Stat510(Section001): Homework #1

Due on Spet. 15, 2021 at 11:59pm

Instructor:Brandon Legried

Tiejin Chen **tiejin@umich.edu**

Problem 1

Assumed that set $\{A_i\}, i = 1, 2...$ which can represent any infinite but countable sequence of sets which satisfies for any $i \neq j$, A_i and A_j is disjoint. Our goal is to prove:

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

According to finite additivity, we can only get:

$$P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

Now,we create a new sequence $\{S_n\} = \sum_{i=1}^{n} P(A_i)$, according to the defintion, we can know that $\{S_n\}$ is an incresing sequence since for every i,$P(A_i) \geq 0$ and it has a upper bound beacause for any n, $S_n \leq P(\Omega) = 1$. And We have following theroerm in Calculus:

Theorem 1 (Monotone convergence theorem)**.** *if a sequence is increasing and bounded above by a upper bound, then the sequence will converge.*

this theorem means $\{S_n\}$ will converge, which implies the right-side of our equation in finite additivity situation above will converge, Hence we get:

$$\lim_{n\to\infty} P(\cup_{i=1}^{n} A_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{\infty} P(A_i)$$

So, we only need to prove the equation below to complete our proof:

$$\lim_{n\to\infty} P(\cup_{i=1}^{n} A_i) = P(\cup_{i=1}^{\infty} A_i)$$

First, We need to get a new conclusion from the Axiom of Continuity(from above). Assumed we have set $\{X_i\}$ which is a decresing sequence of sets, and $X_i \searrow B$.Now we consider a new set$\{X_i \backslash B\}$.We can easily know:

$$\cap_{i=1}^{\infty}(X_i \backslash B) = (\cap_{i=1}^{\infty} X_i)\backslash B = B \backslash B = \emptyset$$

which means that $\{X_i \backslash B\} \searrow \emptyset$. Then we can have:

$$\lim_{i\to\infty} P(X_i) = \lim_{i\to\infty} P(X_i \backslash B \cup B) = \lim_{i\to\infty}(P(X_i \backslash B) + P(B)) = \lim_{i\to\infty} P(X_i \backslash B) + P(B)$$

We can apply Axiom of Continuity (from above) to the first term of the equation above,and we can get:

$$\lim_{i\to\infty} P(X_i) = 0 + P(B) = P(B)$$

Hence we derive a new conclusion:

Conclusion 1 (Extend of Axiom of Continuity (from above))**.** *For a decreasing sequence of sets* $\{X_i\}_1^{\infty}$ *and* $\{X_i\} \searrow B$*, then* $\lim_{i\to\infty} P(X_i) = P(B) = P(\cap_{i=1}^{\infty} X_i)$

Now assumed that there is a set $\{X_i\}_1^{\infty}$ which is an incresing sequence of sets.Let $X_i \nearrow B$ means $\cup_{i=1}^{\infty} X_i = B$. We consider their complement set $\{X_i^c\}$. Since the set $\{X_i\}_1^{\infty}$ is an incresing sequence of sets, their complement set will be a decresing sequence, And from De Morgan's Law, we can know that:

$$\cap_i^{\infty} X_i^c = (\cup_i^{\infty} X_i)^c = B^c$$

Problem 1 continued on next page...

Apply the conclusion we mention above,we can get:

$$\lim_{i\to\infty} P(X_i^c) = P(B^c) = \lim_{i\to\infty}(1 - P(X_i)) = 1 - \lim_{i\to\infty} P(X_i) = 1 - P(B)$$

we get:

$$\lim_{i\to\infty} P(X_i) = P(B) = P(\cup_{i=1}^{\infty} X_i)$$

Now we have $\{X_i\}$ is an incresing sequence of sets, which means that for every $i < j$, $X_i \subseteq X_j$. so it is obvious that:

$$\cup_{i=1}^{\infty} X_i = \lim_{i\to\infty} X_i$$

Thus, We can have another conclusion:

Conclusion 2 (Axiom of Continuity (from below))**.** *For an increasing sequence of sets* $\{X_i\}_1^{\infty}$ *and* $\{X_i\} \nearrow B$, $\lim_{i\to\infty} P(X_i) = P(B) = P(\cup_{i=1}^{\infty} X_i) = P(\lim_{i\to\infty} X_i)$

Let us back to the original problem, We now create a new sequence $\{Z_n\}$ that for every n, $Z_n = \cup_{i=1}^{n} A_i$. And thus, $\{Z_n\}$ is an incresing sequence of sets without doubt.Apply our Axiom of Continuity (from below), We can get:

$$\lim_{n\to\infty} P(\cup_{i=1}^{n} A_i) = \lim_{n\to\infty} P(Z_n) = P(\lim_{n\to\infty} Z_n) = P(\lim_{n\to\infty} \cup_{i=1}^{n} A_i) = P(\cup_{i=1}^{\infty} A_i)$$

Thus we prove:

$$\lim_{n\to\infty} P(\cup_{i=1}^{n} A_i) = P(\cup_{i=1}^{\infty} A_i)$$

According to what we have proved at the beginning, we have done our proof. □

Problem 2

To prove **P** is a well-defined set function on $\mathcal{A}$, we need to prove for every A $\in \mathcal{A}$, there is only one value for **P**(A). That is to say when A is countable,**P**(A) should be only 0 and cannot be 1. which means A^c cannot be countbale according to the defintion of **P**. And we have the following fact:

Fact 1. *the countable union of countable sets will be countable.*

With the fact, we can know that if A^c is also countable when A is countable, $A^c \cup A = \Omega$ will be countable, which contradicts with the fact that Ω is unconutable. So with contradiction, A^c could not be a countable set. Hence we prove that **P** is a well-defined set function on $\mathcal{A}$.
Next we need to prove **P** is a probability measure on $\mathcal{A}$. First we need to prove $\mathbf{P}(\Omega)$ is 1. We have already known that $\mathcal{A}$ is a $\sigma - algebra$, So $\Omega \in \mathcal{A}$.And since Ω is uncountable and $\Omega^c = \emptyset$ is countable, according to the defintion of **P**, We can get $\mathbf{P}(\Omega) = 1$.
Second, we need to prove the countable additivity of **P**.Assumed that we have a countable sequence of sets$\{A_i\}_{i=1}^{\infty}$which satisfies for any $i \neq j$, A_i and A_j is disjoint. Now we have the following two situations:

1. All A $\in \{A_i\}$ is countable:
   According to fact1 we mentioned above, we know that $\cup_{i=1}^{\infty} A_i$ is also countable,so we get:

$$\sum_{i=1}^{\infty} \mathbf{P}(A_i) = \sum_{i=1}^{\infty} 0 = 0 = \mathbf{P}(\cup_{i=1}^{\infty} A_i)$$

2. there exists some A $\in \{A_i\}$ is unconutable, Here we have another fact:

   Fact 2. *the union of one countable and one uncountable set is uncountable.*

with fact2, we can know that $\cup_{i=1}^{\infty} A_i$ is an uncountable set,so we have:

$$\mathbf{P}(\cup_{i=1}^{\infty} A_i) = 1$$

Now, we need to prove that it is impossible for $\{A_i\}_1^{\infty}$ to have more than 1 uncountable set. If we have A_i and A_j for $i \neq j$ are both uncountable set. According to defintion, we know that $A_i \cap A_j = \emptyset$. Using this equation and De Morgan's Law, we can get:

$$\Omega = (\emptyset)^c = (A_i \cap A_j)^c = A_i^c \cup A_j^c$$

Since A_i and A_j are both uncountable set, their complements must be countable set. Hence A_i^c and A_j^c are both countable set. With fact1, we know that $A_i^c \cup A_j^c$ is a countable set which means Ω is a countable set. It contradicts with the fact that Ω is uncountable. Thus, We prove that it is impossible for our $\{A_i\}_{i=1}^{\infty}$ to have more than 1 uncountable set. Let A_n be this only uncountable set, we have:

$$\sum_{i=1}^{\infty} \mathbf{P}(A_i) = \mathbf{P}(A_n) + \sum_{i \neq n} \mathbf{P}(A_i) = 1 + 0 = 1 = \mathbf{P}(\cup_{i=1}^{\infty} A_i)$$

Now we have proved for both situation, countable additivity established. Thus we prove that $\mathbf{P}$ is a probability measure on $\mathcal{A}$. □

Problem 3

Part One

In this part, we will calculate probability that a single(now sepcific) box receives all 8 balls. Here begin our analysis:

1. To the first ball, the ball can be thrown to any box of 20 boxes. So we do not need to calculate the probability for the first ball
2. for the rest balls, every ball need to be thrown to the exact box which the first ball is thrown in. The probability for every throw is 1/20 since they are all randomly throw.

Hence, we get the result:

$$P(8balls\ in\ one\ single\ box) = \frac{1}{20^7}$$

Part Two

We will calculate probability that no box receives more than one ball. Here begin our analysis:

1. For the first ball, the ball can be thrown to any box of 20 boxes. So we do not need to calculate the probability for the first ball
2. For the second ball, it cannot be thrown to the box which the first ball is thrown in, which means it should be thrown to any of left 19 boxes, so the probability here is $\frac{19}{20}$
3. For the n-th($n \leq 20$) ball, it cannot be thrown to the former n-1 box which the first n-1 ball is thrown in, which means it should be thrown to any of left 21-n boxes, so the probability here is $\frac{21-n}{20}$.

Hence, we get the result:

$$P(no\ box\ more\ than\ one\ ball) = \frac{19!}{12! \times 20^7}$$

Problem 3 continued on next page. . .

Or we can have another view. The counted number of all possible outcomes is 20^8 by ordered sampling with replacement. And let no box receive more than one ball means we need to pick up 8 subsamples from 20 boxes without replacement, And it is order since one same box has one different ball is a different desired outcome. So, we can get the count number of the desired outcome is ${}_{20}P_8$. Hence we have:

$$P(no\ box\ more\ than\ one\ ball) = \frac{20!}{12!} \div 20^8 = \frac{19!}{12! \times 20^7}$$

We get the same answer from different methods.

Problem 4

It is obvious to have:

$$P(defective\ bulbs\ in\ one\ person) = P(defective\ bulbs\ in\ first\ person) + P(defective\ bulbs\ in\ second\ person)$$

Let us calculate this two part one by one.
The probability of all the defective bulbs in the first person means that the first person should take 4 defective bulbs and other 6 bulbs. The counted number of all possible outcomes for the first person is ${}_{24}C_{10}$ by unorder sampling without replacement. The counted number of all desired outcomes is ${}_{20}C_6$ because 4 defective bulbs is fixed, and the different situations only occur in left 6 bulbs and 20 total bulbs by unorder sampling without replacement. Hence we get:

$$P(defective\ bulbs\ in\ first\ person) = \frac{{}_{20}C_6}{{}_{24}C_{10}} = \frac{10!14!20!}{6!14!24!}$$

The probability of all the defective bulbs in the second person means that the first person takes 10 good bulbs.The counted number of all possible outcomes for the first person is ${}_{24}C_{10}$ as same as the situation above. And the counted number of desired outcomes is ${}_{20}C_{10}$ because the first person needs to take 10 good bulbs in 20 total good bulbs by unorder sampling without replacement. Hence we get:

$$P(defective\ bulbs\ in\ second\ person) = \frac{{}_{20}C_{10}}{{}_{24}C_{10}} = \frac{14!20!}{10!24!}$$

Hence,we can get:

$$P(defective\ bulbs\ in\ one\ person) = \frac{10!14!20!}{6!14!24!} + \frac{14!20!}{10!24!} \simeq 0.1139657$$

Problem 5

The problem can been seen as the first person pick up 13 cards, the second person pick up 13 cards in left 39cards and so on. Let us begin our analysis with the first to pick up 13 cards.The counted number of the first person possible outcomes is ${}_{52}C_{13}$ by unorder sampling without replacement. And when any outcome for the first person is settled down, second person begins to pick up. The counted number of second person possible outcomes is ${}_{39}C_{13}$ under this outcome. In other words, one possible outcome from first person leads to ${}_{39}C_{13}$ possible outcomes from second person. Hence the counted number of the first and second person possible outcomes is $({}_{52}C_{13}) \times ({}_{39}C_{13})$. With the same logic, we can get the counted number of all situation is $({}_{52}C_{13}) \times ({}_{39}C_{13}) \times ({}_{26}C_{13}) \times ({}_{13}C_{13}) = ({}_{52}C_{13}) \times ({}_{39}C_{13}) \times ({}_{26}C_{13})$.
Now, we want to get the counted number of all desired outcomes. For the first person, he or she must take exact 1 ace. Hence the different situations only occur in left 12 cards, and the first person needs to pick up 12 cards in all non-ace cards(52-4 cards). And because of the different suit of cards, the suit of ace can influence the situation. That is to say, the first person needs to pick up 12 cards in 48 non-ace cards and pick up 1 card

Problem 5 continued on next page...

in 4 ace card, So the counted number of the first person desired outcomes is $({}_{48}C_{12}) \times ({}_{4}C_{1}) = 4 \times ({}_{48}C_{12})$. And with the same logic we mentioned above in this problem, we have the counted number of first and second person desired outcomes is $4 \times 3 \times ({}_{48}C_{12}) \times ({}_{36}C_{12})$.Hence, we get the counted number of all desired number is$4! \times ({}_{48}C_{12}) \times ({}_{36}C_{12}) \times ({}_{24}C_{12})$:

$$P(everyone\ get\ one\ ace) = \frac{4! \times ({}_{48}C_{12}) \times ({}_{36}C_{12}) \times ({}_{24}C_{12})}{({}_{52}C_{13}) \times ({}_{39}C_{13}) \times ({}_{26}C_{13})}$$

Problem 6

Let us consider a new situation first. For any $i, j \leq n, and i \neq j$, we use $P(i, j)$ to present the probability of i-th person and j-th person meet in one tournament.And we let A be 1 and B be 2. So our goal is to calculate $P(1, 2)$.
Now we consider all the situation of P(i,j). First, we consider how many this kind of P we have? The answer is ${}_{n}C_{2}$ by unorder sampling without replacement. Thus, we have ${}_{n}C_{2}\ P(i, j)\ for\ i \neq j$.
Now we consider the relationship between all the P. Since there is no any difference between the first person, second person, and third person. $P(1, 2)$ should equal to $P(1, 3)$ and $P(2, 3)$. With the same logic, since there is no difference between all the people in the tournament. All P should be equal and we use p to present all the $P(i, j)$ Hence, our goal now is to calculate this p.
Next we need to calculate the sum of all the P now. Since in one tournament, everyone except the champion will lose and only lose one game. Hence there are total $n - 1$ games have a loser. As we all know, a game must have a loser. Hence, the total number of game in tournament is $n - 1$. Also, it is obvious that two person can only meet one time in one tournament. Hence, the every game in tournament is every different pairwise person. That is to say, in one tournament, for $i \neq j$ and i meet j, this event happen and only happen n-1 times in one tournament.Thus, We have:

$$\sum_{{}_{n}C_{2}} p = \sum_{i \neq j} P(i, j) = n - 1$$

We can get the result:

$$p = P(1, 2) = \frac{n - 1}{{}_{n}C_{2}} = \frac{2}{n}$$

Problem 7

Part(a)
Let us use induction to prove the result. And we start from n = 2:
If there is an unrooted binary tree with 2 leaf vertices. Then these vertices can just connect to each other. Then their degrees are both 1, so they are leaf vertices. And such a tree is an unrooted binary tree and they have one edge and two vertices. Hence the$|E| = 1 = 2 * 2 - 3, |V| = 2 = 2 * 2 - 2$ satisfy the formula.
Now we assumed that when in the situation of $n - 1$,the formula established, we then prove the situation of n the formula is also estabilished.
We will consider the situation that a new leaf vertice want to join an unrooted binary tree with n-1 leaf vertices. We know that we must connect the new leaf(use nl to present) to an old leaf(use ol to present), because if the nl connected to the an interior vertice, the degree of this interior vertice will be greater than 3. So this tree is not an unrooted binary tree anymore. Hence it must be connected with an ol. However, after a nl is added to an ol, the degree of this ol will be 2 instead of 1 or 3. In order to let this tree still be binary tree, we need to add a new leaf besides our nl to ol. Then the ol will be a new interior vertice to make its degree become 3. And the tree is still a unrooted binary tree. Now the number of this new tree's leaf vertices will be $n - 1 - 1 + 2 = n$ that statisfy our siutation. And, we add two edges and two vertices

into this tree obviously. Thus, $|E| = 2(n-1) - 3 + 2 = 2n-3, |V| = 2(n-1) - 2 + 2 = 2n-2$, which statisfy our formula. Hence we have proved the formula established. $\square$

Part(b)

First,we consider the situation of $n = 0$ and $n = 1$. Now we prove an unrooted binary tree cannot have 0 or 1 leaf vertice. If we want to create a tree without leaf vertice, then all the vertices need to be interior point. Then every pair points which are connected, the rest of point they connect should not have overlap. Otherwise, it will have more than one path between the overlaping point and anyone of the original pair points. Then we can conclude that $|V|$ need to be ∞ since this process is endless. For situation $n = 1$,we have the one interior vertice(use v_1 to present) connected to the only leaf vertice. And we use v_2 and v_3 to present the interior vertices connected to v_1. Then we can find that the vertices connected to v_2 and the vertices connected to v_3 should not overlap except v_1. With the same logic we use when $n = 0$ because the rest of vertices are all interior vertices only connected to other interior vertices, we can know that $|V|$ need to be ∞. Thus, we prove that it is impossible for an unrooted binary tree have 0 or 1 leaf vertice.

Second we consider the situation of $n = 2$, from the analysis on Part(a), we can know that the number of phylogenetic tree is 1 since they only have one connecting way.

Third we consider the situation of $n = 3$. There is only one structure of this situation, that is one interior connected to the 3 leaf vertices. Hence we can get when $n = 3$, the number of phylogenetic tree is 1.

Now we consider the situation that when $n > 3$. We can know from Part(a) that we will have $|E| = 2n - 3$ and $|V| = 2n-2$. Hence we can conclude that the number of all vertices are $2n-2$ and the number of edges between vertices are $2n - 3$.

Assumed that n-leaf tree have n_x different phylogenetic trees. Then, with the situation of $n+1$ leaf. we can find that for every distinct n-leaf phylogenetic tree, the new leaf can add to any edge. That is to say if there is an edge between v_1 and v_2, We break the edge and introduce a new interior vertice v_3. We connect v_1, v_3 and v_3, v_2. And we connect v_3 with new leaf vertice. We traverse all the edge. And each new n+1 leaf tree is a new phylogenetic tree. Hence we get:

$$n_{x+1} = (2n-3)n_x$$

Now we also have $n_2 = n_3 = 1$, statisfy the formula above. So we have:

$$n_x = \begin{cases} 0 & n = 0 \ or \ n = 1 \\ 1 & n = 2 \\ \prod_{i=3}^{n}(2i-5) & n \geq 3 \end{cases}$$

Problem 8

Write the probability the walk after step n is at 0 means that we need to walk exact $\frac{n}{2}$ step to right and exact $\frac{n}{2}$ to left. It is obvious that $\frac{n}{2}$ must be an interger. Hence when n is odd, it is impossible for us to stay at 0 when We walk n step.

Now we consider the situation when n is even. For every step, I have 2 possible outcome, which is go left or go right. Hence the counted number of all possible outcomes is 2^n by ordered sampling with replacement. The number of sequences with exactly $\frac{n}{2}$ going left(which implies $\frac{n}{2}$ going right) is in correspondence with the subsets of size $\frac{n}{2}$ from the n steps.By unordered sampling without replacement, the number of desired outcomes is ${}_nC_{\frac{n}{2}}$. Hence when n is even, the probability is $\frac{{}_nC_{\frac{n}{2}}}{2^n}$. Thus, we get the result:

$$P(stay\ 0\ after\ n\ steps) = \begin{cases} 0 & n\ is\ odd \\ \frac{{}_nC_{\frac{n}{2}}}{2^n} & n\ is\ even \end{cases}$$