Jae Hong Lee
Prof. C.E. Tsourakakis
Boston University

March 1, 2024

— CAS CS 365 - Assignment 1

1 Probability Problems

1.1 Problem 1.1

Two fair dice are rolled. Show that the event that their sum is 7 is independent of the score shown by the first die.

To prove independence of two die's score, we need to show the relation: P(A|B) = P(A) and P(B|A) = P(B) between two probability P(A) and P(B), which each die shows to add towards 7. As we know P(A|B) = $\frac{P(A \cap B)}{P(B)}$ This is a pair to make 7 with two dies sum (A∩B).

A	B	SUM
1	6	7
2	5	7
3	4	7
4	3	7
5	2	7
6	1	7

So total 6 pairs to make 7. So the probability to make 7 with A and B independent dice, p(A∩B) is $\frac{6}{36}$. As we can see from the table, In order to make 7, the die B can have the variable from 1 6. Since each number 1 - 6 have a pair to make 7, the probability of die b to be to make 7 is $6 \cdot \frac{1}{6} = 1$. So the P(A|B) is $\frac{6}{36} \cdot 1 = \frac{1}{6}$. As corresponds to die B, The probability of die A to make 7 with the summation from die B is one of 1 - 6, depends on die B's result. So P(A) = $\frac{1}{6}$. As we have shown that P(A|B) = P(A), with the same approach P(B|A) is equal to P(B).

Now we know P(A|B) = P(B) and P(B|A) = P(B), the event that their sum is 7 is independent of the score shown by the first die.

1.2 Problem 1.2

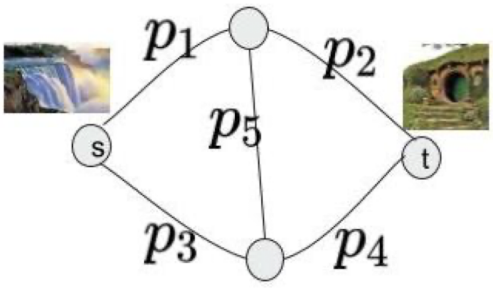

Figure 1: How likely is water to reach t from s?

Consider the network shown in Figure 1. Recall that p_i is the probability pipe i breaks down, and that pipes break down independently.

What is the probability water can go from the water source s to the destination village t? Explain your answer. You do not need to simplify it algebraically.(Figure 1)

When we look at the figure above, we can observe that p_5 is allowing outputs of p_1 and p_3 go across to p_2 and p_4 to reach t. So we can split the probability of water go from source to target to either p_5 works or fails.

First: When p_5 fails.

When p_5 fails, the only possible ways to reach s to t are either go through $p_1 -> p_2$ or $p_3 -> p_4$. The probability of p_i represents the probability of pipe i's failure, so let's say $\bar{p_i}$ represents the probability of pipe i not being broken. Let's say event that go through p_1 *and* p_2 is A_1, which is $(\bar{p_1} \cdot \bar{p_2})$ and go through p_3 *and* p_4 is $A_2 = (\bar{p_3} \cdot \bar{p_4})$. So the overall probability of water reach t from s is $p_5 \cdot (A_1 \cup A_2)$. The inclusion - exclusion tells us $A \cup B = A + B - (A \cap B)$. The probability of water reach t from source s when p_5 is broken is $p_5 \cdot (\bar{p_1} \cdot \bar{p_2} + \bar{p_3} \cdot \bar{p_4} - (\bar{p_1} \cdot \bar{p_2} \cdot \bar{p_3} \cdot \bar{p_4}))$.

Second: When p_5 works.

When we know the probability of failure, we can calculate the probability of success by 1- failure. When p_5 works, the ways of failure are when p_1 *and* p_3 fail or p_2 *and* p_4 fails. Let's say event that p_1 *and* p_3 broke A_3, which is $(p_1 \cdot p_3)$ and p_2 *and* p_4 broke is $A_4 = (p_2 \cdot p_4)$. Then the probability of water can reach s to t when p_5 is working: $\bar{p_5} \cdot (1 - (A_3 \cup A_4)) = (1 - p_5) \cdot (1 - (p_1 \cdot p_3 + p_2 \cdot p_4 - p_1 \cdot p_2 \cdot p_3 \cdot p_4))$

So the overall probability of water reach s to t is

$$p_5 \cdot (\bar{p_1} \cdot \bar{p_2} + \bar{p_3} \cdot \bar{p_4} - (\bar{p_1} \cdot \bar{p_2} \cdot \bar{p_3} \cdot \bar{p_4})) + (1 - p_5) \cdot (1 - (p_1 \cdot p_2 + p_3 \cdot p_4 - p_1 \cdot p_2 \cdot p_3 \cdot p_4))$$

1.3 Problem 1.3

Suppose X has an exponential distribution, that is the pdf is given by

$$f_X(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, 0 \leq x < \infty, \lambda > 0$$

You obtain iid samples $x_1, ..., x_n$ from $f_X(x)$

1.3.1 (a) Compute the expectation and variance of X.

Expectation of X,

$$E[x] = \int_0^\infty x \cdot f_X(x) \cdot dx = \int_0^\infty x \cdot \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx$$

Let's say $u = -\frac{x}{\lambda}$, then

$$\int_0^\infty x \cdot \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \cdot dx = -\int_0^\infty u \cdot e^u \cdot dx$$

since $u = -\frac{x}{\lambda}, du = -\frac{1}{\lambda} dx, dx = -\lambda du$

$$-\int_0^\infty u \cdot e^u \cdot dx = \lambda (\int_0^\infty u \cdot e^u \cdot du)$$

$$Partial - Integral : \int f(x) g\acute{(}x) \cdot dx = f(x)g(x) - \int f\acute{(}x) g(x) \cdot dx$$

Let's say f(x) = u, $g\acute{(}x) = e^u$. Then we can get $f\acute{(}x) = 1, g(x) = e^u$

So

$$\lambda(\int u \cdot e^u \cdot dx) = \lambda(u \cdot e^u - \int 1 \cdot e^u \cdot du) = \lambda(u \cdot e^u - e^u) = \lambda(\Big[-\frac{x}{\lambda} \cdot e^{-\frac{x}{\lambda}} - e^{-\frac{x}{\lambda}}\Big]_0^\infty) = \lambda(1) = \lambda$$

The expectation of X is λ

In order to find the variance, I will use the equation

$$Variance\ of\ X = E[X^2] - (E[X])^2$$

Let find $E[X^2]$ first

$$E[x^2] = \int_0^\infty x^2 \cdot f_X(x) \cdot dx = \int_0^\infty x^2 \cdot \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \cdot dx$$

Once again, let's say $u = -\frac{x}{\lambda}$, then

$$\int_0^\infty x^2 \cdot \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx = \lambda \cdot \int_0^\infty u^2 \cdot e^u \cdot dx$$

From finding expectation of X, We know $dx = -\lambda \cdot du$

$$\lambda \cdot \int_0^\infty u^2 \cdot e^u \cdot dx = -\lambda^2 \cdot \int_0^\infty u^2 \cdot e^u \cdot du$$

Using Partial Integral,

$$Partial - Integral : \int f(x)g\acute{(}x) \cdot dx = f(x)g(x) - \int f\acute{(}x)g(x) \cdot dx$$

Let's say f(x) = u^2, $g\acute{(}x) = e^u$. Then we can get $f\acute{(}x) = 2u,\ g(x) = e^u$

$$-\lambda^2 \cdot \int_0^\infty u^2 \cdot e^u \cdot du = -\lambda^2 \cdot (u^2 \cdot e^u - \int 2u \cdot e^u \cdot du) = -\lambda^2 \cdot (u^2 \cdot e^u - 2\int u \cdot e^u \cdot du)$$

For $\int u \cdot e^u \cdot du$, we have done the integral for calculating the expectation of X

$$-\lambda^2 \cdot (u^2 \cdot e^u - 2\int u \cdot e^u \cdot du) = -\lambda^2 \cdot (u^2 \cdot e^u - 2(u \cdot e^u - e^u)) = -\lambda^2 \cdot (u^2 \cdot e^u - 2u \cdot e^u + 2e^u)$$

$$-\lambda^2 \cdot (u^2 \cdot e^u - 2u \cdot e^u + 2e^u) = -\lambda^2 \cdot (\frac{x^2}{\lambda^2} \cdot e^{-\frac{x}{\lambda}} + 2\frac{x}{\lambda} \cdot e^{-\frac{x}{\lambda}} + 2e^{-\frac{x}{\lambda}})$$

$$-\lambda^2 \cdot (\left[\frac{x^2}{\lambda^2} \cdot e^{-\frac{x}{\lambda}} + 2\frac{x}{\lambda} \cdot e^{-\frac{x}{\lambda}} + 2e^{-\frac{x}{\lambda}}\right]_0^\infty) = -\lambda^2 \cdot (0 - (2)) = 2\lambda^2$$

We got $E[x^2] = 2\lambda^2$, so the variance of the $f_X(x)$ is $2\lambda^2 - (\lambda)^2 = \lambda^2$

The Variance of X is λ^2

1.3.2 (b) Use the maximum likelihood principle to learn λ

Using the given pdf $f_X(x)$ to find likelihood function. Since the problem said I obtain iid samples $x_1, ..., x_n$ from $f_X(x)$ the likelihood function is

$$Likelihood\ function\ f_X(x) = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \frac{1}{\lambda} e^{\frac{-x_i}{\lambda}} = \frac{1}{\lambda^n} e^{\frac{-\sum_{i=1}^{n} x_i}{\lambda}}$$

The log likelihood function is

$$log(\frac{1}{\lambda^n} e^{\frac{-\sum_{i=1}^{n} x_i}{\lambda}}) = log(\frac{1}{\lambda^n}) + log(e^{\frac{-\sum_{i=1}^{n} x_i}{\lambda}}) = n \cdot log(\frac{1}{\lambda}) + \frac{-\sum_{i=1}^{n} x_i}{\lambda} \cdot log(e) = n \cdot log(\frac{1}{\lambda}) + \frac{-\sum_{i=1}^{n} x_i}{\lambda}$$

In order to find the max value of Lambda we need to derivative the above equation and find when the equation equals to 0.

$$(n \cdot log(\frac{1}{\lambda}) + \frac{-\sum_{i=1}^{n} x_i}{\lambda}) \cdot \frac{d}{d\lambda} = n \cdot \lambda \cdot -\frac{1}{\lambda^2} + -\sum_{i=1}^{n} x_i \cdot -\frac{1}{\lambda^2} = -\frac{n}{\lambda} + \frac{\sum_{i=1}^{n} x_i}{\lambda^2}$$

$$-\frac{n}{\lambda} + \frac{\sum_{i=1}^{n} x_i}{\lambda^2} = 0,\ so\ \frac{n}{\lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda^2},\ n \cdot \lambda = \sum_{i=1}^{n} x_i,\ \lambda = \frac{\sum_{i=1}^{n} x_i}{n}$$

So we learn $\lambda = \frac{\sum_{i=1}^{n} x_i}{n}$, which is the average of the $x_1, ..., x_n$

1.3.3 (c) Use the method of moments to learn λ

Method of moments equating sample moments with theoretical moments. The equation of method of moment is

$$a_j = \frac{1}{n} \sum_{i=1}^{n} x_i^j, for j = 1, ..., k$$

.

j is the series of moments. From 1.3.1 (a) we learn the expectation of the exponential distribution is λ. In the method of moment, expectation is the 1st moment. When j = 1, $a_1 = \frac{\sum_{i=1}^{n} x_i}{n}$, which is the same as we got from MLE method.

so $\lambda = \frac{\sum_{i=1}^{n} x_i}{n}$

1.4 Problem 1.4

Suppose that 5% of men and .25% of women are color-blind. A person is chosen at random and that person is color-blind. What is the probability that the person is male? Assume males and females to be in equal numbers.

With the given information, we can say that out of male, the probability of color-blind male is P(B|M) = 0.05 and out of female, the probability of color-blind women is $P(B|\bar{M}) = 0.0025$. The problem is asking the probability of the person is color-blind given that person is male $Pr(M|B)$.

From Bayes rule, $Pr(M|B) = \frac{Pr(B|M) \cdot Pr(M)}{Pr(B)}$. $Pr(B) = Pr(B|M) \cdot Pr(M) + Pr(B|\bar{M}) \cdot Pr(\bar{M}) = 0.05 \cdot 0.5 + 0.0025 \cdot 0.5 = 0.02625$.

So, $Pr(M|B) = \frac{0.05 \cdot 0.5}{0.02625} = \frac{20}{21} \approx 0.95238$.

The probability that the person is make is $\frac{20}{21} \approx 0.95238$

1.5 Problem 1.5

Prove that correlation between two random variables is between [-1, 1]. When is it equal to -1 and +1 respectively?

Reminder The correlation between two random variables can be defined in terms of their covariance and standard deviations. Specifically, the Pearson correlation coefficient, r, between two variables X and Y is given by:

$$r = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

where Cov(X, Y) is the covariance between X and Y, measuring how much the variables change together, and σ_X and σ_Y are the standard deviations of X and Y, respectively, measuring the spread of each variable around their mean.

This formula normalizes the covariance by the product of the standard deviations of the variables, thus scaling the correlation to lie between -1 and 1, where -1 indicates perfect negative linear correlation, 1 indicates perfect positive linear correlation, and 0 indicates no linear correlation.

For random variable X and Y, if their expectation is E[X] = μ_x and E[Y] = μ_y, the covariance of X and Y is $E(x-\mu_x)(y-\mu_y)$. Standard deviation of X is $\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i-\mu_x)^2}{n}}$, and Y is $\sigma_y = \sqrt{\frac{\sum_{i=1}^n (y_i-\mu_y)^2}{n}}$.

Rewrite Pearson correlation:

$$\frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\frac{\sum_{i=1}^n (x_i-\mu_x)(y_i-\mu_y)}{n}}{\sqrt{\frac{\sum_{i=1}^n (x-\mu_x)^2}{n}} \cdot \sqrt{\frac{\sum_{i=1}^n (y-\mu_y)^2}{n}}}$$

Square it:

$$\frac{(\frac{\sum_{i=1}^n (x_i-\mu_x)(y_i-\mu_y)}{n})^2}{(\sqrt{\frac{\sum_{i=1}^n (x_i-\mu_x)^2}{n}} \cdot \sqrt{\frac{\sum_{i=1}^n (y_i-\mu_y)^2}{n}})^2} = \frac{\frac{(\sum_{i=1}^n (x_i-\mu_x)(y_i-\mu_y))^2}{n^2}}{\frac{\sum_{i=1}^n (x_i-\mu_x)^2}{n} \cdot \frac{\sum_{i=1}^n (y_i-\mu_y)^2}{n}} = \frac{(\sum_{i=1}^n (x_i-\mu_x)(y_i-\mu_y))^2}{\sum_{i=1}^n (x_i-\mu_x)^2 \cdot \sum_{i=1}^n (y_i-\mu_y)^2}$$

If we say $u_i = (x_i - \mu_x)$ and $v_i = (y_i - \mu_y)$, then we can observe $r^2 = \frac{(\sum_{i=1}^n u_i \cdot v_i)^2}{\sum_{i=1}^n u_i^2 \cdot \sum_{i=1}^n v_i^2}$.

In order the bound of correlation to [-1, 1] Then we need to show $r^2 \leq 1$, $\frac{(\sum_{i=1}^n u_i \cdot v_i)^2}{\sum_{i=1}^n u_i^2 \cdot \sum_{i=1}^n v_i^2} \leq 1$, so $(\sum_{i=1}^n u_i \cdot v_i)^2 \leq \sum_{i=1}^n u_i^2 \cdot \sum_{i=1}^n v_i^2$

Through **Cauchy–Schwartz inequality prove**, we can show $(\sum_{i=1}^n u_i \cdot v_i)^2 \leq \sum_{i=1}^n u_i^2 \cdot \sum_{i=1}^n v_i^2$.

Let's prove Cauchy–Schwartz inequality:

For example, say there is a quadratic equation, f(x) = $ax^2 + bx + c$. Using quadratic formula, $\frac{-b \pm \sqrt{b^2-4ac}}{2a}$ we can find the solutions of equation. Discriminant(D): $b^2 - 4ac$ tells us how many solutions exists in this equation. When $b^2 - 4ac > 0$, the solutions are $\frac{-b-\sqrt{b^2-4ac}}{2a}$ and $\frac{-b+\sqrt{b^2-4ac}}{2a}$. When $b^2 - 4ac = 0$, the solution is $\frac{-b}{2a}$. And lastly when $b^2 - 4ac < 0$, there is no solution.

If f(x) is defined as follow:

$$f(x) = (a_1 x - b_1)^2 + (a_2 x - b_2)^2 +, ..., (a_n x - b_n)^2$$

$$= (a_1^2 + a_2^2 +, ..., + a_n^2)x^2 - 2(a_1 b_1 + a_2 b_2 +, ..., + a_n b_n)x + (b_1^2 + b_2^2 +, ..., b_n^2)$$

f(x) is the summation of perfect square equations, so for x, discriminant is less than 0.

$$D = 4(a_1b_1 + a_2b_2+, ..., +a_nb_n)^2 - 4(a_1^2 + a_2^2+, ..., +a_n^2) \cdot (b_1^2 + b_2^2+, ..., b_n^2) \leq 0$$

$$\frac{D}{4} = (a_1b_1 + a_2b_2+, ..., +a_nb_n)^2 - (a_1^2 + a_2^2+, ..., +a_n^2) \cdot (b_1^2 + b_2^2+, ..., b_n^2) \leq 0$$

so

$$(a_1b_1 + a_2b_2+, ..., +a_nb_n)^2 \leq (a_1^2 + a_2^2+, ..., +a_n^2) \cdot (b_1^2 + b_2^2+, ..., b_n^2)$$

$\therefore$ this is the form of $(u \cdot v)^2 \leq u^2 \cdot v^2$ when $u = (a_1 + a_2+, ..., +a_n)$, $v = (b_1 + b_2+, ..., +b_n)$

Using Cauchy–Schwartz inequality, $(\sum_{i=1}^{n} u_i \cdot v_i)^2 \leq \sum_{i=1}^{n} u_i^2 \cdot \sum_{i=1}^{n} v_i^2$

so the Pearson correlation's range is [-1, 1].

When random variable Y = X, the covariance of X and Y is $E(x - \mu_x)(x - \mu_x)$, so the correlation is

$$\frac{\frac{\sum_{i=1}^{n}(x_i - \mu_x)(x_i - \mu_x)}{n}}{\sqrt{\frac{\sum_{i=1}^{n}(x - \mu_x)^2}{n}} \cdot \sqrt{\frac{\sum_{i=1}^{n}(x - \mu_x)^2}{n}}}$$

leads to r = $\frac{u \cdot u}{\sqrt{u^2} \cdot \sqrt{u^2}} = 1$.

When Y = -X, the covariance of X and Y is $E(x - \mu_x)(-x - \mu_{-x})$, so the correlation is

$$\frac{\frac{\sum_{i=1}^{n}(x_i - \mu_x)(-x_i - \mu_{-x})}{n}}{\sqrt{\frac{\sum_{i=1}^{n}(x - \mu_x)^2}{n}} \cdot \sqrt{\frac{\sum_{i=1}^{n}(-x - \mu_{-x})^2}{n}}}$$

leads to r = $\frac{u \cdot -u}{\sqrt{u^2} \cdot \sqrt{u^2}} = -1$.

When Y = X and Y = -X, correlation is +1 and -1 respectively

1.6 Problem 1.6

Player A and B are solving a puzzle independently. The time needed to solve this puzzle for anyone is random, and it follows the same distribution, i.e., a uniform distribution between 10 seconds and 60 seconds. Player agree the game is a deuce if the difference between their finish time is within 5 seconds. What is the probability they have a deuce?

There are two ways to approach this problem. In the first case, use the area underneath the coordinate graph, and in the second case, you use the range of the boundary to find the integral.

(1)-Graph

The time needed to solve this puzzle for player A and B is the uniform distribution. So, for the uniform distribution between 10 seconds and 60 seconds, we can make graph like this(Figure 2).

Since the Player B's solving time is in the range with in 5 seconds of Player A and Player A's solving time is in the range within 5 seconds of Player B. The graph range looks like this. (Figure 3)

The area under the graph is the probability that they have a deuce. The colored area is the subtraction of two triangles of width 45 from the square of width 50. So the area is $Big\ square - 2 \cdot trainges = (50) \cdot (50) - 2 \cdot \frac{(45) \cdot (45)}{2} = 50^2 - 45^2 = 2500 - 2025 = 475$.

So the colored area ratio from the Player A and B's uniform distribution is $\frac{475}{2500} = \frac{19}{100} = 0.19$

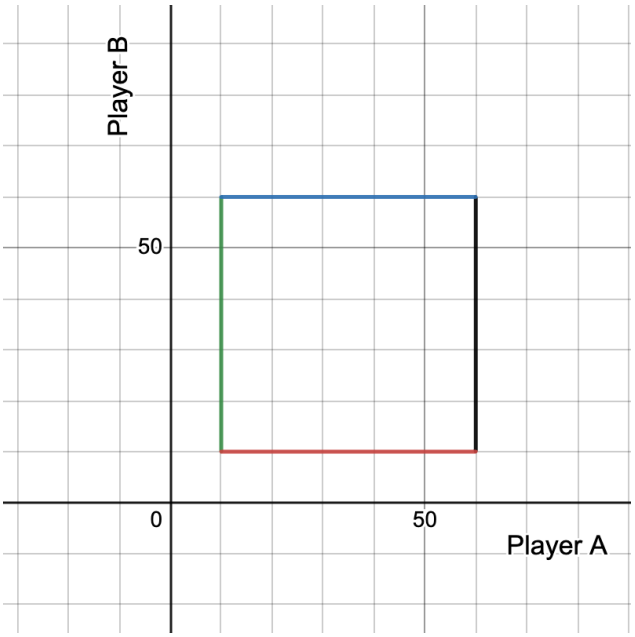

Figure 2: Player A and B uniform distribution

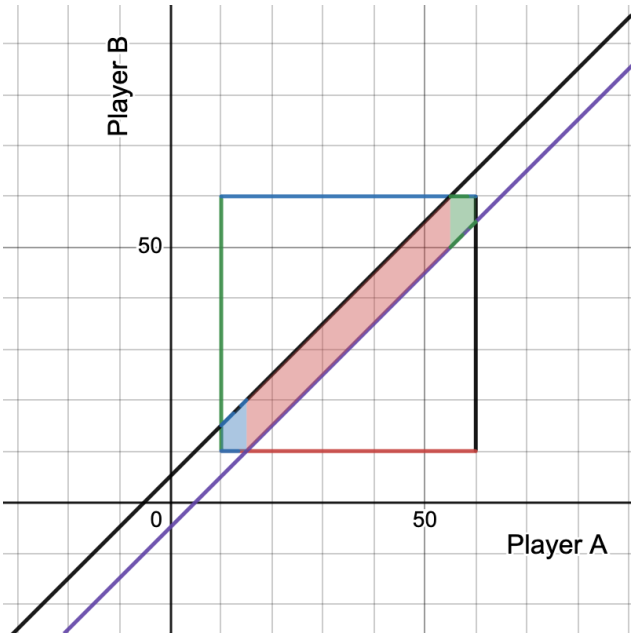

Figure 3: Probability area of Player A and B's deuce

(2) - Using Joint Distribution

In order to support that the answer is 0.19, let's try integral approach. Let's say player A trying to solve the puzzle first. Then Player B follows after. Since Player A solving the puzzle and Player B solving the puzzle is an independent event, Their Joint Distribution is multiple of each uniform probability density function: $\frac{1}{50} * \frac{1}{50} = \frac{1}{2500}$

Let's say player A's solving time is A, Player B's solving time is B. For range of A, I can say that B's

range is A-5 to A + 5 ($A-5 \leq B \leq A+5$). However, in A's range of 10 - 15 seconds ($10 \leq A \leq 15$), then B's range is 10 to A + 5 ($10 \leq A \leq A+5$) because B cannot be less than 10. Similar to this. in A's range of 55 - 60 seconds ($55 \leq A \leq 60$), B's range is A-5 to 60 ($A-5 \leq B \leq 60$) since B cannot exceed 60 seconds. So we can make three different equations and add up those integrals to get the probability.

$$\int_{10}^{15}\int_{10}^{A+5}\frac{1}{2500}\cdot dB\cdot dA+\int_{15}^{55}\int_{A-5}^{A+5}\frac{1}{2500}\cdot dB\cdot dA+\int_{55}^{60}\int_{A-5}^{60}\frac{1}{2500}\cdot dB\cdot dA$$

$$=\frac{1}{2500}\cdot(\int_{10}^{15}\int_{10}^{A+5}1\cdot dB\cdot dA+\int_{15}^{55}\int_{A-5}^{A+5}1\cdot dB\cdot dA+\int_{55}^{60}\int_{A-5}^{60}1\cdot dB\cdot dA)$$

$$\int_{10}^{15}\int_{10}^{A+5}1\cdot dB\cdot dA=\int_{10}^{15}|B|_{10}^{A+5}\cdot dA=\int_{10}^{15}(A+5-10)\cdot dA=\int_{10}^{15}A-5\cdot dA=\left|\frac{1}{2}A^2-5A\right|_{10}^{15}=\frac{125}{2}-25$$

$$\int_{15}^{55}\int_{A-5}^{A+5}1\cdot dB\cdot dA=\int_{15}^{55}|B|_{A-5}^{A+5}\cdot dA=\int_{10}^{15}(A+5-A+5)\cdot dA=\int_{15}^{55}10\cdot dA=|10A|_{15}^{55}=550-150=400$$

$$\int_{55}^{60}\int_{A-5}^{60}1\cdot dB\cdot dA=\int_{55}^{60}|B|_{A-5}^{60}\cdot dA=\int_{55}^{60}(60-A+5)\cdot dA=\int_{55}^{60}-A+65\cdot dA=\left|-\frac{1}{2}A^2+65A\right|_{55}^{60}=-\frac{575}{2}+325$$

$$=\frac{1}{2500}\cdot(\frac{125}{2}-25+400-\frac{575}{2}+325)=\frac{1}{2500}\cdot(475)=\frac{475}{2500}=\frac{19}{100}=0.19$$

So the probability of Player A and Player B have a deuce is 0.19.

1.7 Problem 1.7

You are drawing a train ticket from a box. Four different types of tickets are in the box with equal number, including tickets only for the first class, the third class, the coach class, and VIP tickets that can sit anywhere on the train. Consider A_k as the event you can sit at $k-th$ class according to the ticket you draw, where k = 1, 2, 3. Are the three events A_1, A_2, A_3 pairwise independent? Are they mutually independent?

There are four kinds of tickets in the box. The probability to get each ticket is as follows:

	Y total
1st	$\frac{1}{4}$
3rd	$\frac{1}{4}$
Coach	$\frac{1}{4}$
VIP	$\frac{1}{4}$

For the event A_1 that I can sit in 1st class the probability is $Pr(1st)+Pr(VIP)=\frac{1}{2}$. For each event A_1-A_3 we can get the probability to each as the table:

	Probability
1st	$\frac{1}{2}$
3rd	$\frac{1}{2}$
Coach	$\frac{1}{2}$

In order to A_1, A_2, A_3 to be pairwise independent,

$Pr(A_1 \cap A_2) = Pr(A_1) \cdot Pr(A_2)$, $Pr(A_1 \cap A_3) = Pr(A_1) \cdot Pr(A_3)$, $Pr(A_2 \cap A_3) = Pr(A_2) \cdot Pr(A_3)$

$Pr(A_1) \cdot Pr(A_2) = \frac{1}{4}$ and $Pr(A_1 \cap A_2) = \frac{1}{4}$ since only VIP ticket can let us sit in the 1st class and 3rd class. It works the same as the others. **So the three events A_1, A_3, A_3 are pairwise independent.**

In order to A_1, A_2, A_3 to be mutual independent,

$Pr(A_1 \cap A_2 \cap A_3) = Pr(A_1) \cdot Pr(A_2) \cdot Pr(A_3)$

$Pr(A_1) \cdot Pr(A_2) \cdot Pr(A_3) = \frac{1}{8}$ and $Pr(A_1 \cap A_2 \cap A_{3]}) = \frac{1}{4}$ since the VIP ticket can let us sit in the 1st, 3rd, and Coach class all three. $Pr(A_1 \cap A_2 \cap A_3) \neq Pr(A_1) \cdot Pr(A_2) \cdot Pr(A_3)$. **So the three events A_1, A_3, A_3 are not mutual independent.**

1.8 Problem 1.8

A new type of flu is spreading in the community, with statistics showing that 10% of the population is affected. Research also found this flu causes headaches with a high probability, i.e., 80% of the patients got headaches as a symptom. On the other hand, 15% of the population complains of headaches for various reasons. When Bob wakes up in the morning, he feels a headache; what is the chance he gets the flu?

The probability of flu affect population is given my the problem, $Pr(F) = 0.1$. With in the people who has flu, 80% has headaches as a symptom. Let's say the probability of having headache as Pr(H). The probability of people who has flu experiencing headache is $Pr(H|F) = 0.8$. 15% of population experience headache includes the flu, so $Pr(H) = 0.15$. The problem ask the probability of bob feels headache and the chance he has flu, which is $Pr(F|H)$.

Using Bayes Rule:

$$Pr(F|H) = \frac{Pr(H|F) \cdot Pr(F)}{Pr(H)}$$

so

$$\frac{0.8 \cdot 0.1}{0.15} = \frac{0.08}{0.15} = \frac{8}{15}$$

$\therefore \frac{8}{15}$

2 Linear Algebra

2.1 Problem 2.1

Let $A \in R^{m \times n}$ be a real m x n matrix. Prove that eigenvalues of AA^T and A^TA are real and non-negative.

(i) AA^T

AA^T is an $m \times m$ matrix and $AA^T \cdot x = \lambda \cdot x$

For λ, there exists an normalized eigenvector, which has a length of 1. Let's call it v

so,

$$\|v\|^2 = 1$$

Now let's look at λ

$$\lambda = 1 \cdot \lambda = \|v\|^2 \cdot \lambda$$
$$= v^T \cdot v \cdot \lambda$$

since λ is a constant, so we can move.

$$= v^T \cdot \lambda \cdot v$$

$$AA^T \cdot x = \lambda \cdot x, \text{ so we can substitute } AA^T \text{ as } \lambda.$$
$$= v^T \cdot A \cdot A^T \cdot v$$
$$= (v^T \cdot A) \cdot (A^T \cdot v) = (A^T \cdot v)^T \cdot (A^T \cdot v)$$
$$= \left\|A^T \cdot v\right\|^2$$
$$\lambda \text{ is the length of } A^T \cdot v, \text{ so } \lambda \geq 0$$

(ii) $A^T A$

$A^T A$ is an $n \times n$ matrix and $A^T A \cdot x = \lambda \cdot x$

Similarly, for λ, there exists an normalized eigenvector, which has a length of 1. Let's call it w

so,

$$\|w\|^2 = 1$$

Now let's look at λ

$$\lambda = 1 \cdot \lambda = \|w\|^2 \cdot \lambda$$
$$= w^T \cdot w \cdot \lambda$$
$$\text{since } \lambda \text{ is a constant, so we can move.}$$
$$= w^T \cdot \lambda \cdot w$$
$$A^T A \cdot x = \lambda \cdot x, \text{ so we can substitute } A^T A \text{ as } \lambda.$$
$$= w^T \cdot A^T \cdot A \cdot w$$
$$= (w^T \cdot A^T) \cdot (A \cdot w) = (A \cdot w)^T \cdot (A \cdot w)$$
$$= \|A \cdot w\|^2$$
$$\lambda \text{ is the length of } A \cdot w, \text{ so } \lambda \geq 0$$

2.2 Problem 2.2

Let $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$ **and** $B = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}$ **and** $C = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$**. Verify that the eigenvalues of (i) A are 2 + i and 2 - i where** $i^2 = -1$**, (ii) B are 4 and -2, (iii) C are 1 and 0. Find the eigenvectors by using (i), (ii), (iii) and the definition of an eigenvector.**

(i) $2 + i$ **and** $2 - i$

We have $A \cdot x = \lambda \cdot x$. We can rewrite it to $(A - \lambda \cdot I) \cdot x = 0$

For $2 + i$, $(A - \lambda \cdot I) = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} - \begin{bmatrix} 2+i & 0 \\ 0 & 2+i \end{bmatrix} = \begin{bmatrix} -1-i & -1 \\ 2 & 1-i \end{bmatrix}$. If the determinant of $(A - \lambda \cdot I) = 0$ then λ is the eigen values. $det(\begin{bmatrix} -1-i & -1 \\ 2 & 1-i \end{bmatrix}) = (-1-i) \cdot (1-i) - (-1) \cdot (2) = -2 - (-2) = 0.$

We verified it is the eigen values, so let's find eigen vector x. For $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, let's say $x_1 = i$. Then $i \cdot \begin{bmatrix} -1-i \\ 2 \end{bmatrix} + x_2 \cdot \begin{bmatrix} -1 \\ 1-i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} -i+1-x_2 = 0 \\ 2i + x_2 - ix_2 = 0 \end{bmatrix}$. So $x_2 = -i + 1$. We still need to normalize x_1 and x_2 to be length of 1. $\|x\| = \sqrt{x_1^2 + x_2^2} = \sqrt{(i)^2 + (-i+1)^2} = \sqrt{-1-2i}$. So x norm eigen vector is $\begin{bmatrix} \frac{i}{\sqrt{-1-2i}} \\ \frac{-i+1}{\sqrt{-1-2i}} \end{bmatrix}$

For $2-i$, $(A-\lambda\cdot I)=\begin{bmatrix}1 & -1\\ 2 & 3\end{bmatrix}-\begin{bmatrix}2-i & 0\\ 0 & 2-i\end{bmatrix}=\begin{bmatrix}-1+i & -1\\ 2 & 1+i\end{bmatrix}$. If the determinant of $(A-\lambda\cdot I)=0$ then λ is the eigen values. $det(\begin{bmatrix}-1+i & -1\\ 2 & 1+i\end{bmatrix})=(-1+i)\cdot(1+i)-(-1)\cdot(2)=-2-(-2)=0$.

We verified it is the eigen values, so let's find eigen vector x. For $\begin{bmatrix}x_1\\ x_2\end{bmatrix}$, let's say $x_1=i$. Then $i\cdot\begin{bmatrix}-1+i\\ 2\end{bmatrix}+x_2\cdot\begin{bmatrix}-1\\ 1+i\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$, $\begin{bmatrix}-i-1-x_2=0\\ 2i+x_2+ix_2=0\end{bmatrix}$. So $x_2=-i-1$. We still need to normalize x_1 and x_2 to be length of 1. $\|x\|=\sqrt{x_1^2+x_2^2}=\sqrt{(i)^2+(-i-1)^2}=\sqrt{-1+2i}$. So x norm eigen vector is $\begin{bmatrix}\frac{i}{\sqrt{-1+2i}}\\ \frac{-i-1}{\sqrt{-1+2i}}\end{bmatrix}$

(ii) 4 and -2

We have $B\cdot x=\lambda\cdot x$. We can rewrite it to $(B-\lambda\cdot I)\cdot x=0$

For 4, $(B-\lambda\cdot I)=\begin{bmatrix}1 & 3\\ 3 & 1\end{bmatrix}-\begin{bmatrix}4 & 0\\ 0 & 4\end{bmatrix}=\begin{bmatrix}-3 & 3\\ 3 & -3\end{bmatrix}$. If the determinant of $(B-\lambda\cdot I)=0$ then λ is the eigen values. $det(\begin{bmatrix}-3 & 3\\ 3 & -3\end{bmatrix})=(-3)\cdot(-3)-(3)\cdot(3)=9-(9)=0$.

We verified it is the eigen values, so let's find eigen vector x. For $\begin{bmatrix}x_1\\ x_2\end{bmatrix}$, let's say $x_1=1$. Then $1\cdot\begin{bmatrix}-3\\ 3\end{bmatrix}+x_2\cdot\begin{bmatrix}3\\ -3\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$, $\begin{bmatrix}-3-3x_2=0\\ 3-3x_2=0\end{bmatrix}$. So $x_2=1$. We still need to normalize x_1 and x_2 to be length of 1. $\|x\|=\sqrt{x_1^2+x_2^2}=\sqrt{(1)^2+(1)^2}=\sqrt{2}$. So x norm eigen vector is $\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}$

For -2, $(B-\lambda\cdot I)=\begin{bmatrix}1 & 3\\ 3 & 1\end{bmatrix}-\begin{bmatrix}-2 & 0\\ 0 & -2\end{bmatrix}=\begin{bmatrix}3 & 3\\ 3 & 3\end{bmatrix}$. If the determinant of $(B-\lambda\cdot I)=0$ then λ is the eigen values. $det(\begin{bmatrix}3 & 3\\ 3 & 3\end{bmatrix})=(3)\cdot(3)-(3)\cdot(3)=9-9=0$.

We verified it is the eigen values, so let's find eigen vector x. For $\begin{bmatrix}x_1\\ x_2\end{bmatrix}$, let's say $x_1=1$. Then $1\cdot\begin{bmatrix}3\\ 3\end{bmatrix}+x_2\cdot\begin{bmatrix}3\\ 3\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$, $\begin{bmatrix}3+3x_2=0\\ 3+3x_2=0\end{bmatrix}$. So $x_2=-1$. We still need to normalize x_1 and x_2 to be length of 1. $\|x\|=\sqrt{x_1^2+x_2^2}=\sqrt{(1)^2+(-1)^2}=\sqrt{2}$. So x norm eigen vector is $\begin{bmatrix}\frac{1}{\sqrt{2}}\\ -\frac{1}{\sqrt{2}}\end{bmatrix}$

(iii) 1 and 0 We have $C\cdot x=\lambda\cdot x$. We can rewrite it to $(C-\lambda\cdot I)\cdot x=0$

For 1, $(C-\lambda\cdot I)=\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}-\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}0 & 0\\ 1 & -1\end{bmatrix}$. If the determinant of $(C-\lambda\cdot I)=0$ then λ is the eigen values. $det(\begin{bmatrix}0 & 0\\ 1 & -1\end{bmatrix})=0\cdot(-1)-0\cdot1=0$.

We verified it is the eigen values, so let's find eigen vector x. For $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, let's say $x_1 = 1$. Then $1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} + x_2 \cdot \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0+0=0 \\ 1-x_2=0 \end{bmatrix}$. So $x_2 = 1$. We still need to normalize x_1 and x_2 to be length of 1. $\|x\| = \sqrt{x_1^2 + x_2^2} = \sqrt{(1)^2 + (1)^2} = \sqrt{2}$. So x norm eigen vector is $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$

For 0, $(C - \lambda \cdot I) = C = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$. If the determinant of $C = 0$.

We verified it is the eigen values, so let's find eigen vector x. For $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, let's say $x_1 = 0$. Then $0 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 + 0 \cdot x_2 = 0 \\ 0 + 0 \cdot x_2 = 0 \end{bmatrix}$. So x_2 is all real number but in order to make $\|x\|$ length of 1, we can say $x_2 = 1$.So the vector x is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

2.3 Problem 2.3

Prove that a linear system Ax = b is consistent if and only if $rank(A) = rank([A|b])$. **Comment on the geometric interpretation of equation** $rank(A) = rank([A|b])$**.**

(i) A linear System $Ax = b$ **is consistent then** $rank(A) = rank([A|b])$

Let's say A is $m \times n$ matrix. When $Ax = b$ is consistent then A has pivots in every row. The columns of A span R^m. Columns of A span R^m means that the rank of (A) is m, and A's columns are linearly independent each others. Also $b \in$ span of (A). Since columns of A span R^m, adding b to the matrix A as an additional columns will still span R^m. Rank of A is m, and since b is in span of A, b is linearly dependent to columns of A. ∴ So the rank of ([A|b]) is the same as the rank of A.

(ii) $rank(A) = rank([A|b])$ **then A linear System** $Ax = b$ **is consistent.**

The rank of a matrix is the number of linearly independent columns. When $rank(A) = rank([A|b])$, b is linearly dependent on columns of A since the number of linearly independent columns stays the same. This means b must be in the span of A columns. b is in the span of A, then there exists an vector x that makes $Ax = b$ consistent.

3 Entropy

Consider two discrete random variable X and Y with the following joint distribution table:

	Y = 0	Y = 1	Y =2
X = 0	p_{00}	p_{01}	p_{02}
X = 1	p_{10}	p_{11}	p_{12}
X = 2	p_{20}	p_{21}	p_{22}

where p_{ij} **represents the probability that X = i and Y = j. Assume that the probabilities sum to 1.**

3.1 Problem 3.1

1. Consider the random variable Z = X|Y = 1. What is the distribution of Z?

We have a joint distribution table above. we can get the joint distribution table when Y = 1.

In according to the table, Pr(X=0|Y=1) = $\frac{p_{01}}{p_{01}+p_{11}+p_{21}}$, Pr(X=1|Y=1) = $\frac{p_{11}}{p_{01}+p_{11}+p_{21}}$, Pr(X=2|Y=1) = $\frac{p_{01}}{p_{01}+p_{11}+p_{21}}$

	Y = 1
X = 0	p_{01}
X = 1	p_{11}
X = 2	p_{21}
Total	$p_{01} + p_{11} + p_{21}$

3.2 Problem 3.2

2. Calculate the joint entropy H(X, Y).

The joint Entropy formula is: $-\sum_{x,y} p(x,y) \cdot log_2 p(x,y)$. Followed by the joint distribution table. we can get the calculation :

$$-p_{00} \cdot log_2(p_{00}) - p_{01} \cdot log_2(p_{01}) - p_{02} \cdot log_2(p_{02}) - p_{10} \cdot log_2(p_{10}) - p_{11} \cdot log_2(p_{11})$$
$$-p_{12} \cdot log_2(p_{12}) - p_{20} \cdot log_2(p_{20}) - p_{21} \cdot log_2(p_{21}) - p_{22} \cdot log_2(p_{22})$$

$$= -(p_{00} \cdot log_2(p_{00}) + p_{01} \cdot log_2(p_{01}) + p_{02} \cdot log_2(p_{02}) + p_{10} \cdot log_2(p_{10}) + p_{11} \cdot log_2(p_{11})$$
$$+p_{12} \cdot log_2(p_{12}) + p_{20} \cdot log_2(p_{20}) + p_{21} \cdot log_2(p_{21}) + p_{22} \cdot log_2(p_{22}))$$

3.3 Problem 3.3

3. Calculate the conditional entropies H(X|Y) and H(Y|X).

(1) Conditional Entropy H(Y|X)

The conditional Entropy H(Y|X)'s formula is $\sum_x P_x(X) \cdot H(Y|X = x)$

$$\sum_x P_x(X) \cdot H(Y|X = x) = \sum_x P_x(X) \cdot (-\sum_y P_{Y|X}(Y|X) log_2(P_{Y|X}(Y|X)))$$

so,

$$H(Y|X) = p_x(0) \cdot (H(Y = 0|X = 0) + H(Y = 1|X = 0) + H(Y = 2|X = 0))$$
$$+p_x(1) \cdot (H(Y = 0|X = 1) + H(Y = 1|X = 1) + H(Y = 2|X = 1))$$
$$+p_x(2) \cdot (H(Y = 0|X = 2) + H(Y = 1|X = 2) + H(Y = 2|X = 2))$$

$$= [(p_{00} + p_{01} + p_{02}) \cdot -((\frac{p_{00}}{p_{00} + p_{01} + p_{02}}) log_2(\frac{p_{00} + p_{01} + p_{02}}{p_{00}})$$
$$+(\frac{p_{01}}{p_{00} + p_{01} + p_{02}}) log_2(\frac{p_{00} + p_{01} + p_{02}}{p_{01}})$$
$$+(\frac{p_{02}}{p_{00} + p_{01} + p_{02}}) log_2(\frac{p_{00} + p_{01} + p_{02}}{p_{02}}))$$

$$+(p_{10} + p_{11} + p_{12}) \cdot -((\frac{p_{10}}{p_{10} + p_{11} + p_{12}}) log_2(\frac{p_{10} + p_{11} + p_{12}}{p_{10}})$$
$$+(\frac{p_{11}}{p_{10} + p_{11} + p_{12}}) log_2(\frac{p_{10} + p_{11} + p_{12}}{p_{11}})$$
$$+(\frac{p_{12}}{p_{10} + p_{11} + p_{12}}) log_2(\frac{p_{10} + p_{11} + p_{12}}{p_{12}}))$$

$$+(p_{20} + p_{21} + p_{22}) \cdot -((\frac{p_{20}}{p_{20} + p_{21} + p_{22}}) log_2(\frac{p_{20} + p_{21} + p_2}{p_{20}})$$

$$+(\frac{p_{21}}{p_{20}+p_{21}+p_{22}})log_2(\frac{p_{20}+p_{21}+p_{22}}{p_{21}})$$
$$+(\frac{p_{22}}{p_{20}+p_{21}+p_{22}})log_2(\frac{p_{20}+p_{21}+p_{22}}{p_{22}}))]$$

$$= -(p_{00}\cdot log_2(\frac{p_{00}+p_{01}+p_{02}}{p_{00}}) + p_{01}\cdot log_2(\frac{p_{00}+p_{01}+p_{02}}{p_{01}}) + p_{02}\cdot log_2(\frac{p_{00}+p_{01}+p_{02}}{p_{02}})$$
$$+p_{10}\cdot log_2(\frac{p_{10}+p_{11}+p_{12}}{p_{10}}) + p_{11}\cdot log_2(\frac{p_{10}+p_{11}+p_{12}}{p_{11}}) + p_{12}\cdot log_2(\frac{p_{10}+p_{11}+p_{12}}{p_{12}})$$
$$+p_{20}\cdot log_2(\frac{p_{20}+p_{21}+p_{22}}{p_{20}}) + p_{21}\cdot log_2(\frac{p_{20}+p_{21}+p_{22}}{p_{21}}) + p_{22}\cdot log_2(\frac{p_{20}+p_{21}+p_{22}}{p_{22}}))$$

$$= p_{00}\cdot log_2(\frac{p_{00}}{p_{00}+p_{01}+p_{02}}) + p_{01}\cdot log_2(\frac{p_{01}}{p_{00}+p_{01}+p_{02}}) + p_{02}\cdot log_2(\frac{p_{02}}{p_{00}+p_{01}+p_{02}})$$
$$+p_{10}\cdot log_2(\frac{p_{10}}{p_{10}+p_{11}+p_{12}}) + p_{11}\cdot log_2(\frac{p_{11}}{p_{10}+p_{11}+p_{12}}) + p_{12}\cdot log_2(\frac{p_{12}}{p_{10}+p_{11}+p_{12}})$$
$$+p_{20}\cdot log_2(\frac{p_{20}}{p_{20}+p_{21}+p_{22}}) + p_{21}\cdot log_2(\frac{p_{21}}{p_{20}+p_{21}+p_{22}}) + p_{22}\cdot log_2(\frac{p_{22}}{p_{20}+p_{21}+p_{22}})$$

(2) Conditional Entropy H(X|Y)

The conditional Entropy H(X|Y)'s formula is the same as the previous questions but need to change the conditional variable to H(X|Y). So H(X|Y)'s formula is $\sum_y P_y(Y)\cdot H(X|Y=y)$

$$\sum_y P_y(Y)\cdot H(X|Y=y) = \sum_y P_y(X)\cdot(-\sum_x P_{X|Y}(X|Y)log_2(P_{X|Y}(X|Y)))$$

Using the previous equation, the Conditional Entropy H(X|Y) is

$$= p_{00}\cdot log_2(\frac{p_{00}}{p_{00}+p_{10}+p_{20}}) + p_{10}\cdot log_2(\frac{p_{10}}{p_{00}+p_{10}+p_{20}}) + p_{20}\cdot log_2(\frac{p_{20}}{p_{00}+p_{10}+p_{20}})$$
$$+p_{01}\cdot log_2(\frac{p_{01}}{p_{01}+p_{11}+p_{21}}) + p_{11}\cdot log_2(\frac{p_{11}}{p_{01}+p_{11}+p_{21}}) + p_{21}\cdot log_2(\frac{p_{21}}{p_{01}+p_{11}+p_{21}})$$
$$+p_{02}\cdot log_2(\frac{p_{02}}{p_{02}+p_{12}+p_{22}}) + p_{12}\cdot log_2(\frac{p_{12}}{p_{02}+p_{12}+p_{22}}) + p_{22}\cdot log_2(\frac{p_{22}}{p_{02}+p_{12}+p_{22}})$$

3.4 Problem 3.4

4. Calculate the marginal entropy H(X) and H(Y) of each variable.

Before get two marginal entropys, Let's see the marginal distribution of x and y tables.

	Total
X = 0	$p_{00}+p_{01}+p_{02}$
X = 1	$p_{10}+p_{11}+p_{12}$
X = 2	$p_{20}+p_{21}+p_{22}$

	Y =0	Y = 1	Y = 2
Total	$p_{00}+p_{10}+p_{20}$	$p_{01}+p_{11}+p_{21}$	$p_{02}+p_{12}+p_{22}$

The entropy formula of x is

$$H(x) = -\sum_x p(x)log_2p(x)$$

so H(X) $= -((p_{00} + p_{01} + p_{02})log_2(p_{00} + p_{01} + p_{02}) + (p_{10} + p_{11} + p_{12})log_2(p_{10} + p_{11} + p_{12}) + (p_{20} + p_{21} + p_{22})log_2(p_{20} + p_{21} + p_{22}))$.

Similarly, H(Y) $= -((p_{00} + p_{10} + p_{20})log_2(p_{00} + p_{10} + p_{20}) + (p_{01} + p_{11} + p_{21})log_2(p_{01} + p_{11} + p_{21}) + (p_{02} + p_{12} + p_{22})log_2(p_{02} + p_{12} + p_{22}))$.

Submitted by Jae Hong Lee on March 1, 2024.