MA2040: Probability, Statistics and Stochastic Processes
Problem Set-II

Sivaram Ambikasaran

February 15, 2020

1. Assume that Carlsen meets Ding Liren in the 2020 chess championship. The championship match consists of a sequence of games and each game has three outcomes (i) Ding Liren winning, (ii) Carlsen winning, (iii) A draw. The first player to win a game wins the match. For instance, we could have a sequence of 3 draws followed by a Carlsen victory in the 4^{th} game, which would mean that Carlsen wins the Championship. The probability of a single game ending in

   (a) Carlsen's favour is 0.4

   (b) Ding Liren's favour is 0.2

   (c) draw is 0.4

   i What is the probability of Ding Liren winning the championship?
   Solution: Probability of Ding Liren winning the match on the k^{th} game is given by $0.4^{k-1} \times 0.2$, i.e., the first $k-1$ games have to be drawn and the k^{th} game has to be won by Ding Liren. Hence, the probability of Ding Liren winning the championship is

   $$\sum_{k=1}^{\infty} 0.4^{k-1} \times 0.2 = \frac{0.2}{1-0.4} = 1/3$$

   ii What is the probability of Carlsen winning the championship?
   Solution: Probability of Carlsen winning the match on the k^{th} game is given by $0.4^{k-1} \times 0.2$, i.e., the first $k-1$ games have to be drawn and the k^{th} game has to be won by Carlsen. Hence, the probability of Carlsen winning the championship is

   $$\sum_{k=1}^{\infty} 0.4^{k-1} \times 0.4 = \frac{0.4}{1-0.4} = 2/3$$

   This can also be obtained as $1 - 1/3 = 2/3$, since the probability of Ding Liren or Carlsen winning the Championship is 1.

   iii What is the Probability Mass Function for the number of games played in the championship?
   Solution: For the match to last k games, the first $k-1$ have to be drawn and the k^{th} game has to be won either by Ding Liren or Carlsen. Hence, $P(k) = 0.4^{k-1} \times (0.4 + 0.2) = 0.6 \times 0.4^{k-1}$.

2. Consider rolling a pair of fair dice. Let X denote the difference between the numbers that show up on the dice, i.e., $X = |D_1 - D_2|$, where D_i is the number that shows up on the i^{th} dice.

   - What are the possible values for X?
     Solution: X can take values from 0 to 5.

- What is the probability mass function for X?
  Solution: We have

$$P(0)=\frac{6}{36},\ P(1)=\frac{10}{36},\ P(2)=\frac{8}{36},\ P(3)=\frac{6}{36},\ P(4)=\frac{4}{36},\ P(5)=\frac{2}{36}$$

- Find the expected value and standard deviation of X.
  Solution:

$$\mathbb{E}\left(X\right)=\frac{0\times 6+1\times 10+2\times 8+3\times 6+4\times 4+5\times 2}{36}=\frac{70}{36}=1.9\bar{4}$$

3. A fair die is rolled repeatedly till an odd prime appears. What is the probability that the number of rolls exceed 5?
   Solution: Probability of an odd prime occuring in a single roll is $1/3$ (since only 3 and 5 have to occur). For the number of rolls to exceed 5, we need the first 5 rolls to be not an odd prime. Hence, the probability of this event is $\left(\dfrac{2}{3}\right)^5$.

4. Let X be a discrete random variable with mean μ and variance σ^2. Prove that

$$\mathbb{E}\left[\left(X-a\right)^2\right]=\sigma^2+\left(a-\mu\right)^2$$

   Hence, prove that the mean (or expected value) minimizes $\mathbb{E}\left[\left(X-a\right)^2\right]$.
   Solution: We have

$$\mathbb{E}\left[\left(X-a\right)^2\right]=\mathbb{E}\left[\left(X-\mu+\mu-a\right)^2\right]=\mathbb{E}\left[\left(X-\mu\right)^2\right]+2\mathbb{E}\left[\left(X-\mu\right)\left(\mu-a\right)\right]+\mathbb{E}\left[\left(\mu-a\right)^2\right]$$

   since the expectation is a linear operator. This immediately gives us that

$$\mathbb{E}\left[\left(X-a\right)^2\right]=\sigma^2+\left(a-\mu\right)^2$$

   since $\mathbb{E}\left[\left(X-\mu\right)\left(\mu-a\right)\right]=\left(\mu-a\right)\mathbb{E}\left[\left(X-\mu\right)\right]=0$ and $\mathbb{E}\left[\left(\mu-a\right)^2\right]=\left(\mu-a\right)^2$.

5. A production process is partitioned into two independent sub-processes. The probabilities of a defective component in the first and second sub-processes are 0.01 and 0.02, respectively. If 50 units are produced, what is the probability there will be fewer than 3 defective units?
   Solution: Probability of a unit being non-defective is $(1-0.01)\times(1-0.02)$. Hence, the probability of a unit being defective is $1-(1-0.01)\times(1-0.02)=0.0302$. The desired probability is given as

$$\sum_{k=0}^{2}\binom{50}{k}\left(0.0302\right)^k\left(0.9698\right)^{50-k}\approx 0.8082$$

   We could also approximate this probability with a Poisson random variable with expected defectives to be $50\times 0.0302=1.51$. Hence, the desired probability is

$$\sum_{k=2}^{2}e^{-1.51}\frac{1.51^k}{k!}\approx 0.8063$$

6. Communication channels do not always trasmit the correct signal. Suppose that for a particular channel the error rate is 1 in 100, i.e., the probability of incorrect transmission is $1/100$. If 2000 messages are sent in a given week, and it is assumed that their transmissions are independent, what is the probability

that there will be at least 5 errors?
Solution: Desired probability is

$$1 - \sum_{k=0}^{4} \binom{2000}{k} (0.01)^k (0.99)^{2000-k} \approx 0.999984$$

Approximating it with Poisson, the expected number of incorrect transmissions is 20, we get the probability as

$$1 - e^{-20} \left(\sum_{k=0}^{4} \frac{20^k}{k!} \right) \approx 0.999983$$

7. A casino offers a game of chance for a single player in which a fair coin is tossed at each stage. The initial stake starts at \$1 and is increased by \$1 every time heads appears. The first time tails appears, the game ends and the player wins whatever is in the pot. Thus the player wins \$1 dollar if tails appears on the first toss, \$2 dollars if heads appears on the first toss and tails on the second, \$3 dollars if heads appears on the first two tosses and tails on the third, and so on. Mathematically, the player wins $\$k$ dollars, when we have the first $k-1$ tosses to be heads and the k^{th} toss to be a tail. The casino demands a pay of \$3 to enter the game. Will you play the game?
   Solution: Probability of winning on the k attempt is given by $\dfrac{1}{2^k}$, i.e., the first first $k-1$ tosses are heads and the k^{th} toss is a tail. Hence, the expected return is

$$\sum_{k=1}^{\infty} \frac{k}{2^k} = 2$$

   Since the casino demands a pay of \$3 to enter the game, which is greater than the expected return, it is not advisable to play the game.

8. Repeat the above if the price money was 2^k instead of k and the casino demands a pay of \$100 to enter the game. Will you still be willing to play the game? (For more details, look up St. Petersburg paradox)
   Solution: My expected return in this case is given by

$$\sum_{k=1}^{\infty} \frac{2^k}{2^k}$$

   which diverges. Hence, going by the expected value we could play the game. However, I will make a profit only when my return exceeds 100, which happens with a probability of $1/2^{100}$.

9. If X is a discrete random variable, prove that (i) $\mathbb{E}[aX+b] = a\mathbb{E}[X] + b$ and (ii) $\text{Var}[aX+b] = a^2 \text{Var}(X)$
   Solution: Follows immediately from definition.

$$\mathbb{E}[aX+b] = \sum_x (ax+b)\, p(x) = a \sum_x xp(x) + b \sum_x p(x) = a\mathbb{E}[X] + b$$

10. Data shows that 5% of the individuals reserving tables at a restaurant will not appear. If the restaurant has 50 tables and takes 52 reservations, what is the probability that it will be able to accommodate everyone appearing?
   Solution: Probability of failing to accommodate everyone is when 51 individuals or 52 individuals appear. This happens with a probability of

$$\binom{52}{51} (0.95)^{51} \times (0.05) + \binom{52}{52} (0.95)^{52} = 0.95^{51} \times (0.95 + 52 \times 0.05) = 3.55 \times 0.95^{51} \approx 0.2595$$

   Hence, desired probability is 0.7405.

11. Electrical power failures in a workplace are modeled as a Poisson experiment with a rate of one every two months.

    (a) What is the probability of having more than 10 failures in a year?
    Solution: The mean for a year is 6. Desired probability is $e^{-6}\sum_{k=11}^{\infty}\frac{6^k}{k!}\approx 0.04262$

    (b) What is the probability that the number of failures in a year will differ by more than a standard deviation from the expected number?
    Solution: The standard deviation for a year is $\sqrt{6}\approx 2.45$. Hence, the desired probability is

$$1-e^{-6}\left(\frac{6^4}{4!}+\frac{6^5}{5!}+\frac{6^6}{6!}+\frac{6^7}{7!}+\frac{6^8}{8!}\right)\approx 0.3039$$

12. Table 1 below indicates the joint probabilities per day.

Table 1: Joint probability of weather and power cuts

	Sunny	Rainy
Power cut	0.2	0.15
No power cut	0.6	p

    (a) Find p.
    Solution:Sum must be 1. Hence, $p = 0.05$.

    (b) What is the probability that there won't be rain for one week?
    Solution: Probability that it will be sunny is $0.2 + 0.6 = 0.8$. Hence, probability it won't rain for one week is 0.8^7.

    (c) What is the probability that there will be at least one power in the next three days?
    Solution: Probability of having no power cuts is 0.65. Hence, probability of having at least one power in the next three days $1-0.65^3 = 0.725375$.

    (d) Is there a dependence between weather and power cuts?
    Solution: Let X be the event of having a power cut and Y be the event of the day being sunny. We have

$$P(X) = 0.2 + 0.15 = 0.35$$

    We have

$$P(X \mid Y) = \frac{P(X,Y)}{P(Y)} = \frac{0.2}{0.2+0.6} = 0.25$$

    Hence, we see that there is a dependence between weather and power cuts.

    (e) Find the joint probability, all marginal probabilities, and all conditional probabilities.
    Solution: We have $P(X) = 0.35$, $P(X^c) = 0.65$, $P(Y) = 0.8$, $P(Y^c) = 0.2$.

$$P(X \mid Y) = \frac{P(X,Y)}{P(Y)} = 0.25 \quad P(X^c \mid Y) = 0.75$$

$$P(X \mid Y^c) = \frac{P(X,Y^c)}{P(Y^c)} = \frac{0.15}{0.2} = 0.75 \quad P(X^c \mid Y^c) = 0.25$$

$$P(Y \mid X) = \frac{P(X,Y)}{P(X)} = 4/7 \quad P(Y^c \mid X) = 3/7$$

$$P(Y \mid X^c) = \frac{P(X^c,Y)}{P(X^c)} = 12/13 \quad P(Y^c \mid X^c) = 1/13$$