Final #1

Mark the correct answer in each part of the following questions.

1. An urn contains 29 copper balls and one gold ball. Three players – A, B, and C – compete for the gold ball as follows. First, A draws a random ball, then B, then C, then again A, B, C, and so forth. The drawings are **without replacement**, so the game is over after at most 30 steps. Let p_A, p_B, p_C denote the probability of A, B, and C, respectively, to get the gold ball.

   (a) The numbers p_A, p_B, p_C are ordered as follows:

   (i) $p_A > p_B > p_C$.

   (ii) $p_A < p_B < p_C$.

   (iii) $p_A = p_B = p_C$.

   (iv) $p_A = p_C > p_B$.

   (v) None of the above.

   (b) Now suppose that the balls are drawn **with replacement**. Let p'_A, p'_B, p'_C be the winning probabilities under the new rules. Then:

   (i) $p'_A = \frac{1}{3}$.

   (ii) $p'_B = \frac{p'_A + p'_C}{2}$.

   (iii) $p'_B = \frac{29}{30} \cdot p'_C$.

   (iv) $p'_B = \frac{30}{29} \cdot p'_C$.

   (v) None of the above.

(c) Suppose now that the rules of the game are set in a way that gives each player a winning probability of 1/3. However, there is a referee R in the game. R dislikes A, is neutral towards B, and is a good friend of C. Hence, if A wins – the referee cancels the game with some pretext and instructs to restart it. If B wins, R accepts the result or rejects it with probabilities of 1/2 each. If C wins, R accepts the result. Let X be the number of games held until one of the players wins. (For example, if A wins twice, then B wins but R decides to reject the result, then A wins again twice, and then C wins, we have $X = 6$.) Then X is distributed

(i) $G(1/3)$.

(ii) $G(1/2)$.

(iii) $\bar{B}(2, 1/3)$.

(iv) $\bar{B}(2, 1/2)$.

(v) None of the above.

(d) Continue with the assumptions of the preceding part. Given that R has rejected the outcome of the first round, the probability that the winner of that round was A is

(i) $\frac{1}{3}$.

(ii) $\frac{1}{2}$.

(iii) $\frac{2}{3}$.

(iv) $\frac{3}{4}$.

(v) None of the above.

2. Consider the absent-minded secretary problem. Let $n \geq 2$ denote the number of letters and envelopes and X the number of letters sent to the correct destination.

(a) Suppose that $n \geq 7$. Employing Markov's inequality to bound the probability that at least seven letters are sent to the correct address, we obtain:

(i) $P(X \geq 7) \leq \frac{1}{49}$.

(ii) $P(X \geq 7) \leq \frac{1}{36}$.

(iii) $P(X \geq 7) \leq \frac{1}{7}$.

(iv) $P(X \geq 7) \leq \frac{1}{6}$.

(v) None of the above.

Remark: We mean the best bound that can be obtained. For example, if (i) is correct, then (ii)-(iv) are correct as well, but only (i) should be marked as the correct answer.

(b) We want to bound $P(X \geq n)$ using Chernoff's bound.

(i) The optimal t to use is $t_0 = 1 - 1/n$.

(ii) The moment-generating function of X is undefined (except at the point $t = 0$), so it is impossible to use Chernoff's bound.

(iii) Chernoff's bound gives a trivial result for every value of t.

(iv) When we use Chernoff's bound with $t \to \infty$, the resulting bound converges to $P(X \geq n)$.

(v) None of the above.

(c) Suppose that n is large. Consider the function F_X at the points 10 and 10.5.

(i) F_X is differentiable at both points.

(ii) F_X is differentiable at 10 and undefined at 10.5.

(iii) F_X is continuous and non-differentiable at 10 and is differentiable at 10.5.

(iv) F_X has a discontinuity at 10 and is differentiable at 10.5.

(v) None of the above.

An organization, in which letters are regularly sent as in the absent-minded secretary problem, has designed the following way to increase the percentage of letters sent to the correct address.

A secretary places the letters in the envelopes as usual in the organization. Then a second secretary goes over the letters in the envelopes, leaves in the envelope each letter that matches the envelopes, and then places randomly the remaining letters in the remaining envelopes. Suppose that in the beginning there are $n \geq 2$ letters and matching envelopes. Denote by X the number of letters put in the correct envelopes by the first secretary out of the initial n letters, and by Y the number of additional letters put in the correct envelopes by the second secretary out of the remaining $n - X$ letters. (If $X = n$, then $Y = 0$.)

(d) $E(Y)$ relates to $E(X)$ as follows: $E(Y) =$

(i) $\left(1 - \frac{1}{n}\right) \cdot E(X)$.

(ii) $\left(1 - \frac{1}{n!}\right) \cdot E(X)$.

(iii) $E(X)$.

(iv) $\left(1 + \frac{1}{n!}\right) \cdot E(X)$.

(v) None of the above.

(e) $V(Y|X) =$

(i) $\begin{cases} 1 - \dfrac{1}{n \cdot n!}, & X < n, \\ 0, & X = n. \end{cases}$

(ii) $\begin{cases} 1 - \dfrac{1}{n!}, & X < n, \\ 0, & X = n. \end{cases}$

(iii) $\begin{cases} 1 - \dfrac{1}{n}, & X < n, \\ 0, & X = n. \end{cases}$

(iv) $\begin{cases} 1, & X < n, \\ 0, & X = n. \end{cases}$

(v) None of the above.

3. A drunkard D is located at time 0 at the point 0 on the x-axis. During each time unit, D does one of the following:

- remains at the same location – with probability 1/6;
- moves one unit to the right along the x-axis – with probability 1/2;
- moves two units to the right along the x-axis – with probability 1/3.

D's actions during distinct time units are independent. Denote by R the number of time units, out of the first n, in which D moves to the right (be it one unit or two), and by X his location at time n. (For example, if $n = 8$ and D first took twice option (i), then twice option (ii), and then four times option (iii), we have $R = 6, X = 10$.)

(a) Consider the random variables $R, X, X - R$, and $2R - X$. Out of these:

(i) exactly one is binomial.
(ii) exactly two are binomial.
(iii) exactly three are binomial.
(iv) all four are binomial.
(v) None of the above.

(b) $P_{R,X}(r, x) =$

(i)

$$\begin{cases} \frac{n!}{(n-x)!(x-r)!r!} \cdot \frac{2^{x-r}3^{n-x}}{6^n}, & 0 \le r \le x \le 2r \le 2n, \\ 0, & \text{otherwise.} \end{cases}$$

(ii)

$$\begin{cases} \frac{n!}{(n-r)!(2r-x)!(x-r)!} \cdot \frac{2^{x-r}3^{2r-x}}{6^n}, & 0 \le r \le x \le 2r \le 2n, \\ 0, & \text{otherwise.} \end{cases}$$

(iii)

$$\begin{cases} \frac{n!}{x!r!(n-x-r)!} \cdot \frac{2^{x}3^{n-2r+x}}{6^n}, & 0 \le r \le x \le 2r \le 2n, \\ 0, & \text{otherwise.} \end{cases}$$

(iv)

$$\begin{cases} \frac{n!}{(2r-x)!x!(r+n-2x)!} \cdot \frac{2^{r}3^{n-x}}{6^n}, & 0 \le r \le x \le 2r \le 2n, \\ 0, & \text{otherwise.} \end{cases}$$

(v) None of the above.

(c) Denote by M_X the moment-generating function of X. If $n = 2$, then $M_X(\log 7) =$

(i) 400.

(ii) 500.

(iii) 600.

(iv) 700.

(v) None of the above.

(d) $E(R|X = 2) =$

(i) $2 - \dfrac{3}{9n - 5}$.

(ii) $2 - \dfrac{4}{9n - 5}$.

(iii) $2 - \dfrac{5}{9n - 5}$.

(iv) $2 - \dfrac{6}{9n - 5}$.

(v) None of the above.

(e) $V(RX|R) =$

(i) $\dfrac{6R^2}{25}$.

(ii) $\dfrac{6R^3}{25}$.

(iii) $\dfrac{R^2}{4}$.

(iv) $\dfrac{R^3}{4}$.

(v) None of the above.

(f) $\rho(R, X) =$

   (i) $\frac{3}{\sqrt{85}}$.

   (ii) $\frac{4}{\sqrt{85}}$.

   (iii) $\frac{6}{\sqrt{85}}$.

   (iv) $\frac{7}{\sqrt{85}}$.

   (v) None of the above.

Solutions

1. (a) The gold ball has the same probability of $1/30$ of being the first ball to be drawn, the second to be drawn, and so forth. Hence, each of the players has a probability of $10/30 = 1/3$ to get it.

    Thus, (iii) is true.

    (b) The probability of the gold ball to be drawn for the first time in the n-th drawing is $(29/30)^{n-1} \cdot 1/30$ for every $n \geq 1$. Hence,

    $$p'_A = (29/30)^0 \cdot 1/30 + (29/30)^3 \cdot 1/30 + (29/30)^6 \cdot 1/30 + \cdots$$

    The expression for p'_B is the same, except that the powers of $29/30$ of corresponding terms are larger by 1. For p'_C, they are larger by 2. It follows that:

    $$p'_B = \frac{29}{30} \cdot p'_A, \qquad p'_C = \left(\frac{29}{30}\right)^2 \cdot p'_A.$$

    Thus, (iv) is true.

    (c) The probability for the result of any game to be accepted by the referee is

    $$\frac{1}{3} \cdot 0 + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 1 = \frac{1}{2}.$$

    Moreover, the results of distinct games are independent. Considering the acceptance of the result of a game by the referee as a success, we see that X counts the trials until the first success in a sequence of independent trials of success probability $1/2$.

    Thus, (ii) is true.

    (d) Define the events: A' – A wins, B' – B wins, R' – R rejects the result. Then:

    $$P(A'|R') = \frac{P(A' \cap R')}{P(R')} = \frac{1/3 \cdot 1}{1/3 \cdot 1 + 1/3 \cdot 1/2 + 1/3 \cdot 0} = \frac{2}{3}.$$

    Thus, (iii) is true.

2. (a) We have seen in class that $E(X) = 1$. Therefore:

$$P(X \geq 7) \leq \frac{E(X)}{7} = \frac{1}{7}.$$

Thus, (iii) is true.

(b) X is bounded between 0 and n, so that M_X is well-defined on $\mathbf{R}$. Obviously:

$$P(X \geq n) = P(X = n) = 1/n!.$$

Hence:

$$M_X(t) = P_X(0) + P_X(1)e^t + \cdots + e^{nt}/n!, \quad t \in \mathbf{R}.$$

Chernoff's Inequality gives:

$$\begin{aligned} P(X \geq n) &\leq \frac{M_X(t)}{e^{nt}} \\ &= P_X(0)e^{-nt} + P_X(1)e^{-(n-1)t} + \cdots + 1/n!, \qquad t \in \mathbf{R}. \end{aligned}$$

As $t \to \infty$, all terms on the right-hand side but the last decay to 0. Hence the bound converges to $1/n!$, which is $P(X \geq n)$.

Thus, (iv) is true.

(c) Since X assumes only integer values, F_X is constant along every open interval $(m, m+1)$ between consecutive integers. In particular, it is differentiable at 10.5. On the other hand, since X has a positive probability of assuming the values $0, 1, 2, \ldots, n-2, n$, the function has jump discontinuities at these points. In particular, for $n \geq 12$ it is discontinuous at the point 10.

Thus, (iv) is true.

(d) We have seen in class that $E(X) = 1$. Hence, unless all letters are placed in the correct envelopes in the first stage, the expected of letters being placed correctly in the second stage is again 1. Thus:

$$\begin{aligned} E(Y) &= P(X < n)E(Y|X < n) + P(X = n)E(Y|X = n) \\ &= (1 - 1/n!) \cdot 1 + 1/n! \cdot 0 \\ &= 1 - 1/n!. \end{aligned}$$

Thus, (ii) is true.

(e) We have seen in class that $V(X) = 1$. Hence, unless $X = n$, the variance of the number of letters placed correctly in the second stage is 1 as well.

Thus, (iv) is true.

3. (a) R counts all moves (of positive length). At each time unit, there is a probability of 5/6 of a move, and time units are independent. Hence, $R \sim B(n, 5/6)$.

$X - R$ counts moves of two distance units. Thus, $X - R \sim B(n, 1/3)$. Similarly, $2R - X$ counts moves of one distance unit, so that $2R - X \sim B(n, 1/2)$.

We claim that X is not binomially distributed. Indeed, since it clearly assumes the values $0, 1, \ldots, 2n$, if it was binomial, we would have $X \sim B(2n, p)$ for some p. Now, $P(X = 0) = 1/6^n$ and $P(X = 2n) = 1/3^n$, which imply that $q = 1/\sqrt{6}$ and $p = 1\sqrt{3}$, respectively. The contradiction proves the claim.

Thus, (iii) is true.

(b) The event $(R, X) = (r, x)$ occurs if D stays in place $n-r$ times and makes $2r - x$ single moves and $x - r$ double moves. Consequently, for feasible pairs (r, x):

$$P_{R,X}(r, x) = \binom{n}{n-r}\left(\frac{1}{6}\right)^{n-r}\left(\frac{5}{6}\right)^{r}\cdot\binom{r}{2r-x}\left(\frac{3}{5}\right)^{2r-x}\left(\frac{2}{5}\right)^{x-r}.$$

Canceling out equal terms, we finally obtain:

$$P_{R,X}(r, x) = \frac{n!}{(n-r)!(2r-x)!(x-r)!}\cdot\frac{2^{x-r}3^{2r-x}}{6^n}$$

Thus, (ii) is true.

(c) X may be written in the form

$$X = X_1 + X_2 + \cdots + X_n,$$

where X_i is the number of distance units D covered at the i-th time unit, $1 \le i \le n$. Now

$$M_{X_i}(t) = E\left(e^{tX_i}\right) = \frac{1}{6}\cdot e^{t\cdot 0} + \frac{1}{2}\cdot e^{t\cdot 1} + \frac{1}{3}\cdot e^{t\cdot 2},$$

and since the X_i-s are independent,

$$M_X(t) = \left(\frac{1}{6} + \frac{1}{2}\cdot e^t + \frac{1}{3}\cdot e^{2t}\right)^n.$$

For $n = 2$:

$$\begin{aligned} M(\log 7) &= \left(\frac{1}{6} + \frac{1}{2}\cdot e^{\log 7} + \frac{1}{3}\cdot e^{2\log 7}\right)^2 \\ &= \left(\frac{1}{6} + \frac{1}{2}\cdot 7 + \frac{1}{3}\cdot 7^2\right)^2 \\ &= 20^2 \\ &= 400. \end{aligned}$$

Thus, (i) is true.

(d) Given that $X = 2$, there have been either one double move and $n-1$ non-moves, or two single moves and $n-2$ non-moves. Denoting these two events by A and B, respectively, we have:

$$\begin{aligned} P(R = 2|X = 2) &= \frac{P(R = X = 2)}{P(X = 2)} \\ &= \frac{\binom{n}{2}(1/2)^2(1/6)^{n-2}}{\binom{n}{2}(1/2)^2(1/6)^{n-2} + \binom{n}{1}(1/3)^1(1/6)^{n-1}} \\ &= 1 - \frac{4}{9n-5}. \end{aligned}$$

It follows that:

$$E(R|X = 2) = \frac{4}{9n-5}\cdot 1 + \left(1 - \frac{4}{9n-5}\right)\cdot 2 = 2 - \frac{4}{9n-5}.$$

Thus, (ii) is true.

(e) Given R, each of the R moves is a single move with a probability of $3/5$ and a double move with a probability of $2/5$. Moreover, the lengths of distinct moves are independent. Hence, $X - R|R \sim B(R, 2/5)$. It follows that

$$V(RX|R) = R^2 V(X|R) = R^2 \cdot R \cdot \frac{2}{5} \cdot \frac{3}{5} = \frac{6R^3}{25}.$$

Thus, (ii) is true.

(f) Since $R \sim B(n, 5/6)$, we have

$$E(R) = n \cdot \frac{5}{6} = \frac{5n}{6}$$

and

$$V(R) = n \cdot \frac{5}{6} \cdot \frac{1}{6} = \frac{5n}{36}.$$

Regarding X, since $X - R|R \sim B(R, 2/5)$, we have

$$E(X) = E(E(X|R)) = E\left(R + R \cdot \frac{2}{5}\right) = E\left(R \cdot \frac{7}{5}\right) = \frac{7}{5} \cdot \frac{5n}{6} = \frac{7n}{6}.$$

Also:

$$\begin{aligned} V(X) &= E(V(X|R)) + V(E(X|R)) \\ &= E\left(R \cdot \frac{2}{5} \cdot \frac{3}{5}\right) + V\left(\frac{7R}{5}\right) \\ &= \frac{5n}{6} \cdot \frac{6}{25} + \frac{7^2}{5^2} \cdot \frac{5n}{36} \\ &= \frac{17n}{36}. \end{aligned}$$

Next we calculate $E(RX)$:

$$\begin{aligned} E(RX) &= E(E(RX|R)) \\ &= E(RE(X|R)) \\ &= E\left(\frac{7R^2}{5}\right) \\ &= \frac{7}{5}\left(V(R) + E^2(R)\right) \\ &= \frac{7}{5}\left(\frac{5n}{36} + \frac{25n^2}{36}\right) \\ &= \frac{35n^2}{36} + \frac{7n}{36}. \end{aligned}$$

Hence:

$$\begin{aligned} \mathrm{Cov}(R, X) &= E(RX) - E(R)E(X) \\ &= \frac{35n^2}{36} + \frac{7n}{36} - \frac{5n}{6} \cdot \frac{7n}{6} \\ &= \frac{7n}{36}. \end{aligned}$$

Finally:

$$\rho(R, X) = \frac{7n/36}{\sqrt{17n/36}\sqrt{5n/36}} = \frac{7}{\sqrt{85}}.$$

Alternative Solution: For $1 \leq i \leq n$, put:

$$Y_i = \begin{cases} 1, & \text{the drunkard made one step during the } i\text{-th time unit,} \\ 0, & \text{otherwise,} \end{cases}$$

$$Z_i = \begin{cases} 1, & \text{the drunkard made two steps during the } i\text{-th time unit,} \\ 0, & \text{otherwise.} \end{cases}$$

Clearly,

$$R = Y_1 + \cdots + Y_n + Z_1 + \cdots + Z_n$$

and

$$X = Y_1 + \cdots + Y_n + 2Z_1 + \cdots + 2Z_n.$$

Note that $Y_i Z_i = 0$, so that $\mathrm{Cov}(Y_i, Z_i) = -E(Y_i)E(Z_i)$ for $1 \leq i \leq n$. Also, for $i \neq j$, the pair of variables Y_i and Y_j, the pair Y_i and Z_j, and the pair Z_i and Z_j consist of independent variables each. We have

$$\begin{aligned} V(R) &= nV(Y_1) + nV(Z_1) - 2nE(Y_1)E(Z_1) \\ &= \frac{n}{4} + \frac{2n}{9} - \frac{2n}{2 \cdot 3} \\ &= \frac{5n}{36}, \end{aligned}$$

$$\begin{aligned} V(X) &= nV(Y_1) + 4nV(Z_1) - 4nE(Y_1)E(Z_1) \\ &= \frac{n}{4} + \frac{8n}{9} - \frac{4n}{2 \cdot 3} \\ &= \frac{17n}{36}. \end{aligned}$$

Finally,

$$\begin{aligned} \mathrm{Cov}(R, X) &= nV(Y_1) + 2nV(Z_1) - 3nE(Y_1)E(Z_1) \\ &= \frac{n}{4} + \frac{4n}{9} - \frac{3n}{2 \cdot 3} \\ &= \frac{7n}{36}. \end{aligned}$$

Hence,

$$\rho(R, X) = \frac{7}{\sqrt{85}}.$$

Thus, (iv) is true.