Indian Institute of Technology Guwahati
Statistical Foundation of Data Science (MA589)
Problem Set 01

1. Let S be a sample space of a random experiment. Let A, B, and C be three events. What is the event that only A occurs? What is the event that at least two of A, B, C occur? What is the event that both A, B, but not C occur? What is the event of at most one of the A, B, C occurs?

2. Let $S = \{0, 1, 2, \ldots\}$ be a sample space. Let $\mathcal{F} = \mathcal{P}(S)$. In each of the following cases, verify if $P(\cdot)$ is a probability.

   (a) $P(A) = \sum_{x \in A} \frac{e^{-\lambda}\lambda^x}{x!}$, $A \in \mathcal{F}$, $\lambda > 0$.

   (b) $P(A) = \sum_{x \in A} p(1-p)^x$, $A \in \mathcal{F}$, $0 < p < 1$.

   (c) $P(A) = 0$, if A has a finite number of elements, and $P(A) = 1$, if A has infinite number of elements, $A \in \mathcal{F}$.

3. Let $A_1, A_2, \ldots, A_n$ be $n > 1$ events. Then prove that

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P\left(A_i\right).$$

4. (Inclusion-exclusion principle) Let $A_1, A_2, \ldots, A_n$ be $n > 1$ events. Then prove that

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{\substack{i_1=1 \\ i_1<i_2}}^{n} \sum_{i_2=1}^{n} P\left(A_{i_1} \cap A_{i_2}\right) + \sum_{\substack{i_1=1 \\ i_1<i_2<i_3}}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} P\left(A_{i_1} \cap A_{i_2} \cap A_{i_3}\right) - \ldots$$
$$+ (-1)^{n+1} P\left(\bigcap_{i=1}^{n} A_i\right).$$

5. (Bonferroni's Inequality) Given $n\,(> 1)$ events $A_1, A_2, \ldots, A_n$, prove that

$$\sum_{i=1}^{n} P(A_i) - \sum_{\substack{i=1 \\ i<j}}^{n} \sum_{j=1}^{n} P(A_i \cap A_j) \leq P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

   [Hint: To prove the LHS, use induction.]

6. Let A, B, C, and D be four events such that $P(A) = 0.6$, $P(B) = 0.5$, $P(C) = 0.4$, $P(A \cap B) = 0.3$, $P(A \cap C) = 0.2$, $P(B \cap C) = 0.2$, $P(A \cap B \cap C) = 0.1$, $P(B \cap D) = P(C \cap D) = 0$, $P(A \cap D) = 0.1$, and $P(D) = 0.2$. Find

   (a) $P(A \cup B \cup C)$ and $P(A^c \cap B^c \cap C^c)$. (Ans: 0.9 and 0.1)

(b) $P((A \cup B) \cap C)$ and $P(A \cup (B \cap C))$. (Ans: 0.3 and 0.7)

(c) $P((A^c \cup B^c) \cap C^c)$ and $P((A^c \cap B^c) \cup C^c)$. (Ans: 0.4 and 0.7)

(d) $P(D \cap B \cap C)$ and $P(A \cap C \cap D)$. (Ans: 0 and 0)

(e) $P(A \cup B \cup D)$ and $P(A \cup B \cup C \cup D)$. (Ans: 0.9 and 1.0)

(f) $P((A \cap B) \cup (C \cap D))$. (Ans: 0.3)

7. Suppose that $n\,(\geq 3)$ persons $P_1, \ldots, P_n$ are made to stand in a row at random. Find the probability that there are exactly r persons between P_1 and P_2; here $r \in \{1, 2, \ldots, n-2\}$. (Ans: $2(n-r-1)/(n(n-1))$.)

8. Three numbers are chosen at random and without replacement from the set $\{1, 2, \ldots, 50\}$. Find the probability that the chosen numbers are in arithmetic progression.

9. A class consisting of four graduate and twelve undergraduate students is randomly divided into four groups of four. What is the probability that each group includes a graduate student? [Ans: $(2 \times 3 \times 4^3)/(15 \times 14 \times 13)$.]

10. Suppose that we have $n\,(\geq 2)$ letters and corresponding n addressed envelopes. If these letters are inserted at random in n envelopes, find the probability that no letter is inserted into the correct envelop. (Ans: $\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \ldots + (-1)^n \frac{1}{n!}$.)

11. Find the probability that among three random digits there appear exactly two different digits. (Ans: 0.27)

12. Let r indistinguishable balls are placed in n cells numbered 1 through n. Two distributions are said to be distinguishable only if the corresponding n-tuples $(r_1, r_2, \ldots, r_n)$ are not identical, where r_i stands for the number of balls in the ith cell.

   (a) Show that the number of distinguishable distributions is $\binom{n+r-1}{r}$.

   (b) For $r \geq n$, show that the number of distinguishable distributions in which no cell remain empty is $\binom{r-1}{n-1}$.

   (c) Find the probability that no cell remain empty.

13. A man is given n keys of which only one fits his door. He tries them successively. This procedure may require $1, 2, \ldots, n$ trails. Show that each of these n outcomes has probability $1/n$.

14. Consider an experiment involving two successive rolls of a 4-sided die in which all 16 possible outcomes are equally likely and have probability 1/16.

   (a) Are the events $A = \{\text{1st roll results in 1}\}$ and $B = \{\text{2nd roll results in 2}\}$ independent?

   (b) Are the events $A = \{\text{1st roll results in 1}\}$ and $B = \{\text{sum of the two rolls is a 5}\}$ independent?

   (c) Are the events $A = \{\text{maximum of the two rolls is 2}\}$ and $B = \{\text{minimum of the two rolls is 2}\}$ independent?

15. Let $S = (0, 1)$ and $P(I) =$ length of I, where I is an interval in S. Let $A = (0, 1/2)$, $B = (1/4, 1)$ and $C = (1/4, 11/12)$. Show that $P(A \cap B \cap C) = P(A)P(B)P(C)$, but $P(A \cap B) \neq P(A)P(B)$. (Note: It implies that $P(A \cap B \cap C) = P(A)P(B)P(C)$ is not sufficient for mutual independence of A, B, and C.)

16. Consider two independent fair coin tosses, in which all four possible outcomes are equally likely. Let $H_1 = \{\text{1st toss is a head}\}$, $H_2 = \{\text{2nd toss is a head}\}$, and $D = \{\text{the two tosses have different results}\}$. Find $P(H_1)$, $P(H_2)$, $P(H_1 \cap H_2)$, $P(H_1|D)$, $P(H_2|D)$, and $P(H_1 \cap H_2|D)$. (Ans: $P(H_1) = 0.5$, $P(H_2)0.5$, $P(H_1 \cap H_2) = 0.25$, $P(H_1|D) = 0.5$, $P(H_2|D) = 0.5$, and $P(H_1 \cap H_2|D) = 0$.) (Note: Independent does not imply conditionally independent.)

17. There are two coins, a blue and a red one. We choose one of the two at random, each being chosen with probability 1/2, and proceed with two independent tosses. The coins are biased. With the blue coin, the probability of heads in any given toss is 0.99, whereas for the red coin it is 0.01. Let D be the event that the blue coin was selected. Let H_i, $i = 1, 2$, be the event that the ith toss resulted in head. Find $P(H_1)$, $P(H_2)$, $P(H_1 \cap H_2)$, $P(H_1|D)$, $P(H_2|D)$, and $P(H_1 \cap H_2|D)$. (Ans: $P(H_1) = 0.5$, $P(H_2) = 0.5$, $P(H_1 \cap H_2) = 0.4901$, $P(H_1|D) = 0.99$, $P(H_2|D) = 0.99$, and $P(H_1 \cap H_2|D) = 0.9801$.) (Note: Conditional independence does not imply independence.)

18. Let A, B, and C be three events such that $P(B \cap C) > 0$. Prove or disprove each of the following: (a) $P(A \cap B|C) = P(A|B \cap C)P(B|C)$; (b) $P(A \cap B|C) = P(A|C)P(B|C)$ if A and B are independent events.

19. Let A, B, and C be three events such that A and B are negatively (positively) associated and B and C are negatively (positively) associated. Can we conclude that, in general, A and C are negatively (positively) associated?

20. Let A and B be two events. Show that if A and B are positively (negatively) associated then A and B^c are negatively (positively) associated.

21. (Geometric Probability) A point (X, Y) is randomly chosen on the unit square

$$S = \{(x, y) : 0 \le x \le 1, 0 \le y \le 1\}$$

i.e., for any region $R \subseteq S$ for which the area is defined, the probability that (X, Y) lies on R is $\frac{\text{Area of } R}{\text{Area of } S}$. Find the probability that distance from (X, Y) to the nearest side does not exceed $\frac{1}{5}$ units. (Ans: 16/25)

22. Two persons agree to meet a place at a given time. Each will arrive at the meeting place with a random delay between 0 to 1 hour independent of each other. The first to arrive will wait for 15 minutes and will leave if the other has not yet arrived. What is the probability that they meet? (Hint: Try to use geometric probability.) (Ans: 7/16.)

23. An individual uses the following gambling system. He bets Re. 1. If he wins, he quits. If he loses, he makes the same bet a second time only this time he bets Rs. 2, and then regardless the result of the second match he quits the game. Assuming that he has a probability 0.5 to win each bet, find the probability that he goes home a winner. (Ans: 3/4.)

24. A student is taking a probability course and at the end of each week, she can be either up-to-date or she may have fallen behind. If she is up-to-date in a given week, the probability that she will be up-to-date in the next week is 0.8. If she is behind in a week, the probability that she will be up-to-date in the next week is 0.4. She is up-to-date when she starts the class. Find the probability that she is up-to-date after three weeks. (Ans: 0.688.)

25. (The Monty Hall problem) Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others are goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, which has a goat. He then asks to you, "Do you want to pick the other closed door?" What should be your answer? (Ans: Yes, as the probability of wining the car is $\dfrac{2}{3}$ if I pick the other closed door.)

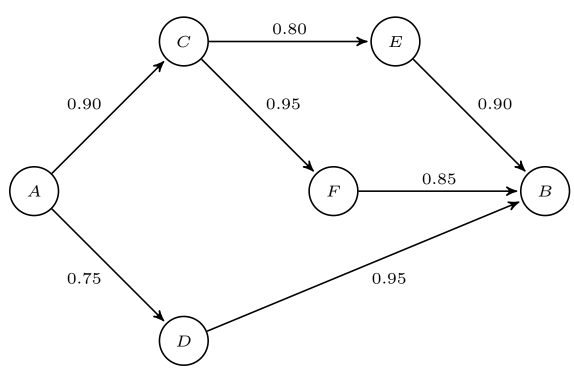
26. Consider four coding machines M_1, M_2, M_3, and M_4 producing binary codes 0 and 1. The machine M_1 produces codes 0 and 1 with respective probabilities $\frac{1}{4}$ and $\frac{3}{4}$. The code produced by machine M_k is fed into machine M_{k+1}, $(k = 1, 2, 3)$, which may either leave the received code unchanged or may change it. Suppose that each of the machines M_2, M_3, and M_4 change the code with probability $\frac{3}{4}$. Given that the machine M_4 has produced code 1, find the conditional probability that the machine M_1 produced code 0. (Ans: 3/10.)

27. A student appears in the examinations of four subjects Biology, Chemistry, Physics and Mathematics. Suppose that probabilities of the student clearing examinations in these subjects are $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{4}$, and $\frac{1}{5}$ respectively. Assuming that the performances of the student in four subjects are independent, find the probability that the student will clear examination(s) of (a) all the subjects; (b) no subject; (c) exactly one subject; (d) exactly two subjects; (e) at least one subject.

28. If an aircraft is present in a certain area, a radar detects it and generates an alarm signal with probability 0.99. If an aircraft is not present, the radar generates a (false) alarm, with probability 0.10. We assume that an aircraft is present with probability 0.05. What is the probability of no aircraft presence and a false alarm? What is the probability of aircraft presence and no detection? If the radar generates a alarm, what is the probability of the presence of an aircraft?

29. (The False-Positive Puzzle) A test for a certain rare disease is assumed to be correct 95% of the time. A random person drawn from a certain population has probability 0.001 of having the disease. Given that the person just tested positive, what is the probability of having the disease?

30. A computer networks connects two nodes A and B through intermediate nodes C, D, E, and F as shown in the figure. For every pair of directly connected nodes, say i and j, there is a given probability p_{ij} that the link form i to j is up. Assume that the link failure are independent each other. What is the probability that there is a path from A to B?



(Ans: 0.957.)