Natural Language Processing
Assignment 2: Probability and Vector Exercises

Li-Yi Lin, JHED ID: 5F0CE8

Question 1.
(a) Answer:
Since $Y \subseteq Z$, we have $Z = Y \cup (Z \cap \neg Y)$. So the probability of Z is

$$\begin{aligned} p(Z) &= p(Y \cup (Z \cap \neg Y)) \\ &= p(Y) + p(\emptyset) \\ &\geq p(Y) + 0 && (p(\emptyset) \geq 0) \\ &= p(Y) \end{aligned}$$

Thus, we proved that if $Y \subseteq Z$, then $p(Y) \leq p(Z)$.

(b) Answer:
We know that $p(X \mid Z) = \frac{p(X \cap Z)}{p(Z)}$. Since $(X \cap Z) \subseteq Z$, we have the fact that $p(X \cap Z) \leq p(Z)$. Therefore, $p(X \mid Z) = \frac{p(X \cap Z)}{p(Z)} \leq 1$. Moreover, since probability is larger than or equal to 0, we have $p(X \cap Z) \geq 0$. Because $p(X \mid Z)$ is conditioned on Z, we have $p(Z) > 0$. Therefore, by the property that $0 \leq p(X \cap Z) \leq p(Z)$, we can prove that $p(X \mid Z)$ always fall in the range [0,1].

(c) Answer:
Since $E \cap \emptyset = \emptyset$, we have $p(E \cup \emptyset) = p(E) + p(\emptyset)$ by the given axioms. Because we know that $p(E) = 1$ and probability cannot be larger than 1, $p(E \cup \emptyset) = p(E) + p(\emptyset) = 1 + p(\emptyset) \leq 1$. Therefore $p(\emptyset)$ must be 0.

(d) Answer:
Let $\bar{X}$ denote $E - X$. We have $\bar{X} \cup X = E$. Now we are going to prove that $p(X) = 1 - p(\bar{X})$.
Since $\bar{X} \cap X = \emptyset$ and $\bar{X} \cup X = E$, we have

$$\begin{aligned} p(E) &= p(\bar{X} \cup X) \\ &= p(\bar{X}) + p(X) - p(\bar{X} \cap X) \\ &= p(\bar{X}) + p(X) && (p(\bar{X} \cap X) = 0 \text{ by (c)}) \end{aligned}$$

By the equation above, we have $p(E) = p(\bar{X}) + p(X)$. Since $p(E) = 1$, we have $1 = p(\bar{X}) + p(X)$. Therefore we have proven that $p(X) = 1 - P(\bar{X})$.

(e) Answer:

$$\begin{aligned} p(singing \cap rainy \mid rainy) &= \frac{p((singing \cap rainy) \cap rainy)}{p(rainy)} \\ &= \frac{p(singing \cap (rainy \cap rainy))}{p(rainy)} && \text{(Intersection is an associative operation)} \\ &= \frac{p(singing \cap rainy)}{p(rainy)} && (rainy \cap rainy = rainy) \\ &= p(singing \mid rainy) \end{aligned}$$

By the equation above, we proved that $p(singing \text{ AND } rainy \mid rainy) = p(singing \mid rainy)$.

(f) Answer:

$$\begin{aligned} p(Y) &= p((X \cap Y) \cup (\bar{X} \cap Y)) \\ &= p(X \cap Y) + p(\bar{X} \cap Y) - p((X \cap Y) \cap (\bar{X} \cap Y)) \\ &= p(X \cap Y) + p(\bar{X} \cap Y) && ((X \cap Y) \cap (\bar{X} \cap Y) = \emptyset) \end{aligned}$$

By dividing both sides by $p(Y)$, we proved that

$$1 = \frac{p(X \cap Y)}{p(Y)} + \frac{p(\bar{X} \cap Y)}{p(Y)}$$
$$p(X \mid Y) = 1 - p(\bar{X} \mid Y)$$

(g) Answer:

$$\begin{aligned} &(p(X \mid Y) \cdot p(Y) + p(X \mid \bar{Y}) \cdot p(\bar{Y})) \cdot p(\bar{Z} \mid X)/p(\bar{Z}) \\ &= (\frac{p(X \cap Y)}{p(Y)} \cdot p(Y) + \frac{p(X \cap \bar{Y})}{p(\bar{Y})} \cdot p(\bar{Y})) \cdot \frac{p(\bar{Z} \cap X)}{p(X)}/p(\bar{Z}) \\ &= (p(X \cap Y) + p(X \cap \bar{Y})) \cdot \frac{p(\bar{Z} \cap X)}{p(X)}/p(\bar{Z}) \\ &= p(X) \cdot \frac{p(\bar{Z} \cap X)}{p(X)}/p(\bar{Z}) \\ &= \frac{p(\bar{Z} \cap X)}{p(\bar{Z})} \\ &= p(X \mid \bar{Z}) \end{aligned}$$

(h) Answer:
We know that

$$p(singing \text{ AND } rainy) = p(singing) + p(rainy) - p(singing \cap rainy)$$

So, when $(singing \cap rainy) = \emptyset$, then

$$\begin{aligned} p(singing \text{ AND } rainy) &= p(singing) + p(rainy) - p(singing \cap rainy) \\ &= p(singing) + p(rainy) - p(\emptyset) \\ &= p(singing) + p(rainy) - 0 && \text{(by (c))} \\ &= p(singing) + p(rainy) \end{aligned}$$

Therefore, when set *singing* and *rainy* are disjoint, the given equation is true.

(i) Answer:
Assume the given equation is true, then

$$p(singing \text{ AND } rainy) = p(singing) \cdot p(rainy)$$

We further assume that $p(rainy) > 0$ and divide the above equation by $p(rainy)$

$$\frac{p(singing \text{ AND } rainy)}{p(rainy)} = p(singing)$$

$$p(singing \mid rainy) = p(singing)$$

Therefore, for the $p(singing \text{ AND } rainy) = p(singing) \cdot p(rainy)$ to be true, $p(singing \mid rainy) = p(singing)$ must also be true. It means that the probability event *rainy* will not affect the probability of event *singing*.

(j) Answer:
We are given that

$$p(X \mid Y) = 0$$
$$\frac{p(X \cap Y)}{p(Y)} = 0$$

For the above equation to be 0, $p(X \cap Y)$ must be 0.

Then, we are going to prove that $p(X \mid Y, Z) = 0$

$$\begin{aligned} p(X \mid Y, Z) &= \frac{p(X \cap Y \cap Z)}{p(Y \cap Z)} \\ &\leq \frac{p(X \cap Y)}{p(Y \cap Z)} && \text{(by (a), } (X \cap Y \cap Z) \subseteq (X \cap Y)\text{, so } p(X \cap Y \cap Z) \leq p(X \cap Y)) \\ &= \frac{0}{p(Y \cap Z)} \\ &= 0 \end{aligned}$$

Thus, we have proved that if $p(X \mid Y) = 0$, then $p(X \mid Y, Z) = 0$.

(k) Answer:
By (f), we know

$$\begin{aligned} p(\bar{W} \mid Y) &= 1 - p(W \mid Y) \\ &= 1 - 1 \\ &= 0 \end{aligned}$$

Since $p(\bar{W} \mid Y) = \frac{p(\bar{W} \cap Y)}{p(Y)} = 0$, then $p(\bar{W} \mid Y)$ must be 0.

Then, we are going to prove that $p(W \mid Y, Z) = 1$

$$\begin{aligned} p(W \mid Y, Z) &= 1 - p(\bar{W} \mid Y, Z) && \text{(by (f))} \\ &= 1 - \frac{p(\bar{W} \cap Y \cap Z)}{p(Y \cap Z)} \\ &= 1 - 0 && \text{(since } (\bar{W} \cap Y \cap Z) \subseteq (\bar{W} \cap Y), p(\bar{W} \cap Y \cap Z) \leq p(\bar{W} \cap Y) \text{ by (a))} \\ &= 1 \end{aligned}$$

Thus, we have proven that if $p(W \mid Y) = 1$, then $p(W \mid Y, Z) = 1$.

Question 2. (a) Answer:
Denote event "$Actual = Blue$" as A, "$Claimed = blue$" as C.
By Bayes' theorem, we have

$$p(A \mid C)p(C) = p(C \mid A)p(A)$$

(b) Answer:
Since $p(Actual = blue) = 0.1$ is fixed. We further rewrite the equation in (a) as shown below:

$$p(Actual = blue \mid Claimed = blue) \propto p(Claimed = blue \mid Actual = blue)p(Actual = blue)$$

Then, the prior probability is $p(Actual = blue)$,
the likelihood of the evidence is $p(Claimed = blue \mid Actual = blue)$,
and the posterior probability is $p(Actual = blue \mid Claimed = blue)$.

(c) Answer:
The judge should care about the posterior because it takes not only about the observed events, likelihood, but also the belief about the probability of the event we want to guess. By considering the posterior probability, the judge can avoid the situation that our belief is wrong or we have not enough observed events so that the judge can have more accurate information for judgement.

(d) Answer: