CS2402 Assignment 1 Solution

Question 1 (10pts)
Bill has set a password of 5 characters. Each character can either be a letter or a digit. Suppose the upper case and lower case of a letter are different cases. Write down numerical expressions:
(a) If repeated characters are allowed, find the probability that the password contains at least one lower case letter.
(b) Suppose each character be a letter, a digit, or a special character (assume there are 10 possible special characters to choose from). If no repeated characters are allowed, find the probability that the password contains at least one special character.

Answer 1:
(a) $(62^5 - 36^5) / 62^5$
(b) $1 - (62 * 61 * 60 * 59 * 58) / (72 * 71 * 70 * 69 * 68)$

Question 2 (10pts)
In a lottery game, 48 balls numbered 1, 2, 3, …, 48 are placed in a machine and 6 of them are drawn at random. If the 6 numbers drawn match the numbers that a player had chosen, the player wins the first prize. In this lottery, the order of the numbers doesn't matter.
(a) Compute the probability that a player wins the first prize if he/she plays the lottery game once.
(b) If 5 of the 6 numbers drawn match the numbers that a player has chosen, the player wins a second prize. Compute the probability that the player wins the second prize if he/she plays the lottery game once.

Answer 2:
(a) $C(6,6) / C(48,6) = 1 / 12271512 = 0.0000000815$
(b) $C(6,5) * C(42,1) / C(48,6) = 252 / 12271512 = 0.0000205$

Question 3 (10pts)
For the unfinished game, let us consider a four-person game based on "three-up". More specifically, three coins are placed on a wooden stick, which is used to toss the three fair coins in the air. Therefore, we can get eight sequences of tosses: **HHH, HHT, HTH, HTT, THH, THT, TTH**, and **TTT**. (H for Heads and T for Tails). Four players, Emily, Frank, Grace, and Harry, each chooses two sequences before the game begins, with no two players choosing the same sequence.
For example: Emily - HHH, HHT / Frank - HTH, HTT / Grace - THH, THT / Harry - TTH, TTT.
They agree to the following rules:

- The pot is 768 dollars.
- In each round, the three coins are tossed, and the player whose chosen sequence appears wins one point.
- The first player to reach 4 points wins the game and the entire pot.

The game is stopped early when Emily has 3 points, Frank has 3 points, Grace has 2 points, and Harry has 1 point. Assuming the game cannot continue and they must split the pot based on their chances of winning from this point, how should the pot be divided among the players?

Answer 3:

To get the final results (one wins), you should toss at most 4 times and each time you have 4 cases (anyone wins that round of the game), therefore, there are 256 cases for the final results.

For Grace who has 2 points, there are 27 out of 256 cases that she can win. For Harry who has 1 point, there are 7 out of 256 cases that he can win, and both Emily and Frank have 111 cases to win.

Hence, the probabilities of Grace, Harry, Frank, Emily can win is 27/256, 7/256, 111/256, 111/256 and they should divide the pot like this {Grace:81, Harry:21, Frank:333, Emily:333}.

Question 4 (10pts)

At the production of a certain item, two types of defects, A and B, can occur. We know that P(A) = 0.1, P(B) = 0.2, and P(A ∩ B) = 0.05. Compute the probability that a produced item has

(a) At least one of the defects;
(b) Defect A but not defect B;
(c) None of the defects;
(d) Precisely one of the defects.

Answer 4:

(a) $\mathrm{P}(\mathrm{A} \cap \overline{B}) + \mathrm{P}(\overline{A} \cap \mathrm{B}) + \mathrm{P}(\mathrm{A} \cap \mathrm{B}) = \mathrm{P}(\mathrm{A} \cup \mathrm{B}) = \mathrm{P}(\mathrm{A}) + \mathrm{P}(\mathrm{B}) - \mathrm{P}(\mathrm{A} \cap \mathrm{B}) = 0.25$
(b) $\mathrm{P}(\mathrm{A} \cap \overline{B}) = P(A) - P(A \cap B) = 0.05$
(c) $1 - \mathrm{P}(\mathrm{A} \cup \mathrm{B}) = 1 - 0.25 = 0.75$
(d) $\mathrm{P}(\mathrm{A} \cap \overline{B}) + \mathrm{P}(\overline{A} \cap \mathrm{B}) = \mathrm{P}(\mathrm{A} \cup \mathrm{B}) - \mathrm{P}(\mathrm{A} \cap \mathrm{B}) = 0.25 - 0.05 = 0.2$

Question 5 (15pts)

Suppose your utility function is $\mathrm{U}(w) = \ln(w + 200)$, where w is the wealth, and you are offered a bet on flipping a coin:
Heads, you win \$800.
Tails, you lose \$600.
Please compute your expected change in utility when your wealth is
(a) $w = 500$;
(b) $w = 1000$;
(c) $w = 5000$.

Answer 5:

(a)

Outcome	H	T
Prob.	0.5	0.5
Gain	+800	-600
Utility change	ln(800+200+500)-ln(500+200)=0.7621	ln(200 +500-600) − ln(500 +200)=-1.9459
Prob. X utility change	0.38105	-0.97295
Expected utility change	-0.5919	

(b)

Outcome	H	T
Prob.	0.5	0.5
Gain	+800	-600
Utility change	$\ln(800+200+1000)-\ln(1000+200)=0.5108$	$\ln(200+1000-600)-\ln(1000+200)=-0.6931$
Prob. X utility change	0.2554	-0.34657
Expected utility change	-0.0912	

(c)

Outcome	H	T
Prob.	0.5	0.5
Gain	+800	-600
Utility change	$\ln(800+200+5000)-\ln(5000+200)=0.1431$	ln(200+5000-600)-ln(5000+200)=-0.1226
Prob. X utility change	0.07155	-0.0613
Expected utility change	0.0102	

Question 6 (15pts)
In a box of 10 tickets numbered 0, 1, 2, …, 9. You pick at random 2 tickets from this box. Let X be the larger number in these two tickets, and Y be the smaller one.
(a) Find the distributions of X, Y and their joint distribution.
(b) Find the distribution of Z=X-Y.
(c) Compute E(X), Var(X), E(Z) and Var(Z).

Answer 6:
(a) Joint distribution of X and Y

		Possible value of X								
		1	2	3	4	5	6	7	8	9
Possible value of Y	0	1/45 (0,1) (1,0)	1/45 (0,2) (2,0)	1/45 (0,3) (3,0)	1/45 (0,4) (4,0)	1/45 (0,5) (5,0)	1/45 (0,6) (6,0)	1/45 (0,7) (7,0)	1/45 (0,8) (8,0)	1/45 (0,9) (9,0)
	1		1/45 (1,2) (2,1)	1/45 (1,3) (3,1)	1/45 (1,4) (4,1)	1/45 (1,5) (5,1)	1/45 (1,6) (6,1)	1/45 (1,7) (7,1)	1/45 (1,8) (8,1)	1/45 (1,9) (9,1)
	2			1/45 (2,3) (3,2)	1/45 (2,4) (4,2)	1/45 (2,5) (5,2)	1/45 (2,6) (6,2)	1/45 (2,7) (7,2)	1/45 (2,8) (8,2)	1/45 (2,9) (9,2)
	3				1/45 (3,4) (4,3)	1/45 (3,5) (5,3)	1/45 (3,6) (6,3)	1/45 (3,7) (7,3)	1/45 (3,8) (8,3)	1/45 (3,9) (9,3)
	4					1/45 (4,5) (5,4)	1/45 (4,6) (6,4)	1/45 (4,7) (7,4)	1/45 (4,8) (8,4)	1/45 (4,9) (9,4)
	5						1/45 (5,6) (6,5)	1/45 (5,7) (7,5)	1/45 (5,8) (8,5)	1/45 (5,9) (9,5)

	6							1/45 (6,7) (7,6)	1/45 (6,8) (8,6)	1/45 (6,9) (9,6)
	7								1/45 (7,8) (8,7)	1/45 (7,9) (9,7)
	8									1/45 (8,9) (9,8)

Distribution of X

X	1	2	3	4	5	6	7	8	9
P(X=a)	1/45	2/45	3/45	4/45	5/45	6/45	7/45	8/45	9/45

Distribution of Y

Y	0	1	2	3	4	5	6	7	8
P(Y=a)	9/45	8/45	7/45	6/45	5/45	4/45	3/45	2/45	1/45

(b) Distribution of X-Y

Z	1	2	3	4	5	6	7	8	9
P(Z=a)	9/45 (1,0) (2,1) … (9,8)	8/45 (2,0) (3,1) … (9,7)	7/45 (3,0) (4,1) … (9,6)	6/45 (4,0) (5,1) … (9,5)	5/45 (5,0) (6,1) … (9,4)	4/45 (6,0) (7,1) … (9,3)	3/45 (7,0) (8,1) (9,2)	2/45 (8,0) (9,1)	1/45 (9,0)

(c) **E(X)=**1*1/45+2*2/45+3*1/15+4*4/45+5*1/9+6*2/15+7*7/45+8*8/45+9*1/5=**19/3**
Var(X) = E(X^2) – E(X)^2 = (1+8+27+64+125+216+343+512+729) / 45 – (19/3)2 =
45 – 361/9 = (405-361)/9 = **44/9**

E(Z) = 1*9/45 + 2*8/45 + 3*7/45 + 4*6/45 + 5*5/45 + 6*4/45 + 7*3/45 + 8*2/45 + 9/45 = **11/3**
Var(Z) = E(Z^2) – E(Z)^2 = 1*9/45 + 4*8/45 + 9*7/45 + 16*6/45 + 25*5/45 + 36*4/45 + 49*3/45 + 64*2/45 + 81*1/45 – (11/3)2 = 825/45 – 121/9 = 220/45 = **44/9**

Question 7 (15pts)
In n+m independent Bernoulli(p) trials, let S_n be the number of successes in the first n trials, T_m the number of successes in the last m trials.
(a) What is the distribution of S_n? Why?
Requirement for solution for why part:
You are asked to prove it for the general case. A proof for a special case is given in Lecture note 4. If you cannot give a proof for the general case, you should give a proof for a special case, say, n=5, r=3, p=0.2.
(b) What is the distribution of T_m? Why?
Requirement: Write down your formula. One or two sentences to explain why.
(c) What is the distribution of $S_n + T_m$? Why?
Requirement: Write down your formula. One or two sentences to explain why.
(d) Are S_n and T_m independent? Why?
Requirement: Give your answer and one or two sentences to explain why. You don't need to prove it mathematically. Use of the common sense should be ok.

Answer 7:

(a) We can ignore the last m trials. S_n follows the definition of Binomial Distribution.
Then $P(S_n = r) = C_n^r p^r (1-p)^{(n-r)}$
Proof:

The number of all the possible outcomes in n independent Bernoulli trials is 2^n.

The number of outcomes with r successes is C_n^r.

The probability of each case is $p^r(1-p)^{(n-r)}$.

Therefore, $P(S_n = r) = C_n^r p^r (1-p)^{(n-r)}$

(b) We can ignore the first n trials. T_m follows the definition of Binomial Distribution.
$P(T_m = r) = C_m^r p^r (1-p)^{(m-r)}$

(c) $S_n + T_m$ follows the definition of Binomial Distribution.
$P(S_n + T_m = r) = C_{n+m}^r p^r (1-p)^{(n+m-r)}$

(d)Yes, the function of disjoint blocks of independent variables are independent.

Question 8 (15pts)

Let Z = X + Y be the sum of two independent binomial random variables, $X \sim B(m_0, p_0)$ and $Y \sim B(m_1, p_1)$. Prove that:

$$\mathrm{Var}(Z) \leq E[Z](1 - \frac{E[Z]}{m_0 + m_1})$$

Specify the condition that makes the inequality hold with equality.

Answer 8:

The random variables X and Y are independent, so the variance of the sum is equal to the sum of the variances:

$$\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y) = E[X]\left(1 - \frac{E[X]}{m_0}\right) + E[Y](1 - \frac{E[Y]}{m_1})$$

In order to prove the inequality, it is therefore sufficient to prove that:

$$E[X]\left(1 - \frac{E[X]}{m_0}\right) + E[Y](1 - \frac{E[Y]}{m_1}) \leq E[Z](1 - \frac{E[Z]}{m_0 + m_1})$$

Substituting E[X] + E[Y] for E[Z]:

$$E[X]\left(1 - \frac{E[X]}{m_0}\right) + E[Y](1 - \frac{E[Y]}{m_1}) \leq (E[X] + E[Y])(1 - \frac{E[X] + E[Y]}{m_0 + m_1})$$

Multiplying out the brackets:

$$E[X] - \frac{E[X]^2}{m_0} + E[Y] - \frac{E[Y]^2}{m_1} \leq E[X] + E[Y] - \frac{(E[X] + E[Y])^2}{m_0 + m_1}$$

Subtracting E[X] and E[Y] from both sides and reversing the inequality:

$$\frac{E[X]^2}{m_0} + \frac{E[Y]^2}{m_1} \geq \frac{(E[X] + E[Y])^2}{m_0 + m_1}$$

Expanding the right-hand side:

$$\frac{E[X]^2}{m_0} + \frac{E[Y]^2}{m_1} \geq \frac{E[X]^2 + 2E[X]E[Y] + E[Y]^2}{m_0 + m_1}$$

Multiplying by $m_0 m_1 (m_0 + m_1)$:

$$(m_0 m_1 + m_1^2)E[X]^2 + (m_0^2 + m_0 m_1)E[Y]^2 \geq m_0 m_1 (E[X]^2 + 2E[X]E[Y] + E[Y]^2)$$

Deducting the right-hand side:

$$m_1^2 E[X]^2 - 2m_0 m_1 E[X]E[Y] + m_0^2 E[Y]^2 \geq 0$$

Equivalently:

$$(m_1 E[X] - m_0 E[Y])^2 \geq 0$$

The square of a real number is always greater than or equal to zero, so this is true for all independent binomial distributions that X and Y could take. This is sufficient to prove the inequality.

When $m_1 E[X] = m_0 E[Y]$, the equality holds.