170A Midterm 1 Solutions[1]

1. Question 1

Label the following statements as TRUE or FALSE. If the statement is true, explain your reasoning. If the statement is false, provide a counterexample and explain your reasoning.

In the following statements, let A, B, C be subsets of some universe Ω. Let $\mathbf{P}$ denote a probability law on Ω. Also, let $A_1, A_2, A_3, \ldots$ be subsets of Ω.

(a) $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$.

TRUE; see Proposition 2.33 in the notes.

(b) $\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(A^c \cap B) + \mathbf{P}(A^c \cap B^c \cap C^c)$.

FALSE. If $A = B = \emptyset$ and if $C = \Omega$, then this statement says $1 = 0$.

(c) $\mathbf{P}\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mathbf{P}(A_k)$.

FALSE. This statement only holds for disjoint sets. For example if $A_1 = A_2 = \Omega$ and if $A_3 = A_4 = \cdots = \emptyset$, then this statement says $1 = 2$.

(d) If $\mathbf{P}(A) > 0$ and if $\mathbf{P}(A^c) > 0$, then $\mathbf{P}(B) = \mathbf{P}(B|A)\mathbf{P}(A) + \mathbf{P}(B|A^c)\mathbf{P}(A^c)$.

TRUE. This follows from the total probability theorem.

(e) If $\mathbf{P}(A_1 \cap A_2 \cap A_3) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3)$, then the sets A_1, A_2, A_3 are independent.

FALSE. There is not enough information to determine whether or not the sets are independent. For example, if $\Omega = \{1, 2, 3\}$, if $\mathbf{P}$ is the uniform probability distribution on Ω, and if $A_1 = \{1, 2\}$, $A_2 = \{2, 3\}$, $A_3 = \emptyset$, then the statement holds since both sides are zero, but $\mathbf{P}(A_1 \cap A_2) = \mathbf{P}(\{2\}) = 1/3 \neq (2/3)^2 = \mathbf{P}(A_1)\mathbf{P}(A_2)$. So, these sets are not independent.

2. Question 2

Let $\mathbf{R}$ denote the set of real numbers. Prove that

$$\mathbf{R} = \bigcup_{j=1}^{\infty} [-j, j].$$

Solution. Let $x \in \cup_{j=1}^{\infty}[-j, j]$. Then there exists some positive integer J such that $x \in [-J, J]$. Since $[-J, J] \subseteq \mathbf{R}$, we know $x \in \mathbf{R}$. That is, we have shown that $\cup_{j=1}^{\infty}[-j, j] \subseteq \mathbf{R}$. It remains to show that $\mathbf{R} \subseteq \cup_{j=1}^{\infty}[-j, j]$. So, let $x \in \mathbf{R}$. We need to show that $x \in \cup_{j=1}^{\infty}[-j, j]$. That is, using the definition of the union, we need to find some positive integer J such that $x \in [-J, J]$. Let J be any integer larger than $|x|$. Then $x \in [-J, J]$, since $|x| \leq J$. That is, $x \in \cup_{j=1}^{\infty}[-j, j]$ so that $\mathbf{R} \subseteq \cup_{j=1}^{\infty}[-j, j]$.

Since $\cup_{j=1}^{\infty}[-j, j] \subseteq \mathbf{R}$ and $\mathbf{R} \subseteq \cup_{j=1}^{\infty}[-j, j]$, we conclude that $\mathbf{R} = \cup_{j=1}^{\infty}[-j, j]$.

3. Question 3

Two people take turns throwing darts at a board. Person A goes first, and each of her throws has a probability of $1/4$ of hitting the bullseye. Person B goes next, and each of her throws has a probability of $1/3$ of hitting the bullseye. Then Person A goes, and so on. With what probability will Person A hit the bullseye before Person B does?

Solution. Person A hits the bullseye on her first try with probability $1/4$. If both A and B miss their first throw, then Person A hits the bullseye on her second try with probability

$(1-1/4)(1-1/3)(1/4)$. If both A and B miss their first two throws, then Person A hits the bullseye on her third try with probability $(1-1/4)^2(1-1/3)^2(1/4)$. For any positive integer k, let C_k be the event that both A and B miss their first k throws, and Person A hits the bullseye on the $(k+1)^{st}$ try. Then $\mathbf{P}(C_k) = (1-1/4)^k(1-1/3)^k(1/4) = (3/4)^k(2/3)^k(1/4) = (1/2)^k(1/4)$. Let C be the event that person A hits the bullseye before person B. Then $C = \cup_{k\geq 0} C_k$, and $C_k \cap C_{k'} = \emptyset$ if $k \neq k'$. So, from the axioms for a probability law,

$$\mathbf{P}(C) = \mathbf{P}(\cup_{k\geq 0} C_k) = \sum_{k=0}^{\infty} \mathbf{P}(C_k) = (1/4)\sum_{k=0}^{\infty}(1/2)^k = (1/4)(2) = 1/2.$$

4. Question 4

Suppose you have 100 cubes in a container, and they are labelled with the integers 1 through 100, so that each cube has a distinct integer label.

Suppose you remove one cube uniformly at random from the container, and then put the cube back into the container. Then, you take one cube out of the container uniformly at random, and you then take another cube out of the container uniformly at random, you swap the integer labels of these cubes, and you then put the cubes back into the container. Then, once again, you take one cube out of the container uniformly at random, and you then take another cube out of the container uniformly at random, you swap the integer labels of these cubes, and you then put the cubes back into the container

Finally, remove one cube from the container uniformly at random. What is the probability that this cube has the integer label 57?

Solution. The probability is 1/100. The swapping of labels doesn't change the fact that there are 100 cubes with 100 different integer labels. When we finally take one cube from the container, we select the cube uniformly at random, so that each one is equally likely to be removed from the container. So, the probability of finding any fixed label (e.g. 57) is 1/100, since there are 100 cubes.

5. Question 5

A community has $m > 0$ families. Each family has at least one child. The largest family has $k > 0$ children. For each $i \in \{1, \ldots, k\}$, there are n_i families with i children. So, $n_1 + \cdots + n_k = m$. Choose a child randomly in the following two ways.

Method 1. First, choose one of the families uniformly at random among all of the families. Then, in the chosen family, choose one of the children uniformly at random.

Method 2. Among all of the $n_1 + 2n_2 + 3n_3 + \cdots + kn_k$ children, choose one uniformly at random.

What is the probability that the chosen child is the first-born child in their family, if you use Method 1?

What is the probability that the chosen child is the first-born child in their family, if you use Method 2?

Solution. In method 1, let A be the event that the chosen child is first born, and for each $i \in \{1, \ldots, k\}$, let B_i be the event that a family with i children is chosen. Then $B_i \cap B_j = \emptyset$

for $i \neq j$, $i, j \in \{1, \ldots, m\}$, and $\cup_{i=1}^{m} B_i = \Omega$. So, the total probability theorem says

$$\mathbf{P}(A) = \sum_{i=1}^{k} \mathbf{P}(A|B_i)\mathbf{P}(B_i).$$

In method 1, we have $\mathbf{P}(B_i) = n_i/m$ for every $i \in \{1, \ldots, m\}$, since each of the m families is equally likely to be chosen, and there are n_i families with i children. Also, given that B_i occurs, we have $\mathbf{P}(A|B_i) = 1/i$, since if the chosen family has i children, the probability of choosing the first born child is $1/i$. In conclusion,

$$\mathbf{P}(A) = \sum_{i=1}^{k} \mathbf{P}(A|B_i)\mathbf{P}(B_i) = \sum_{i=1}^{k} \frac{n_i}{im} = \frac{1}{m} \sum_{i=1}^{k} \frac{n_i}{i}.$$

In method 2, the total population of children is sampled uniformly at random. So, the probability is the number of first-born children, divided by the total number of children. That is, the probability is the number of families, divided by the total number of children. That is, the probability the chosen child is first-born is

$$\frac{m}{n_1 + 2n_2 + 3n_3 + \cdots + kn_k} = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} in_i}.$$

Another solution. In method 1, the probability of choosing any family is equally likely. After one family is selected, the first-born child is chosen with probability equal to the inverse of the number of children in that family. So, summing over all m families, the probability the chosen child is first-born is

$$\frac{1}{m} \sum_{\text{families } F} \frac{1}{\#\,\text{of children in family}\, F} = \frac{1}{m} \sum_{i=1}^{k} \frac{\#\,\text{of families of size}\, i}{i} = \frac{1}{m} \sum_{i=1}^{k} \frac{n_i}{i}.$$

In method 2, the total population of children is sampled uniformly at random. So, the probability is the number of first-born children, divided by the total number of children. That is, the probability is the number of families, divided by the total number of children. That is, the probability the chosen child is first-born is

$$\frac{m}{n_1 + 2n_2 + 3n_3 + \cdots + kn_k} = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k} in_i}.$$