Probability & Statistics for EECS: Homework #01

Due on Oct 15, 2023 at 23:59

Name: **Runkang Yang**
Student ID: 2022533080

Problem 1

1. (a)

consider the ways of divide n+1 students into k groups, which is $\begin{Bmatrix} n+1 \\ k \end{Bmatrix}$ .If I'm in a group by myself, then the rest is to divide the other n students into k-1 groups, which is $\begin{Bmatrix} n \\ k-1 \end{Bmatrix}$. If I'm not in a group by myself,then I can be in any of the k group and if I was in a given specific, the rest can be seen as to divide n students into k groups. Since the given group has k possibilities, thus the result is multiple by k, which is k $\begin{Bmatrix} n \\ k \end{Bmatrix}$. So adding the two results up, we have $\begin{Bmatrix} n+1 \\ k \end{Bmatrix}$= $\begin{Bmatrix} n \\ k-1 \end{Bmatrix}$+k $\begin{Bmatrix} n \\ k \end{Bmatrix}$.

(b)

We have n+1 students and for the n students (except me), If j people are not going to be in my group, then we can divede the j students into other k groups, then the number of ways to group the j students is first choose j students from the total n students (which is $\begin{pmatrix} n \\ j \end{pmatrix}$) and divide the j students into k groups (which is $\begin{Bmatrix} j \\ k \end{Bmatrix}$). And multiply them we obtain $\begin{pmatrix} n \\ j \end{pmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix}$, which means that we divide n+1 students into k+1(other k groups and my group) groups.

Problem 2

1. We can find that the total number of norepeatword is 26+26x25+26x25x24+...+26!= $\sum_{k=1}^{26} \frac{26!}{(26-k)!}$. And the total number of norepeatword with 26 letters is 26!.

   So the probability that it uses all 26 letters is

   P=$\frac{26!}{\sum_{k=1}^{26} \frac{26!}{(26-k)!}}$
   =$\frac{26!}{26! \sum_{k=1}^{26} \frac{1}{(26-k)!}}$
   =$\frac{1}{\sum_{k=1}^{26} \frac{1}{(26-k)!}}$
   =$\frac{1}{\sum_{k=0}^{25} \frac{1}{k!}}$
   $\approx \frac{1}{e}$

   (Since $\sum_{k=0}^{\infty} \frac{1}{k!}$ is equivalent to e according to taylor's formula and here we just regard 25 as a big number to perform approximate calculations.)

Problem 3

1. (a)

   For any $1 \leq$ j $\leq$ n, j $\in \mathbb{Z}^+$, a_j has n possibilities ranging from a_1 to a_n for us to choose. So the number of the bootstrap samples is n^n.

   (b)

   Since the sum of the total number of a_j is n, let's assume that x_i denotes the number of one a_j. Then we can transfer the problem into this : find all the non-negative integer solution set for the equation $\sum_{i=1}^{n} x_i$=n. By using the method of stars and bars(also called the Bose-Einstein Counting). First, we can transfer the problem into this:find all the positive integer solution set for the equation $\sum_{i=1}^{n} x_i$=2n. And by the conclusion shown in our slides, the result is $\begin{pmatrix} 2n-1 \\ n-1 \end{pmatrix}$

   (essentially, we just inserting n-1 partitions into 2n-1 gaps)

   (c)

   From the example given above, we can find that (3,1) and (1,3) are considered to be the same, so all the possible bootstrap can be divided into three situations, and each of the possibility is $\frac{1}{4}$,$\frac{1}{2}$,$\frac{1}{2}$, and it's obvious that they're not equally likely.

   And we can find that if the bootstrap is formed by all the different element(ranging from a_1 to a_n), the sample is as likely as possible.If we want all the different elements to appear in the bootstrap, there exists n! situations and each situation has the possibility of $(\frac{1}{n})^n$, so p_1 is equal to $n!(\frac{1}{n})^n$.

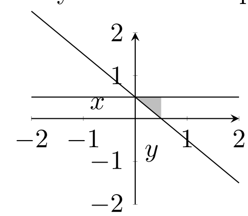
   And we can find that if the bootstrap is formed by all the same element, the sample is as unlikely as possible. If we select a specific element a_j, (the possibilityy is $\frac{1}{n}$), then the number of a_j must be n to satisfy the situation. So p_2 is equal to $(\frac{1}{n})^n$.

   So $\frac{p_1}{p_2} = \frac{n!(\frac{1}{n})^n}{(\frac{1}{n})^n} = n!$.

   It's easy to find that the number of the bootstrap consisted of the elements ranging from a_1 to a_n (like $\{a_1, a_2, a_3, ..., a_n\}$ or $\{a_n, a_{n-1}, a_{n-2}, ..., a_1\}$ and so on) is only one (are all be seen as an identical situation). However, the number of the bootstrap consisted of only one kind of element (like $\{a_1, a_1, a_1, ..., a_1\}$ or $\{a_2, a_2, a_2, ..., a_2\}$ and so on) is n, so the ratio of the probability of getting an unordered bootstrap sample whose probability is p1 to the probability of getting an unordered sample whose probability is p2 is $\frac{1 * p_1}{n * p_2} = (n-1)!$

Problem 4

1. Since the property of a triangle is that the sum of any two sides is longer than the other side, then we can transfer the problem into assuming that the length of a line(the stick) is 1, and the length of the first side is x, the length of the second side is y, the length of other side is 1-x-y, under which circumstance that the three sides can form a triangle? The sample space is that x and y and 1-x-y can be any real number in (0,1). And they should satisfy that $\begin{cases} x+y<1 \end{cases}$ So we can calculate total square of the sample space is 1x1x$\frac{1}{2}$=$\frac{1}{2}$. Besides, they should also satisfy the below requirements. $\begin{cases} x+y>1-x-y \\ x+1-x-y>y \\ y+1-x-y>x \end{cases}$ .By simplify this, we can get $\begin{cases} y>-x+\frac{1}{2} \\ 0<y<\frac{1}{2} \\ 0<x<\frac{1}{2} \end{cases}$



.We can solve this problem using a geometric approach by plotting it on a two-dimensional coordinate system. The resulting figure shows that the shaded area (above the line y=-x+$\frac{1}{2}$, below y=$\frac{1}{2}$, and to the left of x=$\frac{1}{2}$) represents the region that satisfies the condition. And the square of the shade is $\frac{1}{2}$x$\frac{1}{2}$x$\frac{1}{2}=\frac{1}{8}$.Since the total square of the sample space is $\frac{1}{2}$, we get the probability p=$\frac{\frac{1}{8}}{\frac{1}{2}}=\frac{1}{4}$.

Problem 5

1. (a)

   If $k > 365$, then according to the pigeonhole principle,we can conclude that the probability that there is at least one birthday match is 1.
   If $k \leq 365$, we can first consider the probability of no birthday match, then transfer the problem into 1-the probability of no birthday match. Since we can use $e_k(p)$ to denote the concept of the probability of having distinct birthdays on k out of all possible 365 days, then just multiply by k!,which is all possible combinations of k individuals for k days, we get the result n for the probability that there is at least one birthday match p=1-$k!e_k(p)$

   (b)

   Using the method of considering simple and extreme cases, let's say for some specific j, p_j is very close to 1, and we can just regard it as 1 to simplify our calculation, so choose any two people, they must have the same birthday and P(at least one birthday match) is maxmized to 1. And we can find that the higher the possibility of p_j, then the higher the possibility of the same birthday on p_j.

   (c)

   First,let's verify that

$$\begin{aligned} e_k(x_1, ..., x_n) = x_1 x_2 e_{k-2}(x_3, ..., x_n) \\ + (x_1 + x_2)e_{k-1}(x_3, ..., x_n) + e_k(x_3, ..., x_n), \end{aligned}$$

   If we consider both x_1 and x_2, it can be denoted as $x_1 x_2 e_{k-2}(x_3, ..., x_n)$, if we consider x_1 or x_2, it can be denoted as $(x_1 + x_2)e_{k-1}(x_3, ..., x_n)$, if we don't consider both x_1 and x_2, it can be denoted as $e_k(x_3, ..., x_n)$. So we just divide the left hand side into three parts.

   Based on the conditions provided in the question, $p_1 p_2 \leq (\frac{p_1+p_2}{2})^2 = (\frac{2r_1}{2})^2 = r_1^2 = r_1 r_2$.
   And $p_1 + p_2 = 2r_1 = r_1 + r_2$, so we can find that

$$\begin{aligned} e_k(p_1, ..., p_n) &= p_1 p_2 e_{k-2}(p_3, ..., p_n) + (p_1 + p_2)e_{k-1}(p_3, ..., p_n) + e_k(p_3, ..., p_n) \\ &\leq r_1 r_2 e_{k-2}(r_3, ..., r_n) + (r_1 + r_2)e_{k-1}(r_3, ..., r_n) + e_k(r_3, ..., r_n) \\ &= e_k(r_1, ..., r_n) \end{aligned}$$

   Then we have

$$\begin{aligned} -e_k(p_1, ..., p_n) \geq e_k(r_1, ..., r_n) \\ 1 - e_k(p_1, ..., p_n) \geq 1 - e_k(r_1, ..., r_n) \end{aligned}$$

   Which means that

$$P(\text{at least one birthday match}|p) \geq P(\text{at least one birthday match}|r)$$

   And the equal sign holds only when r and p are equal. Since $r_1 = r_2 = \frac{r_1+r_2}{2}$, we can conclude that the value of p that minimizes the probability of at least one birthday match only occurs if each p_j shares the equal part (1 divided into 365 parts, which is $\frac{1}{365}$ for all j).

Problem 6

1. Since each coupon has 108 possibilities, then the number of all the situation is 108^n . In our last term, the discrete mathematic introduce to us a counting method called the Stirling number of the second kind to handle the problem of distributing n labled objects into k unlabled boxes. And we can see the problem as this, that the satisfied situation is to divide the n identical box into 108 different groups, (though they maybe apart from each other and we just see the same type as one group) we can denote this as $S_2(n, 108)$, besides, each group has 108! possibilities, so just multiply by 108!, we get the satisfied situations, which is $108!S_2(n, 108)$.

   So the probability is

$$\begin{aligned} P &= \frac{108!S_2(n,108)}{108^n} \\ &= \frac{1}{108^n}\sum_{k=0}^{108}(-1)^k \left(\begin{array}{c} 108 \\ k \end{array} \right) (108-k)^n \end{aligned}$$

   It seems like too difficulty for us to simplify the expression, so we use the computer to calculate the result, the ploted figure is shown as below:
   And when such probability is no less than 95%, the minimum number of n is 823 by the calculation of computer.