INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination

M. Tech (CS) - II Year, 2013-2014 (Semester - I)

Probability and Stochastic Processes

Date : 06.09.2013 | Maximum Marks : 60 | Duration : 3 Hours

Note: The question is of 75 marks.

Answer as much as you can, but the maximum you can score is 60.

(Q1) Let there be n sticks each of which is broken into one long and one short part. The $2n$ parts are arranged into n pairs from which new sticks are formed. Find the probability that

(a) the parts will be joined in the original order.

(b) that all long parts are paired with short parts.

[5+3=8]

(Ans:) This problem has similarities with the problem of pairing couples that we did in the class. Think of the long part of the stick as female and the other part as male.

(Ans a:) The possible number of arrangements is $(2n)!$. Of them, fix a couple (a long and a short part of the same stick). There are n of them. They can be permuted in $n!$ ways. Now, in an arrangement, each couple can be arranged in 2 ways, giving 2^n for n couples. So, the said probability is $\frac{2^n n!}{(2n)!}$.

(Ans b:) Now permute n long sticks in $n!$ ways and n short sticks in $n!$ ways and pair them up in $(n!)^2$ ways. For each such arrangement, one can again order the long and short in 2 ways, leading to 2^n for n couples. Thus, the said probability is $\frac{2^n (n!)^2}{(2n)!}$ which for a better form is $\frac{2^n}{\binom{2n}{n}}$.

◀

(Q2) A parallel system having n components functions when at least one of the components functions. A component i, independent of other components, functions with probability p_i, $i = 1, \ldots, n$. What is the probability that the system functions? [5]

(Ans:) Let E_i be the event that the component i functions properly. The parallel system functions if at least one of the n independent component functions. Put in other words, the system does not function if all of the n components do not function. That probability is

$$\Pr(\cap_{i=1}^{n} \overline{E_i}) = \Pi_{i=1}^{n}(1 - p_i)$$

as the components are independent and each functions properly with a probability p_i. Therefore, the probability that the system functions is

$$1 - \Pi_{i=1}^{n}(1 - p_i).$$

◀

(Q3) State and prove Bayes' formula. [7]

(Ans:) Let $B_1, B_2, \ldots, B_n$ be n mutually exclusive and exhaustive events. We have

$$A = \bigcup_{i=1}^{n} AB_i.$$

As events AB_i are mutually exclusive, we have

$$\begin{aligned} \Pr(A) &= \sum_{i=1}^{n} \Pr(AB_i) \\ &= \sum_{i=1}^{n} \Pr(A \mid B_i)\ \Pr(B_i) \end{aligned}$$

Bayes' formula finds the conditional probability of any B_j, $j = 1, \ldots, n$, given that A has occurred.

$$\begin{aligned} \Pr(B_j \mid A) &= \frac{\Pr(AB_j)}{\Pr(A)} \\ &= \frac{\Pr(A \mid B_j)\ \Pr(B_j)}{\sum_{i=1}^{n} \Pr(A \mid B_i)\ \Pr(B_i)} \end{aligned}$$

◀

(Q4) Consider a collection of $N+1$ boxes, each containing a total of N red and white balls. The box number i contains i red and $N-i$ white balls ($i = 0, 1, \ldots, N$). A box is chosen at random and n random draws are made from it, the ball drawn being replaced each time. Find the conditional probability that the $n+1$-th drawing from the box will also yield a red ball given that all the prior n balls turn out to be red. [8]

(Ans:) Let A be the event that all the n random draws turn out to be red balls from a randomly selected box. Let B be the event that the $n+1$-th draw from the same box is also red. We want to find $\Pr(B|A)$. Let $E_i, i = 1, \ldots, N+1$ denote the event that the i-th box was selected. $\Pr(E_i) = \frac{1}{N+1}$ and $\Pr(A|E_i) = \left(\frac{i}{N}\right)^n$. Notice that $\Pr(AB)$ denotes that all the $n+1$ random draws are red from the same box.

$$\begin{aligned} \Pr(B|A) &= \frac{\Pr(AB)}{\Pr(A)} \\ &= \frac{\Pr(AB)}{\sum_{i=0}^{N} \Pr(A|E_i)\Pr(E_i)} \\ &= \frac{\Pr(AB)}{\frac{1}{N+1}\sum_{i=0}^{N}\left(\frac{i}{N}\right)^n} \\ &= \frac{\frac{1}{N+1}\sum_{i=0}^{N}\left(\frac{i}{N}\right)^{(n+1)}}{\frac{1}{N+1}\sum_{i=0}^{N}\left(\frac{i}{N}\right)^n} \\ &= \frac{1^{n+1} + 2^{n+1} + 3^{n+1} + \ldots + N^{n+1}}{N(1^n + 2^n + 3^n + \ldots + N^n)} \end{aligned}$$

For large N, the above expression can be approximated as $\frac{n+1}{n+2}$. ◀

(Q5) Independent trials, each resulting in a success with probability p or a failure with probability $q = 1 - p$, are performed. Compute the probability that a run of n consecutive successes occurs before a run of m consecutive failures. [10]

(Ans:) Notice a nature of this problem. Supposing we are having a sequence of successes. Once we reach n consecutive successes before m consecutive failures, we are done. But say before reaching n consecutive successes, we encounter a failure. This wipes out the run of successes and brings us back to square one. This characteristic of the problem suggests that there is a recursive nature to this problem. Thus the solution idea is to condition on some trials and then deduce a recursive formula.

Let E be the event that a run of n consecutive successes occur before a run of m consecutive failures. Let T_1 be the event that the first trial results in a success. Then,

$$\Pr(E) = p\Pr(E \mid T_1) + q\Pr(E \mid \overline{T_1}) \tag{1}$$

Let us determine $\Pr(E \mid T_1)$ and $\Pr(E \mid \overline{T_1})$. We condition again on $S_{2,n}$, the event that all trials between 2 and n result in successes. So, we have

$$\Pr(E \mid T_1) = \Pr(E \mid S_{2,n}T_1) \cdot \Pr(S_{2,n} \mid T_1) + \Pr(E \mid \overline{S_{2,n}}T_1) \cdot \Pr(\overline{S_{2,n}} \mid T_1) \tag{2}$$

Notice that $\Pr(E \mid S_{2,n}T_1) = 1$. Trials are independent, so events $S_{2,n}$ and T_1 are independent and hence, $\Pr(S_{2,n} \mid T_1) = \Pr(S_{2,n}) = p^{n-1}$ and $\Pr(\overline{S_{2,n}} \mid T_1) = \Pr(\overline{S_{2,n}}) = 1 - p^{n-1}$. We are left to tackle $\Pr(E \mid \overline{S_{2,n}}T_1)$. This event says that we started with a success and kept on having successes starting from the second trial onwards and had a failure before the n-th success. So, this wipes out all previous successes and brings us to a situation where we start with a failure, i.e. $\Pr(E \mid \overline{S_{2,n}}T_1) = \Pr(E \mid \overline{T_1})$. Replacing the above observations in Equation 2, we have

$$\begin{aligned}\Pr(E \mid T_1) &= \Pr(E \mid S_{2,n}T_1) \cdot \Pr(S_{2,n} \mid T_1) + \Pr(E \mid \overline{S_{2,n}}T_1) \cdot \Pr(\overline{S_{2,n}} \mid T_1) \\ \Pr(E \mid T_1) &= p^{n-1} + (1 - p^{n-1})\Pr(E \mid \overline{T_1}) \end{aligned} \tag{3}$$

So, now we need to determine $\Pr(E \mid \overline{T_1})$. We again condition on $F_{2,m}$, the event that all trials between 2 and m result in failures. We have

$$\Pr(E \mid \overline{T_1}) = \Pr(E \mid F_{2,m}\overline{T_1}) \cdot \Pr(F_{2,m} \mid \overline{T_1}) + \Pr(E \mid \overline{F_{2,m}}\,\overline{T_1}) \cdot \Pr(\overline{F_{2,m}} \mid \overline{T_1}) \tag{4}$$

Notice that $\Pr(E \mid F_{2,m}\overline{T_1}) = 0$ as $F_{2,m}\overline{T_1}$ denotes the event that m failures have occurred before n successes. By similar logic as earlier, we would have $\Pr(E \mid \overline{F_{2,m}}\,\overline{T_1}) = \Pr(E \mid T_1)$ and $\Pr(\overline{F_{2,m}} \mid \overline{T_1}) = \Pr(\overline{F_{2,m}}) = 1 - q^{m-1}$. So, we have

$$\begin{aligned}\Pr(E \mid \overline{T_1}) &= \Pr(E \mid F_{2,m}\overline{T_1}) \cdot \Pr(F_{2,m} \mid \overline{T_1}) + \Pr(E \mid \overline{F_{2,m}}\,\overline{T_1}) \cdot \Pr(\overline{F_{2,m}} \mid \overline{T_1}) \\ \Pr(E \mid \overline{T_1}) &= (1 - q^{m-1})\Pr(E \mid T_1) \end{aligned} \tag{5}$$

Replacing Equation 5 in Equation 3, we have

$$\begin{aligned}\Pr(E \mid T_1) &= p^{n-1} + (1 - p^{n-1})(1 - q^{m-1})\Pr(E \mid T_1) \\ \Pr(E \mid T_1) &= \frac{p^{n-1}}{p^{n-1} + q^{m-1} - p^{n-1}q^{m-1}} \end{aligned} \tag{6}$$

Replacing Equation 6 in Equation 1, we have

$$\begin{aligned}
\Pr(E) &= p\Pr(E \mid T_1) + q\Pr(E \mid \overline{T_1}) \\
&= p\Pr(E \mid T_1) + q\{(1-q^{m-1})\Pr(E \mid T_1)\} \\
&= \Pr(E \mid T_1)(p + q + q^{m-1}) \\
&= \Pr(E \mid T_1)(1 + q^{m-1}) \\
&= \frac{p^{n-1}(1-q^m)}{p^{n-1} + q^{m-1} - p^{n-1}q^{m-1}} \qquad (7)
\end{aligned}$$

◀

(Q6) Consider the following gambling game. A player holds a bet on any one of the numbers $\{1, 2, 3, 4, 5, 6\}$. Three dice are then rolled, and if the number bet by the player appears i times, $i = 1, 2, 3$, then the player wins i units. On the other hand, if the number bet by the player does not appear on any of the dice, then the player loses 1 unit. Find out if the game is fair to the player. [5]

(Ans:) Let X be a random variable that denotes the amount won by the player. X takes the value -1, 1, 2 and 3. The dice are independent. Let Y be a random variable that denotes the number of times the number bet by the player matches the number in the three dice rolled. Y can take the values 0, 1, 2 and 3. Also, Y follows binomial with parameters $(3, \frac{1}{6}$.

So, $\Pr(X = -1) = \Pr(Y = 0) = \binom{3}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^3 = \frac{125}{216}$. Similarly, $\Pr(X = 1) = \Pr(Y = 1) = \binom{3}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^2 = \frac{75}{216}$; and $\Pr(X = 2) = \Pr(Y = 2) = \binom{3}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^1 = \frac{15}{216}$; and $\Pr(X = 3) = \Pr(Y = 3) = \binom{3}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^0 = \frac{1}{216}$.

$E[X]$, the expectation of X, is basically the amount won by the player. $E[X] = \frac{-125+75+30+3}{216} = \frac{-17}{216}$. As $E[X] < 0$, the player would lose in the expected case and as such the game is not fair to the player. ◀

(Q7) Let X be a random variable defined over a sample space Ω such that $\text{E}[X] = \mu$. Show that $\Pr(X \geq \mu) > 0$ and $\Pr(X \leq \mu) > 0$. [3+3=6]

[Hints: Can you try to prove using contradiction?]

(Ans:) Assume, for a contradiction, $\Pr(X \geq \mu) = 0$. Then,

$$\mu = \text{E}[X] = \sum_x x\Pr(X = x) = \sum_{x<\mu} x\Pr(X = x) < \sum_{x<\mu} \mu\Pr(X = x) = \mu$$

which cannot be. So, $\Pr(X \geq \mu) \neq 0$ and as $\Pr(\cdot) \geq 0$, we have the result.

Similarly, assume $\Pr(X \leq \mu) = 0$. Then,

$$\mu = \text{E}[X] = \sum_x x\Pr(X = x) = \sum_{x>\mu} x\Pr(X = x) > \sum_{x>\mu} \mu\Pr(X = x) = \mu$$

which cannot be. So, $\Pr(X \leq \mu) \neq 0$ and as $\Pr(\cdot) \geq 0$, we have the result. ◀

(Q8) Suppose we roll a standard fair die 200 times. Let X be the sum of the numbers that appear over the 200 rolls. Use Chebyshev's inequality to bound $\mathbf{Pr}[X \geq 750]$. [6]

(Ans:) Let X_i be the random variable that denotes the value obtained in the i^{th} roll of the die, X_i can take any integral value in the range $[1, 6]$ with equal probability. So, $E[X_i] = \frac{7}{2}$ and $\text{Var}[X_i] = E[X_i^2] - (E[X_i])^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$. Since $X = \sum_{i=1}^{200} X_i$, $E[X] = 200 \times \frac{7}{2} = 700$

and as X_i's are independent, $\text{Var}[X] = 200 \times \frac{35}{12} = \frac{1750}{3}$. Now using Chebyshev's inequality, we have

$$\begin{aligned}
\text{Pr}[X \geq 750] &\leq \text{Pr}[|X - 700| \geq 50] \\
&\leq \frac{\text{Var}[X]}{50^2} \\
&= \frac{1750}{3 \times 2500} \\
&= \frac{7}{30}.
\end{aligned}$$

◀

(Q9) Let X be a uniform random variable on $[u, v]$. Show that $\text{Pr}\left(X \leq m \mid X \leq n\right) = \frac{m-u}{n-u}$, where $u \leq m \leq n \leq v$. [6]

(Ans:)

$$\begin{aligned}
\text{Pr}\left(X \leq m \mid X \leq n\right) &= \frac{\text{Pr}\left((X \leq m) \cap (X \leq n)\right)}{\text{Pr}\left(X \leq n\right)} \\
&= \frac{\text{Pr}\left(X \leq m\right)}{\text{Pr}\left(X \leq n\right)} \\
&= \frac{m-u}{n-u}.
\end{aligned}$$

X being conditioned on $X \leq n$, has a distribution function that is uniform on the reduced interval $[u, n]$. ◀

(Q10) Let $X_1, \ldots, X_n$ be independent uniform random variables over $[0, 1]$. Let $Y_1 = \min(X_1, \ldots, X_n)$. Show that $\text{E}[Y_1] = \frac{1}{n+1}$. [8]

(Ans:) We know for a continuous random variable X, that takes only nonnegative values, and has a density function $f(x)$,

$$\begin{aligned}
\int_{x=0}^{\infty} \text{Pr}\left(X \geq x\right)\, dx &= \int_{x=0}^{\infty} \int_{y=x}^{\infty} f(y)\, dy\, dx \\
&= \int_{y=0}^{\infty} \int_{x=0}^{y} f(y)\, dx\, dy \\
&= \int_{y=0}^{\infty} y f(y)\, dy \\
&= E[X].
\end{aligned}$$

Now, as per the question, $Y_1 = \min(X_1, \ldots, X_n)$ and Y_1 takes value in the range $[0, 1]$. We want to find $E[Y_1]$ which is $\int_{y=0}^{1} \text{Pr}(Y_1 \geq y)$. So, we need to figure out the probability of $\text{Pr}(Y_1 \geq y)$.

$$\begin{aligned}
\text{Pr}(Y_1 \geq y) &= \text{Pr}(\min(X_1, \ldots, X_n) \geq y) \\
&= \text{Pr}\left((X_1 \geq y) \cap (X_2 \geq y) \cap \cdots \cap (X_n \geq y)\right)
\end{aligned}$$

$$
\begin{aligned}
&= \prod_{i=1}^{n} \Pr(X_i \geq y) \\
&= (1-y)^n
\end{aligned}
$$

Therefore, $E[Y_1] = \int\limits_{y=0}^{1} (1-y)^n \, dy = \frac{1}{n+1}$. ◀

(Q11) For an exponential random variable show that $\Pr(X > s + t \mid X > t) = \Pr(X > s)$. [6]

(Ans:)

$$
\begin{aligned}
\Pr(X > s + t \mid X > t) &= \frac{\Pr((X > s+t) \cap (X > t))}{\Pr(X > t)} \\
&= \frac{\Pr((X > s+t)}{\Pr(X > t)} \\
&= \frac{1 - \Pr(X \leq s+t)}{1 - \Pr(X \leq t)} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} \\
&= e^{-\lambda s} \\
&= 1 - \Pr(X \leq s) \\
&= \Pr(X > s)
\end{aligned}
$$

◀