# Final #1

Mark the correct answer for each part of each question.

1. Consider the following two-stage experiment. In the first stage, a die is rolled until it first shows 6. Denote by $X$ the number of rolls. In the second stage, each of two mailmen, Reuven and Shimon, gets $X$ letters and $X$ matching addressed envelopes, and needs to put the letters in the envelopes. Both of them work according to the negligent secretary method. Let $R$ be the number of letters Reuven puts in the matching envelopes and $S$ the analogous number for Shimon. Let $R' = X - R$ and $S' = X - S$ be the number of letters Reuven and Shimon put in non-matching envelopes.

   (a) Markov's Inequality yields:

      (i) $P(R \geq 100) \leq \dfrac{1}{100!}$.

      (ii) $P(R \geq 100) \leq \dfrac{1}{2^{100}}$.

      (iii) $P(R \geq 100) \leq \dfrac{1}{100^2}$.

      (iv) $P(R \geq 100) \leq \dfrac{1}{100}$.

      (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) $P(R = 100 \mid X = 102) =$

(i) $\dfrac{1}{2 \cdot 100!}$.

(ii) $\dfrac{1}{100!}$.

(iii) $\dfrac{2}{100!}$.

(iv) $\dfrac{4}{100!}$.

(v) None of the above.

(c) The probability that Reuven puts all the letters he receives in the matching envelopes is

(i) $\dfrac{e^{5/6} - 1}{12}$.

(ii) $\dfrac{e^{5/6} - 1}{6}$.

(iii) $\dfrac{e^{5/6} - 1}{5}$.

(iv) $\dfrac{e^{5/6} - 1}{2}$.

(v) None of the above.

(d) $\rho(R, S)$ lies in the interval:

(i) $[-1, -0.99]$.

(ii) $[-0.01, 0.01]$.

(iii) $[1/e - 0.01, 1/e + 0.01]$.

(iv) $[0.99, 1]$.

(v) None of the above.

(e) $\rho(R', S') =$

(i) $\dfrac{29}{30}$.

(ii) $\dfrac{30}{31}$.

(iii) $\dfrac{31}{32}$.

(iv) $\dfrac{32}{33}$.

(v) None of the above.

2. In a certain game, the win $X$ in shekels is distributed according to the following table:

| $x$ | $-1$ | $1$ | $5$ |
|---|---|---|---|
| $p$ | 0.8 | 0.15 | 0.05 |

A player participates in the game $n$ times. Let $S$ denote his total winnings and $N_{-1}, N_1$, and $N_5$ denote the numbers of times he won $-1, 1$, and 5 shekels, respectively.

(a) $M_X(\log 2) =$

(i) 2.3.

(ii) 2.4.

(iii) 2.5.

(iv) 2.6.

(v) None of the above.

(b) $V(V(N_5 \mid N_1)) =$

(i) $\dfrac{3 \cdot 4}{5 \cdot 17^3} n$.

(ii) $\dfrac{3 \cdot 4^2}{5 \cdot 17^3} n$.

(iii) $\dfrac{3 \cdot 4}{5^2 \cdot 17^3} n$.

(iv) $\dfrac{3 \cdot 4^2}{5^2 \cdot 17^3} n$.

(v) None of the above.

(c) $F_{N_1,S}(2, 4-n) =$

(i) $0.8^{n-2}\left(0.15^2\binom{n}{2} + 0.12n + 0.64\right)$.

(ii) $0.8^{n-2}\left(0.15^2\binom{n}{2} + 0.15n + 0.64\right)$.

(iii) $0.8^{n-2}\left(0.15^2\binom{n}{2} + 0.16n + 0.64\right)$.

(iv) $0.8^{n-2}\left(0.15^2\binom{n}{2} + 0.2n + 0.64\right)$.

(v) None of the above.

(d) Suppose that $n = 1000$ and it is known that $N_5 = 100$. Then the number of times the player has won 5 shekels during the first 80 rounds is distributed as

(i) Approximately $P(4)$.

(ii) $H(80, 100, 900)$.

(iii) $B(80, 0.1)$.

(iv) $\bar{B}(100, 0.05)$.

(v) None of the above.

3. Recall Banach's matchbox problem.

(a) Let $R$ and $L$ denote the numbers of remaining matches in the right and the left pocket after the smoker has drawn $N$ matches. Then $E(RL) =$

(i) $\dfrac{N^2 - N}{6}$.

(ii) $\dfrac{N^2 - N}{4}$.

(iii) $\dfrac{N^2 - N}{3}$.

(iv) $\dfrac{N^2 - N}{2}$.

(v) None of the above.

(b) The probability that, when the smoker first draws the last match from one of the pockets, the other pocket still contains $k$ matches ($1 \leq k \leq N$), is:

(i) $\binom{2N-k-1}{N} \cdot \left(\frac{1}{2}\right)^{2N-k}$.

(ii) $\binom{2N-k-1}{N} \cdot \left(\frac{1}{2}\right)^{2N-k-1}$.

(iii) $\binom{2N-k-1}{N-1} \cdot \left(\frac{1}{2}\right)^{2N-k}$.

(iv) $\binom{2N-k-1}{N-1} \cdot \left(\frac{1}{2}\right)^{2N-k-1}$.

(v) None of the above.

(c) Consider the stage at which the smoker first draws the last match from one of the pockets. Denote by $X$ the number of additional times he tries to draw a match until he realizes that one of his pockets is empty. (For example, suppose that $N = 20$, in the first 36 steps he has drawn 19 matches from the right pocket and 17 from the left, and in the 37-th step he draws the last match from the right pocket. If now he draws two matches from the left pocket and then tries to draw from the right, $X = 3$. If, on the other hand, he now draws all three remaining matches from the left pocket, then $X = 4$ in any case.) Consider $E(X)$.

(i) $E(X) \in [0, 1]$.

(ii) $E(X) \in (1, 2)$.

(iii) $E(X) \in [2, 3]$.

(iv) $E(X) \in (3, 4)$.

(v) None of the above.

(d) Now suppose that the number of matches in the right pocket is $N+1$ at the beginning, while the number of those in the left pocket is still $N$. In addition, the matches are of low quality, so that the probability that the smoker will manage to light his cigarette with any match is only $1/2$. (Different matches are independent.) Denote by $p_N$ the probability that the number of cigarettes the smoker will light with the matches in the right pocket is strictly larger than the corresponding number for the left pocket. Then:

(i) For every sufficiently large $N$ we have $p_N < 1/2$. In the limit, $p_N \xrightarrow[N\to\infty]{} 1/2$.

(ii) For every sufficiently large $N$ we have $p_N = 1/2$.

(iii) For every sufficiently large $N$ we have $p_N > 1/2$. In the limit, $p_N \xrightarrow[N\to\infty]{} 1/2$.

(iv) For every sufficiently large $N$ we have $p_N > 1/2$. In the limit, $p_N \xrightarrow[N\to\infty]{} p$, where $p > 1/2$.

(v) None of the above.

## Solutions

1. (a) As we have seen in class, for every number $n \geq 1$ of letters and envelopes, the expected number of letters put in the matching envelope is 1. Hence
$$E(R \mid X) = 1,$$
and therefore
$$E(R) = E(E(R \mid X)) = E(1) = 1.$$
By Markov's Inequality,
$$P(R \geq 100) \leq \frac{E(R)}{100} = \frac{1}{100}.$$
Thus, (iv) is true.

(b) Given $X = 102$, the event $\{R = 100\}$ means that 100 letters are placed in the matching envelopes and the remaining two are transposed. Hence
$$P(R = 100 \mid X = 102) = \frac{\binom{102}{2}}{102!} = \frac{1}{2 \cdot 100!}.$$
Thus, (i) is true.

(c) By the law of total probability:
$$P(R = X) = \sum_{k=1}^{\infty} P(X = k) P(R = k \mid X = k).$$
Since
$$P(X = k) = \left(\frac{5}{6}\right)^{k-1} \frac{1}{6}, \qquad k = 1, 2, \ldots,$$
and
$$P(R = k \mid X = k) = \frac{1}{k!}, \qquad k = 1, 2, \ldots,$$

we obtain

$$\begin{aligned} P(R = X) &= \frac{1}{6}\sum_{k=1}^{\infty}\frac{(5/6)^{k-1}}{k!} \\ &= \frac{1}{5}\sum_{k=1}^{\infty}\frac{(5/6)^{k}}{k!} \\ &= \frac{e^{5/6}-1}{5}. \end{aligned}$$

Thus, (iii) is true.

(d) Given $X$, the random variables $R$ and $S$ are independent. Therefore

$$E(RS \mid X) = E(R \mid X)E(S \mid X) = 1.$$

Hence

$$E(RS) = E(E(RS \mid X)) = E(1) = 1.$$

Since

$$E(R) = E(S) = 1,$$

it follows that

$$\mathrm{Cov}(R, S) = E(RS) - E(R)E(S) = 1 - 1 = 0,$$

and therefore

$$\rho(R, S) = 0.$$

Thus, (ii) is true.

(e) We have

$$\begin{aligned} \mathrm{Cov}(R', S') &= \mathrm{Cov}(X - R, X - S) \\ &= V(X) - \mathrm{Cov}(X, R) - \mathrm{Cov}(X, S) + \mathrm{Cov}(R, S). \end{aligned}$$

Now

$$\begin{aligned} E(XR) &= E(E(XR \mid X)) \\ &= E(XE(R \mid X)) \\ &= E(X), \end{aligned}$$

so that

$$\operatorname{Cov}(X, R) = E(XR) - E(X)E(R) = 0.$$

Similarly,

$$\operatorname{Cov}(X, S) = 0.$$

Hence

$$\operatorname{Cov}(R', S') = V(X) = \frac{1 - 1/6}{(1/6)^2} = 30.$$

As shown in class the variance of the number of letters put in the matching envelopes is 1 for $n \geq 2$ letters and vanishes for $n = 1$. It follows that

$$\begin{aligned} V(R) &= E(V(R \mid X)) + V(E(R \mid X)) \\ &= P(X = 1) \cdot 0 + P(X \geq 2) \cdot 1 + V(1) \\ &= \frac{5}{6}. \end{aligned}$$

Thus

$$\begin{aligned} V(R') &= V(X - R) \\ &= V(X) + V(R) - 2\operatorname{Cov}(X, R) \\ &= 30 + \frac{5}{6} + 2 \cdot 0 \\ &= \frac{185}{6}. \end{aligned}$$

Similarly,

$$V(S') = \frac{185}{6}.$$

Therefore

$$\rho(R', S') = \frac{30}{185/6} = \frac{36}{37}.$$

Thus, (v) is true.

2. (a) We have

$$\begin{aligned} M_X(\log 2) &= E\left(e^{X\log 2}\right) \\ &= E\left(2^X\right) \\ &= 0.8\cdot 2^{-1} + 0.15\cdot 2 + 0.05\cdot 2^5 \\ &= 0.4 + 0.3 + 1.6 \\ &= 2.3. \end{aligned}$$

Thus, (i) is true.

(b) Given $N_1 = k$, the remaining $n-k$ trials contain only outcomes $-1$ and 5, the conditional probabilities of these outcomes being $0.8/0.85 = 16/17$ and $0.05/0.85 = 1/17$, respectively. Hence

$$N_5 \mid N_1 = k \sim B\left(n-k, \frac{1}{17}\right).$$

Therefore

$$V(N_5 \mid N_1) = \frac{1}{17}\cdot\frac{16}{17}(n-N_1) = \frac{16}{17^2}(n-N_1).$$

Thus

$$\begin{aligned} V(V(N_5 \mid N_1)) &= \left(\frac{16}{17^2}\right)^2 V(N_1) \\ &= \frac{16^2}{17^4}\cdot n\cdot 0.15\cdot 0.85 \\ &= \frac{16^2}{17^4}\cdot n\cdot\frac{3}{20}\cdot\frac{17}{20} \\ &= \frac{3\cdot 4^2}{5^2\cdot 17^3}n. \end{aligned}$$

Thus, (iv) is true.

(c) We have

$$F_{N_1,S}(2, 4-n) = P(N_1 \le 2,\ S \le 4-n).$$

Now for $S$ to be at most $4-n$, most of the rounds need to yield the player a revenue of $-1$ shekels. More precisely, if any of the rounds yields 5 shekels, then the player gets at least $5-(n-1) = 6-n$ in total. Hence the only possibilities are $k$ rounds with a yield of 1 and $n-k$ with a $-1$, for $k = 0, 1, 2$. Thus,

$$\begin{aligned} F_{N_1,S}(2, 4-n) &= P(N_1 = 0, N_5 = 0) \\ &\quad + P(N_1 = 1, N_5 = 0) \\ &\quad + P(N_1 = 2, N_5 = 0) \\ &= 0.8^n + \binom{n}{1} 0.15 \cdot 0.8^{n-1} + \binom{n}{2} 0.15^2 \cdot 0.8^{n-2} \\ &= 0.8^{n-2}\left(0.15^2\binom{n}{2} + 0.12n + 0.64\right). \end{aligned}$$

Thus, (i) is true.

(d) Given that $N_5 = 100$, the locations of the 100 occurrences of the value 5 are uniformly distributed among the 1000 trials. We count the number of those of the first 80 rounds that were among the 100 wins of 5 shekels. Therefore the number is distributed $H(80, 100, 900)$.

Thus, (ii) is true.

3. (a) Note that the number of matches remaining in each of the pockets is the same as the number of matches drawn from the other. Hence, $R \sim B(N, 1/2)$, and we have $L = N - R$. Therefore:

$$
\begin{aligned}
E(RL) &= E(R(N-R)) \\
&= NE(R) - E(R^2) \\
&= N \cdot N/2 - N \cdot 1/2 \cdot 1/2 - (N/2)^2 \\
&= \frac{N^2 - N}{4}.
\end{aligned}
$$

Thus, (ii) is true.

(b) For the right pocket to contain $k$ matches when the left one is exhausted, we need to have $N-k$ draws from the right pocket among the first $2N-k-1$ and the $(2N-k)$-th draw has to be from the left pocket. Regarding a draw from the left pocket as a success and from the right as a failure, it requires $2N-k$ tries until we get $N$ successes. The probability of this event is the probability of a $\bar{B}(N, 1/2)$-distributed random variable to assume the value $2N-k$, which is

$$\binom{2N-k-1}{N-1}\left(\frac{1}{2}\right)^{2N-k}.$$

The event in question contains also the possibility of the analogous event with the right and the left pocket interchanged, so that the total probability is

$$2 \cdot \binom{2N-k-1}{N-1}\left(\frac{1}{2}\right)^{2N-k} = \binom{2N-k-1}{N-1}\left(\frac{1}{2}\right)^{2N-k-1}.$$

Thus, (iv) is true.

(c) The smoker needs to draw at least one match to realize that one of the pockets is empty, and there is a positive probability that he will draw at least two matches before realizing that one pocket is empty. Hence $E(X) > 1$.

On the other hand, suppose that the non-empty pocket contains $k$ matches when the other becomes empty. The number of times the smoker will try to draw a match until trying the empty pocket is in principle $G(1/2)$-distributed. However, even if all the first $k$

draws are from the non-empty pocket, the $k+1$-st try will be from an empty pocket. Hence, using the law of total expectation, $E(X) < 2$.

Thus, (ii) is true.

(d) The question is equivalent to the following question, studied in class: Two players toss coins; the first tosses $n+1$ and the second $n$. What is the probability that the first will get strictly more heads than the second? We have seen that the probability is $1/2$.

Thus, (ii) is true.