln n$ boxes, which is

$$\left(1 - \frac{1}{n}\right)^{n \ln n} + \frac{1}{n} \times \left(1 - \frac{1}{n}\right)^{n \ln n - 1} \times \binom{n \ln n}{1} = \left(1 - \frac{1}{n}\right)^{n \ln n} \times \left(1 + \frac{\ln n}{1 - \frac{1}{n}}\right)$$

$$= \left(1 - \frac{1}{n}\right)^{n \ln n} \times \left(1 + \frac{n \ln n}{n-1}\right).$$

And there are n coupons, each of them has the same probability that have been found at most once. Therefore, the expected value of number of coupons that have been found at most once should be $\left(1 - \frac{1}{n}\right)^{n \ln n} \times \left(1 + \frac{n \ln n}{n-1}\right) \times n$.

Of course, we want to simplify this equation by using $1 - x \le e^{-x}$ inequality to an upper bound

$$n \times \left(1 - \frac{1}{n}\right)^{n \ln n} \le n \times e^{-\frac{1}{n} \times n \ln n}$$

$$= n \times e^{-\ln n}$$

$$= n \times \frac{1}{n}$$

$$= 1$$

Therefore, the final result should be $\left(1 - \frac{1}{n}\right)^{n \ln n} \times \left(1 + \frac{n \ln n}{n-1}\right) \times n \le 1 \times \left(1 + \frac{n \ln n}{n-1}\right)$

$$= \quad 1 + \frac{n \ln n}{n-1}$$