Final #1

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

1. The coin-die game with n players is played as follows. At the first stage, each player tosses a coin. If at least one of the coins shows a head, then at the second stage all players who have received a head roll a die. If none of the coins shows a head, then all n players continue to roll a die at the second stage. Denote by D the number of players who move on to the second stage, and by $X_i, 1 \le i \le 6$, the number of those who roll an i at the second stage. (For example, if $n = 30$ and 17 players get a head at the first stage, and at the second stage 3 of them roll a "1", 3 roll a "2", 3 roll a "3", 3 roll a "4", 3 roll a "5", and 2 roll a "6", then $D = 17, X_1 = X_2 = X_3 = X_4 = X_5 = 3, X_6 = 2$.)

   (a) $V(D) =$

   (i) $\frac{n}{4} - \frac{n^2}{2^n}$.
   (ii) $\frac{n}{4} - \frac{n^2}{4^n}$.
   (iii) $\frac{n}{4} + \frac{n^2}{4^n}$.
   (iv) $\frac{n}{4} + \frac{n^2}{2^n}$.
   (v) None of the above.

   (b) For $1 \le i \le 6$ we have $P(X_i = 0) =$

   (i) $\left(\frac{11}{12}\right)^n - \left(\frac{1}{2}\right)^n + \left(\frac{5}{12}\right)^n$.
   (ii) $\left(\frac{11}{12}\right)^n - \left(\frac{1}{12}\right)^n$.
   (iii) $\left(\frac{7}{12}\right)^n - \left(\frac{1}{4}\right)^n$.
   (iv) $2 \cdot \left(\frac{11}{12}\right)^n - \left(\frac{1}{2}\right)^{n-1}$.
   (v) None of the above.

(c) $P(X_6 = D) =$

(i) $\left(\frac{5}{12}\right)^n - \left(\frac{1}{3}\right)^n + \left(\frac{1}{4}\right)^n$.

(ii) $\left(\frac{1}{2}\right)^n - \left(\frac{5}{12}\right)^n + \left(\frac{1}{12}\right)^n$.

(iii) $\left(\frac{1}{3}\right)^n - \left(\frac{1}{6}\right)^n$.

(iv) $\left(\frac{7}{12}\right)^n - \left(\frac{1}{2}\right)^n + \left(\frac{1}{12}\right)^n$.

(v) None of the above.

(d) $E(X_i) =$

(i) $\frac{n}{12} + \frac{n}{12^n}$.

(ii) $\frac{n}{12} + \frac{n}{2 \cdot 6^n}$.

(iii) $\frac{n}{12} + \frac{n}{6 \cdot 2^n}$.

(iv) $\frac{n}{12} + \frac{5n}{6 \cdot 2^n}$.

(v) None of the above.

(e) Suppose we repeat the game 2^n times. Let W be the number of times (out of 2^n) in which all n players move on to the second stage and roll a "6" (namely, the number of times $X_6 = n$). For sufficiently large n we have $P(W = 2) \in$

(i) $[0, 0.01]$.

(ii) $(0.01, 0.05]$.

(iii) $(0.05, 0.1]$.

(iv) $(0.1, 0.5]$.

(v) None of the above.

2. The number of the first generation descendants (children) of a bacterium of type “Probabilitas fortuitus” is distributed $P(2)$. Denote by S_i the total number of children of i random bacteria.

   (a) The probability for a randomly chosen bacterium to have 3 descendants in the second generation (i.e., “grandchildren”) is:

      (i) $\frac{4}{3e^4}e^{2/e^2}\left(1+4/e^2+4/e^4\right)$.

      (ii) $\frac{8}{3e^4}e^{2/e^2}\left(1+6/e^2+4/e^4\right)$.

      (iii) $\frac{12}{3e^4}e^{4/e^4}\left(1+8/e^2+4/e^4\right)$.

      (iv) $\frac{16}{3e^4}e^{4/e^2}\left(1+10/e^2+4/e^4\right)$.

      (v) None of the above.

   (b) It is known that the number of the second generation descendants (“grandchildren”) of a randomly chosen bacterium is 2. Then the probability for the same bacterium to have 3 first generation descendants (“children”) is:

      (i) $\frac{6}{e^2(2+e^2)}e^{+2/e^2}$.

      (ii) $\frac{4}{e^2(2+e^2)}e^{-2/e^2}$.

      (iii) $\frac{6}{e^2(2+e^2)}e^{-2/e^2}$.

      (iv) $\frac{6}{e^2(4+e^2)}e^{-2/e^2}$.

      (v) None of the above.

(c) By Stirling's Formula $\left(n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\right)$, we have $P(S_{100} = 200) \approx$

(i) $\frac{1}{2^{10}e\sqrt{2\pi}}$.

(ii) $\frac{e^2}{2^{10}\sqrt{\pi}}$.

(iii) $\frac{1}{20\sqrt{\pi}}$.

(iv) $\frac{e}{10\sqrt{2\pi}}$.

(v) None of the above.

(d) Let $\psi_{S_{100}}(t)$ be the moment generating function of S_{100}. Then $\psi_{S_{100}}(0.0001) \approx$

(i) 1.002.

(ii) 1.004.

(iii) 1.01.

(iv) 1.02.

(v) None of the above.

(e) A direct application of Chebyshev's inequality to S_{2000} of first generation descendants of 2000 random bacteria yields $P(3800 \leq S_{2000} \leq 4200) \geq$

(i) 0.65.

(ii) 0.71.

(iii) 0.84.

(iv) 0.90.

(v) None of the above.

3. The weight X and height Y of a randomly chosen BGU student are distributed with joint 2-dimensional density function

$$f_{XY}(x,y) = \begin{cases} C(340 - x - y), & 50 \le x \le 90,\ 160 \le y \le 180, \\ 0, & \text{otherwise,} \end{cases}$$

where C is a constant.

(<u>Hint:</u> To avoid huge numbers during the calculation of the integrals, it is worthwhile introducing new integration variables t and s, defined by

$$t = \frac{x - 70}{10}, \quad s = \frac{y - 170}{10},$$

and return to the original variables (x, y) (if necessary) only at the final stage.)

(a) The normalization constant C is:

(i) $\dfrac{3}{9 \cdot 10^4}$.

(ii) $\dfrac{2}{7 \cdot 10^3}$.

(iii) $\dfrac{1}{8 \cdot 10^4}$.

(iv) $\dfrac{4}{11 \cdot 10^3}$.

(v) None of the above.

(b) Let f_X and f_Y be the marginal density functions of X and Y, respectively. Then for every x, y:

(i) $f_{XY}(x,y) = f_X(x) + f_Y(y)$.

(ii) $f_{XY}(x,y) = -f_X(x) + f_Y(y)$.

(iii) $f_{XY}(x, y) = f_X(x) \cdot f_Y(y)$.

(iv) $f_{XY}(x, y) = f_X(x) - f_Y(y)$.

(v) None of the above.

(c) $P(X + 2Y \leq 410) =$

(i) 0.28.

(ii) 0.55.

(iii) 0.63.

(iv) 0.76.

(v)None of the above.

(d) $P(X > 70 | Y > 170) \approx$

(i) 0.1612.

(ii) 0.3296.

(iii) 0.4474.

(iv) 0.6961.

(v) None of the above.

(e) $F_X(70) =$

(i) 0.44.

(ii) 0.55.

(iii) approximately 0.66.

(iv) approximately 0.77.

(v) None of the above.

Solutions

1. (a) For $1 \leq k \leq n$, let $Y_k = 1$ if the coin of k-th player shows a head and 0 otherwise. Thus $Y_k \sim B(1, 1/2)$. Due to the independence of the variables Y_k, the variable $Y = \sum_{k=1}^{n} Y_k$ is distributed $B(n, 1/2)$. As a result

$$P_Y(m) = \binom{n}{m} \frac{1}{2^n}, \qquad m = 0, 1, ...n,$$

and

$$E(Y) = \frac{n}{2}, \qquad V(Y) = \frac{n}{4}.$$

The random variable D may be expressed as a sum, $D = Y + Y_0$, where

$$Y_0 = \begin{cases} 0, & Y \neq 0, \\ n, & Y = 0. \end{cases}$$

Therefore

$$P_{Y_0}(s) = \begin{cases} 1 - \frac{1}{2^n}, & s = 0, \\ \frac{1}{2^n}, & s = n, \end{cases}$$

and

$$E(Y_0) = \frac{n}{2^n}, \qquad V(Y_0) = \frac{n^2}{2^n} - \frac{n^2}{4^n}.$$

Obviously, $E(Y \cdot Y_0) = 0$, so that $\text{Cov}(Y, Y_0) = -\frac{n^2}{2^{n+1}}$. Hence:

$$V(D) = V(Y + Y_0) = V(Y) + V(Y_0) + 2\text{Cov}(Y, Y_0) = \frac{n}{4} - \frac{n^2}{4^n}.$$

Thus, (ii) is true.

(b) By the law of total probability

$$P_{X_i}(0) = \sum_{m=1}^{n} P_D(m) P(X_i = 0 | D = m).$$

According to the definition, the probability function of D is given by:

$$P_D(m) = \begin{cases} \binom{n}{m} \frac{1}{2^n}, & 1 \leq m \leq n-1, \\ \\ \frac{1}{2^{n-1}}, & m = n. \end{cases} \tag{1}$$

Therefore

$$\begin{aligned} P_{X_i}(0) &= P_D(n) P_{X_i|D}(0|n) + \sum_{m=1}^{n-1} P_D(m) P_{X_i|D}(0|m) \\ &= \frac{1}{2^{n-1}} \left(\frac{5}{6}\right)^n + \sum_{m=1}^{n-1} \binom{n}{m} \left(\frac{1}{2}\right)^n \left(\frac{5}{6}\right)^m \\ &= 2\left(\frac{5}{12}\right)^n + \sum_{m=1}^{n-1} \binom{n}{m} \left(\frac{5}{12}\right)^m \left(\frac{1}{2}\right)^{n-m} \\ &= 2\left(\frac{5}{12}\right)^n + \sum_{m=0}^{n} \binom{n}{m} \left(\frac{5}{12}\right)^m \left(\frac{1}{2}\right)^{n-m} - \frac{1}{2^n} - \left(\frac{5}{12}\right)^n \\ &= \left(\frac{5}{12}\right)^n + \left(\frac{11}{12}\right)^n - \frac{1}{2^n}. \end{aligned}$$

Thus, (i) is true.

(c) This part is similar to the preceding one.

$$\begin{aligned} P(X_6 = D) &= P_D(n)P_{X_6|D}(n|n) + \sum_{m=1}^{n-1} P_D(m)P_{X_6|D}(m|m) \\ &= \frac{1}{2^{n-1}}\left(\frac{1}{6}\right)^n + \sum_{m=1}^{n-1}\binom{n}{m}\left(\frac{1}{2}\right)^n\left(\frac{1}{6}\right)^m \\ &= \frac{1}{2^n}\left(\frac{1}{6}\right)^n + \sum_{m=0}^{n}\binom{n}{m}\left(\frac{1}{2}\right)^n\left(\frac{1}{6}\right)^m - \left(\frac{1}{2}\right)^n \\ &= \left(\frac{1}{12}\right)^n + \left(\frac{7}{12}\right)^n - \left(\frac{1}{2}\right)^n. \end{aligned}$$

Thus, (iv) is true.

(d) Obviously, $D = \sum_{i=1}^{6} X_i$, so that $E(D) = \sum_{i=1}^{6} E(X_i)$. By symmetry, all X_i's are identically distributed, and therefore

$$E(X_i) = \frac{1}{6}E(D)$$

for each i. The expectation $E(D)$ can be found either from the solution of part (a),

$$E(D) = E(Y) + E(Y_0) = \frac{n}{2^n} + \frac{n}{2},$$

or by a direct calculation using (1)

$$\begin{aligned} E(D) &= n\frac{1}{2^{n-1}} + \sum_{m=1}^{n-1} m\binom{n}{m}\frac{1}{2^n} \\ &= \frac{n}{2^n} + \sum_{m=1}^{n} m\binom{n}{m}\frac{1}{2^n} \\ &= \frac{n}{2^n} + \frac{n}{2^n}\sum_{m=1}^{n}\binom{n-1}{m-1} \\ &= \frac{n}{2^n} + \frac{n}{2}. \end{aligned}$$

As a result we get

$$E(X_i) = \frac{n}{6 \cdot 2^n} + \frac{n}{12},$$

Thus, (iii) is true.

(e) We have:

$$P(X_6 = n) = P_D(n) \cdot P(X_6 = n|D = n) = \frac{1}{2^{n-1}} \cdot \frac{1}{6^n} = \frac{2}{12^n}.$$

It follows that $W \sim B\left(2^n, 2/12^n\right)$, and in particular

$$P_W(2) = \binom{2^n}{2} \left(\frac{2}{12^n}\right)^2 \cdot \left(1 - \frac{2}{12^n}\right)^{2^n - 2} \le \frac{2^{2n}}{2} \cdot \frac{2^2}{12^{2n}} = \frac{2}{6^{2n}}.$$

The right-hand side decreases monotonically and is less than 0.01 already for $n = 2$.
Thus, (i) is true.

2. Let X_i be the number of descendants of a random bacterium in the i-th generation, and set $P_i(n) \equiv P(X_i = n)$. According to the formulation of the problem, $X_1 \sim P(2)$, so that

$$P_1(n) = e^{-2} \cdot \frac{2^n}{n!}.$$

(a) The required probability is

$$\begin{aligned} P_2(3) &= \sum_{n=1}^{\infty} P_1(n) P(X_2 = 3|X_1 = n) \\ &= \sum_{n=1}^{\infty} P_1(n) \left(\binom{n}{1} P_1(3) P_1^{n-1}(0) + 2\binom{n}{2} P_1(2) P_1(1) P_1^{n-2}(0) \right. \\ &\qquad \left. + \binom{n}{3} P_1^3(1) P_1^{n-3}(0) \right) \\ &= \frac{8}{3e^4}\left(1 + \frac{6}{e^2} + \frac{4}{e^4}\right) e^{2/e^2}. \end{aligned}$$

Thus, (ii) is true.

(b) According to definition of conditional probability:

$$\begin{aligned} P(X_1 = 3 | X_2 = 2) &= \frac{P(X_1 = 3, X_2 = 2)}{P(X_2 = 2)} \\ &= \frac{P(X_1 = 3) P(X_2 = 2 | X_1 = 3)}{P(X_2 = 2)}. \end{aligned}$$

Now

$$P(X_2 = 2 | X_1 = 3) = \binom{3}{1} P_1(2) P_1^2(0) + \binom{3}{2} P_1^2(1) P_1(0) = \frac{18}{e^6},$$

and

$$\begin{aligned} P(X_2 = 2) &= \sum_{n=1}^{\infty} P(X_1 = n) P(X_2 = 2 | X_1 = n) \\ &= \sum_{n=1}^{\infty} P_1(n) \left(\binom{n}{1} P_1(2) P_1^{n-1}(0) + \binom{n}{2} P_1^2(1) P_1^{n-2}(0) \right) \\ &= \frac{4}{e^4} \left(1 + \frac{2}{e^2} \right) e^{2/e^2}. \end{aligned}$$

Finally,

$$P(X_1 = 3 | X_2 = 2) = \frac{6}{e^2(2 + e^2)} e^{-2/e^2}.$$

Thus, (iii) is true.

(c) Enumerate the bacteria by an index k, running from 1 to 100, and let Y_k be the number of first generation descendants of bacterium number k. Then $S_{100} = \sum_{k=1}^{100} Y_k$ is a sum of 100 independent $P(2)$-distributed variables, and therefore $S_{100} \sim P(100 \cdot 2)$. Hence:

$$P(S_{100} = 200) = e^{-200} \frac{(200)^{200}}{200!}.$$

Applying Stirling's formula we obtain

$$P(S_{100} = 200) \approx \frac{1}{20\sqrt{\pi}}.$$

Thus, (iii) is true.

(d) Since $S_{100} \sim P(200)$ we have $\psi_{S_{100}}(t) = e^{200\left(e^t-1\right)}$. For $t \approx 0$ we obtain

$$\psi_{S_{100}}(t) = e^{200(1+t-1)+o(t)} = 1 + 200t + o(t).$$

Alternatively, the same result may be obtained from the expansion of the moment generating function in general:

$$\psi_{S_{100}}(t) = 1+\psi'_{S_{100}}(0)\cdot t+o(t) = 1+E(S_{100})\cdot t+o(t) = 1+200t+o(t).$$

Substituting $t = 0.0001$ we obtain

$$\psi_{S_{100}}(0.0001) \approx 1 + 200 \cdot 0.0001 = 1.02.$$

Thus, (iv) is true.

(e) Similarly to the preceding parts, $S_{2000} \sim P(4000)$. Therefore $E(S_{2000}) = 4000$, and the event we are interested in may be written as

$$\left|\frac{S_{2000} - E(S_{2000})}{2000}\right| \leq 0.1.$$

Applying Chebyshev's inequality to the random variable $S_{2000}/2000$, we obtain:

$$P\left(\left|\frac{S_{2000}}{2000} - \frac{E(S_{2000})}{2000}\right| \geq 0.1\right) \leq \frac{V(S_{2000}/2000)}{0.1^2} = 0.1.$$

Therefore

$$P\left(\left|\frac{S_{2000}}{2000} - \frac{E(S_{2000})}{2000}\right| \leq 0.1\right) \geq 0.9$$

Thus, (iv) is true.

3. The expectation of a function $g(X,Y)$ of X and Y is:

$$E(g(X,Y)) = \int_{50}^{90} \int_{160}^{180} f_{XY}(x,y)\ g(x,y)dydx.$$

In terms of the variables

$$t = \frac{x-70}{10}, \quad s = \frac{y-170}{10}$$

we have

$$E(g(X,Y)) = C \cdot 10^3 \int_{-2}^{2} \int_{-1}^{1} (10-t-s) g(70+10t, 170+10s)\, ds dt. \quad (2)$$

(a) Choosing $g(X,Y)=1$ in (2) we get the normalization constant

$$C^{-1} = 10^3 \int_{-2}^{2} \int_{-1}^{1} (10-t-s)\, ds dt = 8 \cdot 10^4.$$

Thus, (iii) is true.

(b) The calculation of the two marginal density functions $f_X(x) = \int_{160}^{180} f_{XY}(x,y) dy$ and $f_Y(y) = \int_{50}^{90} f_{XY}(x,y) dx$ leads to the results

$$f_X(x) = \begin{cases} \frac{1}{4 \cdot 10^3}(170-x), & 50 \le x \le 90, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$f_Y(y) = \begin{cases} \frac{1}{2 \cdot 10^3}(270-y), & 160 \le y \le 180, \\ 0, & \text{otherwise}. \end{cases}$$

Thus, (v) is true.

(c) We have:

$$\begin{aligned} P(2Y + X \le 410) &= \iint_{2Y+X \le 410} f_{XY}(x,y)\, dx dy \\ &= \int_{160}^{180} \int_{50}^{410-2y} \frac{1}{8 \cdot 10^3}(340-x-y)\, dx dy \\ &= \frac{1}{80} \int_{-1}^{1} \int_{-2}^{-2s} (10-t-s)\, dt ds = 0.55. \end{aligned}$$

Thus, (ii) is true.

(d)

$$P(X > 70|Y > 170) = \frac{P(X > 70, Y > 170)}{P(Y > 170)}.$$

Now

$$P(Y > 170) = \int_{170}^{180} f_Y(y)dy = \int_{170}^{180} \frac{1}{2 \cdot 10^3}(270 - y)dy = \frac{19}{40},$$

and

$$\begin{aligned} P(X > 70, Y > 170) &= \frac{1}{8 \cdot 10^4} \int_{170}^{180} \int_{70}^{90} (340 - x - y)dxdy \\ &= \frac{1}{80} \int_0^1 \int_0^2 (10 - s - t)dtds = \frac{17}{80}. \end{aligned}$$

Consequently:

$$P(X > 70|Y > 170) = \frac{17}{38} \approx 0.4474.$$

Thus, (iii) is true.

(e) The value of the distribution function of the weight at any point $x \in [50, 90]$ is

$$\begin{aligned} F_X(x) &= \int_{50}^{x} f_X(u)du \\ &= \int_{50}^{x} \frac{1}{4 \cdot 10^3}(170 - u)du \\ &= \frac{1}{4 \cdot 10^3} \left(-\frac{1}{2}x^2 + 170x - 7250 \right). \end{aligned}$$

For $x = 70$ we get $F_X(70) = 0.55$.
Thus, (ii) is true.