1 Exam A (Friday 3 PM)

1.1 Problem 1

1. (20 points) You have an urn that initially contains 6 red balls, 2 black balls and 1 green ball. On the first step, you choose one ball uniformly at random from the urn, look at its color, and then return it back to the urn together with one more ball of the same color (e.g., if you pick a red ball, then you put it back to the urn together with another red ball). Then on the second step you choose a ball uniformly at random from the urn (note that on the second step the urn contains the additional ball).

   What is the probability that on the second step you choose a red ball?

   Solution. Denote

   $$A = \{\text{ball chosen on the first step is red}\}, \quad B = \{\text{ball chosen on the second step is red}\}. \tag{1}$$

   Then using the law of total probability

   $$P(B) = P(B|A)P(A) + P(B|\overline{A})P(\overline{A}) = \frac{7}{10} \cdot \frac{6}{9} + \frac{6}{10} \cdot \frac{3}{9} = \frac{2}{3}. \tag{2}$$

2. (20 points) You have an urn that initially contains 5 red balls, 2 black balls and 2 green ball. On the first step, you choose one ball uniformly at random from the urn, look at its color, and then return it back to the urn together with one more ball of the same color (e.g., if you pick a red ball, then you put it back to the urn together with another red ball). Then on the second step you choose a ball uniformly at random from the urn (note that on the second step the urn contains the additional ball).

   What is the probability that on the second step you choose a green ball?

   Solution. Denote

   $$A = \{\text{ball chosen on the first step is green}\}, \quad B = \{\text{ball chosen on the second step is green}\}. \tag{3}$$

   Then using the law of total probability

   $$P(B) = P(B|A)P(A) + P(B|\overline{A})P(\overline{A}) = \frac{3}{10} \cdot \frac{2}{9} + \frac{2}{10} \cdot \frac{7}{9} = \frac{2}{9}. \tag{4}$$

1.2 Problem 2

3. (20 points) Let X and Y be independent random variables uniformly distributed on the interval $[0,1]$, i.e., $X \sim \mathcal{U}[0,1]$, $Y \sim \mathcal{U}[0,1]$.
   (a) (8 points) Compute the moment generating function of the sum $X+Y$.
   (b) (4 points) Show that for any $t \in \mathbb{R}$
   $$(e^t-1)^2 = e^{2t} - 2e^t + 1 = \sum_{k=2}^{\infty} \frac{2^k}{k!} t^k - \sum_{k=2}^{\infty} \frac{2}{k!} t^k. \tag{5}$$
   (c) (8 points) Use the results of (a) and (b) to compute $E((X+Y)^n)$, moments of the sum, for any $n \in \mathbb{N}$.

Solution.

(a) Compute the moment generating function: for any $t \in \mathbb{R}$
$$M_{X+Y}(t) = E(e^{t(X+Y)}) = E(e^{tX})E(e^{tY}) = \int_0^1 e^{tx}dx \int_0^1 e^{ty}dy = \frac{(e^t-1)^2}{t^2} \tag{6}$$

(b) The first equality is the expansion of the square of a difference, and the second equality follows from the power series expansion of the exponential function
$$e^{2t} = 1 + 2t + \sum_{k=2}^{\infty} \frac{2^k}{k!} t^k, \quad 2e^t = 2 + 2t + \sum_{k=2}^{\infty} \frac{2}{k!} t^k, \tag{7}$$
$$e^{2t} - 2e^t + 1 = (1 + 2t - 2 - 2t + 1) + \sum_{k=2}^{\infty} \frac{2^k}{k!} t^k - \sum_{k=2}^{\infty} \frac{2}{k!} t^k. \tag{8}$$

(c) From (a) and (b) we have that
$$M_{X+Y}(t) = \frac{\sum_{k=2}^{\infty} \frac{2^k-2}{k!} t^k}{t^2} = \sum_{k=2}^{\infty} \frac{2^k-2}{k!} t^{k-2}, \tag{9}$$
by changing $n = k-2$
$$\sum_{k=2}^{\infty} \frac{2^k-2}{k!} t^{k-2} = \sum_{n=0}^{\infty} \frac{2^{n+2}-2}{(n+1)(n+2)} \cdot \frac{t^n}{n!} \tag{10}$$
and we conclude that
$$E((X+Y)^n) = \frac{2^{n+2}-2}{(n+1)(n+2)}. \tag{11}$$

1.3 Problem 3

4. (20 points) Let X and Y be a pair of jointly continuous random variables with joint density

$$f_{X,Y}(x,y) = \begin{cases} c(x+3y), & 0 \le x, y \le 1, \\ 0, & \text{otherwise,} \end{cases} \tag{12}$$

with an unknown parameter $c > 0$.

(a) (5 points) Determine the value of $c > 0$.

(b) (10 points) Compute the marginal densities of X and Y.

(c) (5 points) Determine if random variables X and Y are independent.

Solution.

(a) We have that $c = 1/2$:

$$\int_0^1 \int_0^1 f_{X,Y}(x,y)dxdy = c\int_0^1 \int_0^1 (x+3y)dxdy = c\Big(\frac{1}{2} + \frac{3}{2}\Big) = 2c = 1. \tag{13}$$

(b)

$$f_X(x) = \int_0^1 \frac{1}{2}(x+3y)dy = \frac{x}{2} + \frac{3}{4}, \quad f_Y(y) = \int_0^1 \frac{1}{2}(x+3y)dx = \frac{1}{4} + \frac{3y}{2}. \tag{14}$$

(c) Since $f_X(x)f_Y(y) \neq f_{X,Y}(x,y)$, random variables X and Y are not independent.

5. (20 points) Let X and Y be a pair of jointly continuous random variables with joint density

$$f_{X,Y}(x,y) = \begin{cases} c(5x+y), & 0 \le x, y \le 1, \\ 0, & \text{otherwise,} \end{cases} \tag{15}$$

with an unknown parameter $c > 0$.

(a) (5 points) Determine the value of $c > 0$.

(b) (10 points) Compute the marginal densities of X and Y.

(c) (5 points) Determine if random variables X and Y are independent.

Solution.

(a) We have that $c = 1/3$:

$$\int_0^1 \int_0^1 f_{X,Y}(x,y)dxdy = c\int_0^1 \int_0^1 (5x+y)dxdy = c\Big(\frac{5}{2} + \frac{1}{2}\Big) = 3c = 1. \tag{16}$$

(b)

$$f_X(x) = \int_0^1 \frac{1}{3}(5x+y)dy = \frac{5x}{3} + \frac{1}{6}, \quad f_Y(y) = \int_0^1 \frac{1}{3}(5x+y)dx = \frac{5}{6} + \frac{y}{3}. \qquad (17)$$

(c) Since $f_X(x)f_Y(y) \neq f_{X,Y}(x,y)$, random variables X and Y are not independent.

1.4 Problem 4

6. (20 points) Let X and Y be independent random variables. Suppose that X is Gaussian, $X \sim N(1,1)$ and that Y has Poisson distribution $Y \sim \text{Pois}(2)$.
   (a) (10 points) Compute $E(3X-2Y-5)$ and $\text{Var}(3X-2Y-5)$.
   (b) (10 points) Compute $\text{Var}(XY)$.

Solution.

(a) $E(X) = 1$, $E(Y) = 2$, so by the linearity of the expectation

$$E(3X-2Y-5) = 3 \cdot 1 - 2 \cdot 2 - 5 = -6. \qquad (18)$$

$\text{Var}(X) = 1$,$\text{Var}(Y) = 2$, X and Y are independent, therefore

$$\text{Var}(3X-2Y-5) = 9 \cdot 1 + 4 \cdot 2 = 17. \qquad (19)$$

(b)

$$\text{Var}(XY) = E(X^2Y^2) - (E(X)E(Y))^2 = E(X^2)E(Y^2) - 4, \qquad (20)$$

$$E(X^2) = \text{Var}(X) + (E(X))^2 = 2, \quad E(Y^2) = \text{Var}(Y) + (E(Y))^2 = 6, \qquad (21)$$

so $\text{Var}(XY) = 2 \cdot 6 - 4 = 8$.

7. (20 points) Let X and Y be independent random variables. Suppose that X has Poisson distribution, $X \sim \text{Pois}(2)$ and that Y is exponential, $Y \sim \text{Exp}(1)$.
   (a) (10 points) Compute $E(3X-6Y-1)$ and $\text{Var}(3X-6Y-1)$.
   (b) (10 points) Compute $\text{Var}(XY)$.

Solution.

(a) $E(X) = 2$, $E(Y) = 1$, so by the linearity of the expectation

$$E(3X-6Y-1) = 3 \cdot 2 - 6 \cdot 1 - 1 = -1. \qquad (22)$$

$\text{Var}(X) = 2$,$\text{Var}(Y) = 1$, X and Y are independent, therefore

$$\text{Var}(3X-6Y-5) = 9 \cdot 2 + 36 \cdot 1 = 54. \qquad (23)$$

(b)

$$\mathrm{Var}(XY) = E(X^2Y^2) - (E(X)E(Y))^2 = E(X^2)E(Y^2) - 4, \tag{24}$$

$$E(X^2) = \mathrm{Var}(X) + (E(X))^2 = 6, \quad E(Y^2) = \mathrm{Var}(Y) + (E(Y))^2 = 2, \tag{25}$$

so $\mathrm{Var}(XY) = 2 \cdot 6 - 4 = 8$.

1.5 Problem 5

8. (20 points) Let $(\xi_i)_{i=1}^{\infty}$ be a sequence of independent identically distributed random variables each having Bernoulli distribution with parameter $1/6$, $\xi \sim \mathrm{Ber}(1/6)$, and let $S_n := \xi_1 + \cdots + \xi_n$.
   (a) (5 points) Compute $E(4^{\xi_i})$.
   (b) (5 points) Compute $E(4^{S_n})$.
   (c) (10 points) Using the result of part (b) and Markov's inequality, show that

$$P\Big(\frac{S_n}{n} \geq \frac{1}{2}\Big) \leq \Big(\frac{3}{4}\Big)^n. \tag{26}$$

Solution.

(a)

$$E(4^{\xi_i}) = 4^0 \cdot \frac{5}{6} + 4 \cdot \frac{1}{6} = \frac{3}{2}. \tag{27}$$

(b)

$$E(4^{S_n}) = E(4^{\xi_1+\cdots+\xi_n}) = E(4^{\xi_1}) \cdots E(4^{\xi_n}) = \Big(\frac{3}{2}\Big)^n, \tag{28}$$

where on the second step we used the independence of $\{\xi_i\}$, and on the third step we used the result of part (a).

(c)

$$P\Big(\frac{S_n}{n} \geq \frac{1}{2}\Big) = P\Big(S_n \geq \frac{n}{2}\Big) = P\Big(4^{S_n} \geq 4^{\frac{n}{2}}\Big) \leq \frac{E(4^{S_n})}{2^n} = \Big(\frac{3}{4}\Big)^n, \tag{29}$$

where on the third step we used Markov's inequality, and on the last step we used part (b).

2 Exam B (Thursday 9:30 PM)

2.1 Problem 1

9. (20 points) You have two urns. The first urn has 3 red balls, 2 blue balls and 2 green balls. The second urn has 2 red balls and 4 blue balls. You choose one of the urns at random (with equal probability), and then sample one ball from that urn. The ball that you picked is blue.

   What is the probability that the ball was picked from the first urn?

   Solution. Denote by B_i the event that the ith urn has been chosen, and by A the event that you have picked a blue ball. Then, using the Bayes' rule

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2)} = \frac{\frac{2}{7} \cdot \frac{1}{2}}{\frac{2}{7} \cdot \frac{1}{2} + \frac{4}{6} \cdot \frac{1}{2}} = \frac{3}{10}. \tag{30}$$

10. (20 points) You have three urns. The first urn has 3 red balls, 2 blue balls and 3 green balls. The second urn has 2 red balls and 4 blue balls. The third urn has 6 green balls. You choose one of the urns at random (with equal probability), and then sample one ball from that urn. The ball that you picked is green.

    What is the probability that the ball was picked from the first urn?

    Solution. Denote by B_i the event that the ith urn has been chosen, and by A the event that you have picked a green ball. Then, using the Bayes' rule

$$P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)} \tag{31}$$

$$= \frac{\frac{3}{8} \cdot \frac{1}{3}}{\frac{3}{8} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{3}{11}. \tag{32}$$

2.2 Problem 2

11. (20 points) Let $X \sim N(0, 1)$ be a random variable with standard normal distribution, and let $Y = X^2$.

    Compute the moment generating function of Y and use it to compute $E(Y)$ and $E(Y^2)$.

Solution.

$$M_Y(t) = E(e^{tY}) = E(e^{tX^2}) = \int_{-\infty}^{+\infty} e^{tx^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}(1-2t)} dx, \tag{33}$$

therefore $M_Y(t) = \infty$ for $t \geq 1/2$.

For $t < 1/2$, apply the following change of variable $y = \sqrt{1-2t}\, x$, so that

$$M_Y(t) = \frac{1}{\sqrt{1-2t}} \frac{1}{\cdot\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy = (1-2t)^{-1/2}, \tag{34}$$

where on the last step we used the formula for the Gaussian integral. Now

$$E(Y) = M_Y'(0) = (1-2t)^{-3/2}\Big|_{t=0} = 1. \tag{35}$$

$$E(Y^2) = M_Y''(0) = 3(1-2t)^{-5/2}\Big|_{t=0} = 3. \tag{36}$$

2.3 Problem 3

12. (20 points) Let X and Y be independent random variables with (marginal) densities

$$f_X(x) = \begin{cases} \frac{1}{2}, & x \in (-1,1), \\ 0, & \text{otherwise}, \end{cases} \tag{37}$$

$$f_Y(y) = \begin{cases} 2y, & y \in (0,1), \\ 0, & \text{otherwise}. \end{cases} \tag{38}$$

(a) (5 points) Compute the variance of the product $\mathrm{Var}(XY)$.

(b) (5 points) Compute the probability $P(X > Y)$.

Solution.

(a)

$$\mathrm{Var}(XY) = E((XY)^2) - (E(XY))^2 = E(X^2)E(Y^2) - (E(X))^2(E(Y))^2. \tag{39}$$

Since

$$E(X) = \int_{-1}^{1} \frac{dx}{2} = 0, \tag{40}$$

we have that $\mathrm{Var}(XY) = E(X^2)E(Y^2)$. Now compute

$$E(X^2) = \int_{-1}^{1} \frac{x^2}{2} dx = \frac{1}{3}, \quad E(Y^2) = \int_{0}^{1} y^2 \cdot 2y dy = \frac{1}{2}. \tag{41}$$

We conclude that

$$\mathrm{Var}(XY) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}. \tag{42}$$

(b) Since X and Y are independent, their joint density is given by

$$f_{X,Y}(x,y) = f_X(x)f_Y(y), \tag{43}$$

thus

$$P(X > Y) = \int_0^1 \int_0^x y\,dy\,dx = \int_0^1 \frac{x^2}{2}dx = \frac{1}{6}. \tag{44}$$

13. (20 points) Let X and Y be independent random variables with (marginal) densities

$$f_X(x) = \begin{cases} \frac{1}{2}, & x \in (-1,1), \\ 0, & \text{otherwise}, \end{cases} \tag{45}$$

$$f_Y(y) = \begin{cases} 3y^2, & y \in (0,1), \\ 0, & \text{otherwise}. \end{cases} \tag{46}$$

(a) (5 points) Compute the variance of the product $\mathrm{Var}(XY)$.

(b) (5 points) Compute the probability $P(X > Y)$.

Solution.

(a)

$$\mathrm{Var}(XY) = E((XY)^2) - (E(XY))^2 = E(X^2)E(Y^2) - (E(X))^2(E(Y))^2. \tag{47}$$

Since

$$E(X) = \int_{-1}^1 \frac{dx}{2} = 0, \tag{48}$$

we have that $\mathrm{Var}(XY) = E(X^2)E(Y^2)$. Now compute

$$E(X^2) = \int_{-1}^1 \frac{x^2}{2}dx = \frac{1}{3}, \quad E(Y^2) = \int_0^1 y^2 \cdot 3y^2 dy = \frac{3}{5}. \tag{49}$$

We conclude that

$$\mathrm{Var}(XY) = \frac{1}{3} \cdot \frac{3}{5} = \frac{1}{5}. \tag{50}$$

(b) Since X and Y are independent, their joint density is given by

$$f_{X,Y}(x,y) = f_X(x)f_Y(y), \tag{51}$$

thus

$$P(X > Y) = \int_0^1 \int_0^x \frac{3y^2}{2}dy\,dx = \int_0^1 \frac{x^3}{2}dx = \frac{1}{8}. \tag{52}$$

2.4 Problem 4

14. (20 points) You roll a fair die n times. Each time when the number you get is different from the number obtained on the previous roll, you win 2 dollars. For example, the sequence $(1, 2, 1, 3, 3, 5)$ results in winning 8 dollars (2 dollars on 2nd, 3rd, 4th and 6th rolls).

    Compute your expected winnings after n rolls.

    Solution. Let W_n denote your winnins after n rolls. Denote by A_i the event that ith roll gives different outcome than the $(n+1)$th roll. Then

$$W_n = \sum_{i=1}^{n-1} 2 \cdot \mathbf{1}_{A_i}. \tag{53}$$

From the linearity of the expectation,

$$E(W_n) = \sum_{i=1}^{n-1} 2E(\mathbf{1}_{A_i}) = \sum_{i=1}^{n-1} 2P(A_i). \tag{54}$$

For any $i \in \{1, \dots, n-1\}$

$$P(A_i) = \frac{5}{6}, \tag{55}$$

therefore we conclude that the expected winning after the nth game are equal to

$$E(W_n) = 2(n-1)\frac{5}{6} = \frac{5}{3}(n-1) \tag{56}$$

15. (20 points) You roll a fair die n times. Each time when the number you get is the same as the number obtained on the previous roll, you win 3 dollars. For example, the sequence $(2, 2, 1, 3, 3, 2)$ results in winning 6 dollars (3 dollars on 2nd and 5th rolls).

    Question: Compute your expected winnings after n rolls.

    Solution. Let W_n denote your winnins after n rolls. Denote by B_i the event that ith roll gives the same outcome the the $(i+1)$th roll. Then

$$W_n = \sum_{i=1}^{n-1} 3 \cdot \mathbf{1}_{B_i}. \tag{57}$$

From the linearity of the expectation,

$$E(W_n) = \sum_{i=1}^{n-1} 3E(\mathbf{1}_{B_i}) = \sum_{i=1}^{n-1} 3P(B_i). \tag{58}$$

For any $i \in \{1, \dots, n-1\}$

$$P(B_i) = \frac{1}{6}, \tag{59}$$

therefore we conclude that the expected winning after the nth game are equal to

$$E(W_n) = 3(n-1)\frac{1}{6} = \frac{1}{2}(n-1) \tag{60}$$

2.5 Problem 5

16. (20 points) UCSD Bookstore sells on average 10 books per day. After some research, UCSD Bookstore concluded that the variance of the random variable describing the number of books sold daily is 3.

   (a) (points) Estimate (provide an upper bound) the probability that on a given day UCSD Bookstore sells more than 12 books.

   (b) (points) Assuming that the numbers of books sold on different days are independent and identically distributed, estimate (provide an upper bound) the probability that during the first ten days of December the UCSD Bookstore sells more than 120 books (without using the Central Limit Theorem).

Solution.

(a) Denote by X the number of books sold during a day by the bookstore. Then $E(X) = 10$, $\mathrm{Var}(X) = 3$. Using Markov's inequality

$$P(X > 12) = P(X \geq 13) \leq \frac{E(X)}{13} = \frac{10}{13}. \tag{61}$$

If we now apply Chebyshev's inequality, we get

$$P(X > 12) = P(X \geq 13) = P(X - 10 > 3) \leq P((X-10)^2 > 9) \leq \frac{\mathrm{Var}(X)}{9} = \frac{1}{3}, \tag{62}$$

which gives a much better bound. (Estimating $P(X \geq 12)$ will also get a full credit).

(b) Denote by X_i the number of books sold on December i. Then X_i are independent, identically distributed with $E(X_i) = 10$, $\mathrm{Var}(X) = 3$. Denote by $S_{10} := X_1 + \cdots + X_{10}$ the number of books sold during the first ten days of December. Then using Markov's inequality

$$P(S_{10} > 120) \leq \frac{E(S_{10})}{121} = \frac{100}{121}. \tag{63}$$

Chebyshev's inequality gives

$$P(S_{10} > 120) = P(S_{10} - 100 \geq 21) \leq \frac{\mathrm{Var}(S_n)}{21^2} = \frac{30}{441}, \tag{64}$$

where we used that $\mathrm{Var}(X_1 + \cdots + X_{10}) = 10 \cdot \mathrm{Var}(X_1)$ by independence. (Again, estimating $P(S_{10} \geq 120)$ will be given full credit).