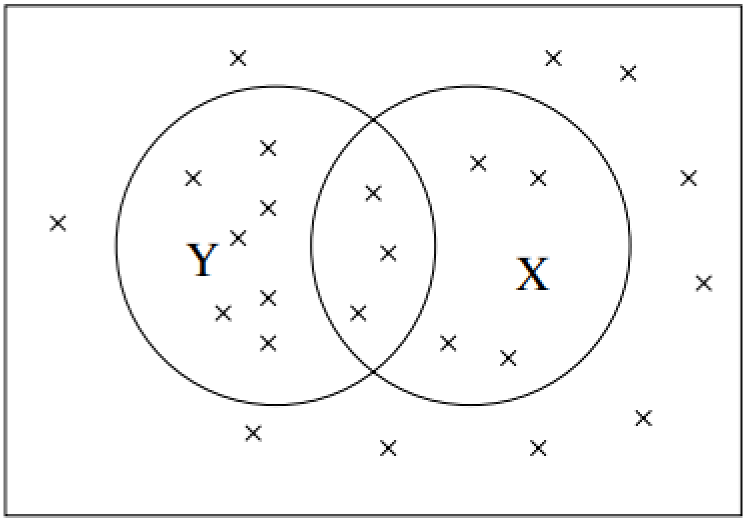
INSTRUCTIONS

- **Due: Wednesday, 15 Feb 2023 at 23:59 PM PDT** Remember that you have 5 slip days to use at your discretion BUT you can use no more than 1 per homework.
- **Format:** Submit the answer sheet containing your answers in PDF. You can provide either typed or handwritten answers to this homework. Handwritten answers must be legible and scanned into a PDF.
- **Note: Please DO NOT FORGET to include your name and WSU ID in your submission.**
- **How to submit:** Submit a PDF containing your answers on Canvas
- **Policy:** See the course website for homework policies and Academic Integrity.

Last Name	
First Name	
WSU ID	

Q1. [20 pts] Probability

Based on the following Venn diagram, answer the following questions:



	X	$\neg X$
Y	$\frac{3}{24}$	
$\neg Y$		

Figure 1: Venn diagram

(a) [5 pts] Complete the joint probability distribution in the above table.

Answer:
There are 24 points in the diagram.
7 of which belong to Y but not X, thus $P(\neg X, Y) = 7/24$.
4 of which belong to X but not Y, thus $P(X, \neg Y) = 4/24$.
10 of which belong to neither X nor Y, thus $P(\neg X, \neg Y) = 10/24$.

(b) [12 pts] Based on the above joint probability distribution, find the following: $P(X)$, $P(Y)$, $P(\neg X)$, $P(\neg Y)$, $P(X|Y)$, $P(Y|X)$, $P(X|\neg Y)$, $P(Y|\neg X)$, $P(\neg Y|X)$, $P(\neg X|Y)$, $P(\neg X|\neg Y)$ and $P(\neg Y|\neg X)$.

Answer:
$P(X) = P(X \wedge Y) + P(X \wedge \neg Y) = 3/24 + 4/24 = 7/24$
$P(Y) = P(Y \wedge X) + P(Y \wedge \neg X) = 3/24 + 7/24 = 10/24 = 5/12$
$P(\neg X) = P(\neg X \wedge Y) + P(\neg X \wedge \neg Y) = 7/24 + 10/24 = 17/24$
$P(\neg Y) = P(\neg Y \wedge X) + P(\neg Y \wedge \neg X) = 4/24 + 10/24 = 14/24 = 7/12$
$P(X \mid Y) = P(X \wedge Y)/P(Y) = (3/24)/(5/12) = 3/10$
$P(Y \mid X) = P(X \wedge Y)/P(X) = (3/24)/(7/24) = 3/7$
$P(X \mid \neg Y) = P(X \wedge \neg Y)/P(\neg Y) = (4/24)/(14/24) = 4/14 = 2/7$
$P(\neg Y \mid X) = P(X \wedge \neg Y)/P(X) = (4/24)/(7/24) = 4/7$
$P(\neg X \mid Y) = P(\neg X \wedge Y)/P(Y) = (7/24)/(10/24) = 7/10$
$P(Y \mid \neg X) = P(\neg X \wedge Y)/P(\neg X) = (7/24)/(17/24) = 7/17$
$P(\neg X \mid \neg Y) = P(\neg X \wedge \neg Y)/P(\neg Y) = (10/24)/(14/24) = 5/7$
$P(\neg Y \mid \neg X) = P(\neg X \wedge \neg Y)/P(\neg X) = (10/24)/(17/24) = 10/17$

(c) [3 pts] Prove the followings using only definition of conditional and marginal probability:

$$\begin{aligned} P(X|Y) &= 1 - P(\neg X|Y) \\ P(X|\neg Y) &= 1 - P(\neg X|\neg Y) \\ P(\neg Y|X) &= \frac{P(X|\neg Y)P(\neg Y)}{P(X|\neg Y)P(\neg Y) + P(X|Y)P(Y)} \end{aligned}$$

Answer:
By definition, we have:

$$P(\neg X \mid Y) + P(X \mid Y) = \frac{P(\neg X \wedge Y)}{P(Y)} + \frac{P(X \wedge Y)}{P(Y)} = \frac{P(Y)}{P(Y)} = 1$$

Thus, re-arranging shows that $P(X \mid Y) = 1 - P(\neg X \mid Y)$.
Using similar argument, we also have $P(X \mid \neg Y) = 1 - P(\neg X \mid \neg Y)$.
Finally, we derive the last equality as follows,

$$\begin{aligned} P(\neg Y \mid X) &= \frac{P(\neg Y \wedge X)}{P(X)} = \frac{P(X \mid \neg Y)P(\neg Y)}{P(X \wedge Y) + P(X \wedge \neg Y)} \\ &= \frac{P(X \mid \neg Y)P(\neg Y)}{P(X \mid Y)P(Y) + P(X \mid \neg Y)P(\neg Y)} \end{aligned}$$

Q2. [20 pts] Probabilistic Reasoning

Assume that 2% of the population in a country carry a particular virus. A test kit developed by a pharmaceutical firm is able to detect the presence of the virus from a patient's blood sample. The firm claims that the test kit has a high accuracy of detection in terms of the following conditional probabilities obtained from their quality test:

- P(the kit shows positive | the patient is a carrier) = 0.998
- P(the kit shows negative | the patient is not a carrier) = 0.996

(a) [10 pts] Given that a patient is tested to be positive using this kit, what is the probability that he is not a carrier? Give your answer to 3 decimal places.

Answer:
Let X and $\neg X$ denote the events where the test shows positive and negative, respectively. Let Y and $\neg Y$ denote the events where the patient is a carrier and not a carrier, respectively. Then,

$$\begin{aligned} P(Y) &= 0.02 \quad \text{(as given above)} \Rightarrow P(\neg Y) = 1 - 0.02 = 0.98 \\ P(X \mid Y) &= 0.998 \quad \text{(as given above)} \\ P(\neg X \mid \neg Y) &= 0.996 \quad \text{(as given above)} \Rightarrow P(X \mid \neg Y) = 1 - 0.996 = 0.004 \end{aligned}$$

Thus, applying Bayes rule,

$$\begin{aligned} P(\neg Y \mid X) &= \frac{P(X \mid \neg Y)P(\neg Y)}{P(X)} = \frac{P(X \mid \neg Y)P(\neg Y)}{P(X \mid Y)P(Y) + P(X \mid \neg Y)P(\neg Y)} \\ &= \frac{0.004 \times 0.98}{0.998 \times 0.02 + 0.004 \times 0.98} \simeq 0.164 \end{aligned}$$

There is an approximately 16.4% chance that when the test shows positive, the patient is not a carrier.

(b) [10 pts] Suppose that the patient does not entirely trust the result offered by the first kit (perhaps because it has expired) and decides to use another test kit. If the patient is again tested to be positive using this second kit, what is the (updated) likelihood that he is not a carrier? You can assume conditional independence between results of different test kits given the patient's state of virus contraction. Give your answer to 4 decimal places.

Answer:
Let the two test events be X_1 and X_2, which are independent instantiations of the original text X,

$$\begin{aligned} P(\neg Y \mid X_1 \wedge X_2) &= \frac{P(\neg Y \wedge X_1 \wedge X_2)}{P(X_1 \wedge X_2)} = \frac{P(X_1 \mid \neg Y \wedge X_2)P(X_2 \mid \neg Y)P(\neg Y)}{P(X_1 \wedge X_2)} && (1) \\ &= \frac{P(X_1 \mid \neg Y)P(X_2 \mid \neg Y)P(\neg Y)}{P(X_1 \wedge X_2)} && (2) \end{aligned}$$

The transition from Eq. (1) to Eq. (2) is due to the fact that X_2 is conditionally independent of X_1 given Y, so $P(X_2 \mid \neg Y \wedge X_1) = P(X_2 \mid \neg Y)$. Next, by the law of total probability:

$$\begin{aligned} \frac{P(X_1 \mid \neg Y)P(X_2 \mid \neg Y)P(\neg Y)}{P(X_1 \wedge X_2)} &= \frac{P(X_1 \mid \neg Y)P(X_2 \mid \neg Y)P(\neg Y)}{P(X_1 \wedge X_2 \wedge Y) + P(X_1 \wedge X_2 \wedge \neg Y)} && (3) \\ &= \frac{P(X_1 \mid \neg Y)P(X_2 \mid \neg Y)P(\neg Y)}{P(X_1 \mid Y)P(X_2 \mid Y)P(Y) + P(X_1 \mid \neg Y)P(X_2 \mid \neg Y)P(\neg Y)} && (4) \end{aligned}$$

where the transition from Eq. (3) to Eq. (4) is again due to the fact that X_1 and X_2 are conditional if we know Y. Finally, we note that X_1 and X_2 are identically distributed with X, which means

$$\frac{P(X_1 \mid \neg Y)P(X_2 \mid \neg Y)P(\neg Y)}{P(X_1 \mid Y)P(X_2 \mid Y)P(Y) + P(X_1 \mid \neg Y)P(X_2 \mid \neg Y)P(\neg Y)} = \frac{P(X \mid \neg Y)^2 P(\neg Y)}{P(X \mid Y)^2 P(Y) + P(X \mid \neg Y)^2 P(\neg Y)}$$

Plugging in the corresponding values, the above evaluate to 0.0008.

Q3. [25 pts] Conditional Probability

Prove the following:

(a) [10 pts] (a) $P(a|a \wedge b) = 1$ from using only the definition of conditional probability.

Answer:
Using the definition of conditional probability,

$$P(a \mid a \wedge b) \;=\; P(a \wedge a \wedge b)/P(a \wedge b) \;=\; P(a \wedge b)/P(a \wedge b) \;=\; 1 \tag{5}$$

(b) [15 pts] (b) Show that (1) $P(a|b) = P(a)$, (2) $P(b|a) = P(b)$ and (3) $P(a \wedge b) = P(a)P(b)$ are equivalent. That is, any two of these statements are equivalent.

Answer:
To show their equivalence, we will show that (1) implies (2), (2) implies (3) and (3) implies (1). First, given (1), $P(a \mid b) = P(a)$ which implies

$$P(b \mid a) = P(a \mid b)P(b)/P(a) \;=\; P(b)$$

Thus, (1) indeed implies (2). Next, given (2), $P(b \mid a) = P(b)$ which implies,

$$P(a \wedge b) \;=\; P(b \mid a)P(a) \;=\; P(a)P(b)$$

Hence, we have shown that (2) implies (3). Finally, given (3), $P(a \wedge b) = P(a)P(b)$, it follows that

$$P(a \mid b) \;=\; P(a \wedge b)/P(b) \;=\; P(a)P(b)/P(b) \;=\; P(a)$$

This means that (3) implies (1). Putting these together, we have (1), (2) and (3) are equivalence.

Q4. [35 pts] Bayesian Network

Given the following conditional table:

$P(WetGrass \mid Sprinkler \wedge Rain)$	0.95
$P(WetGrass \mid Sprinkler \wedge \neg Rain)$	0.9
$P(WetGrass \mid \neg Sprinkler \wedge Rain)$	0.8
$P(WetGrass \mid \neg Sprinkler \wedge \neg Rain)$	0.1
$P(Sprinkler \mid RainySeason)$	0.0
$P(Sprinkler \mid \neg RainySeason)$	1.0
$P(Rain \mid RainySeason)$	0.9
$P(Rain \mid \neg RainySeason)$	0.1
$P(RainySeason)$	0.7

(a) [5 pts] Show that $P(S) = P(\neg RS)$ or in other words, $S \equiv \neg RS$.

Answer:
Let RS, S, R and WG denote RainySeason, Sprinkler, Rain and WetGrass, respectively. To simplify the Bayesian network, we first show $S \equiv \neg RS$ or equivalently, $P(S) = P(\neg RS)$. To see this,

$$P(S) \quad = \quad P(S \wedge RS) + P(S \wedge \neg RS) = P(S \mid RS)P(RS) + P(S \mid \neg RS)P(\neg RS)$$

Plugging in the facts that $P(S \mid \neg RS) = 1$ and $P(S \mid RS) = 0$ reduces the above to $P(S) = P(\neg RS)$.

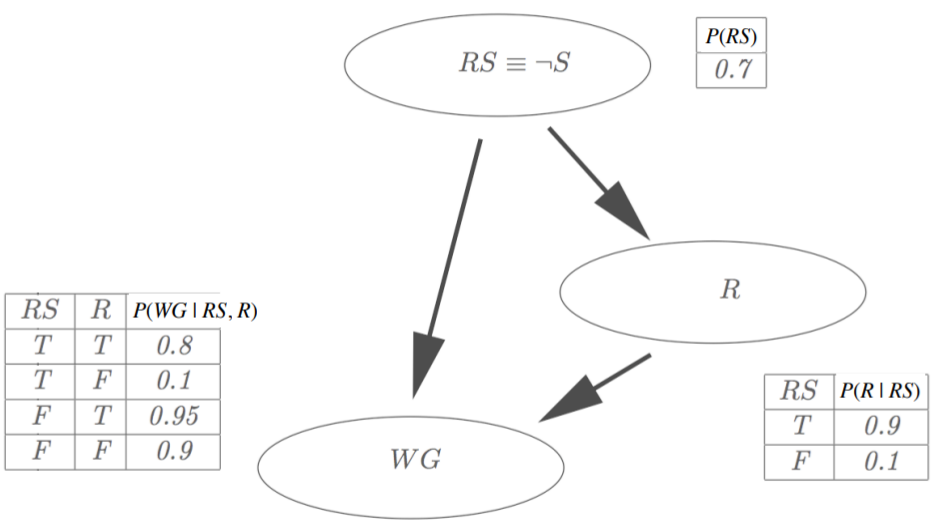


Figure 2: Bayesian Network

(b) [20 pts] Construct a Bayesian Network (BN) with as few parameters as possible (hint: use result of part (a)). You will only get 50% credit for this question if the BN is not optimal in the number of parameters.

Answer:
Using result of part (a), $S \equiv \neg RS$ the Bayesian Network only need to characterize $P(RS)$, $P(WG \mid RS, R)$ and $P(R \mid RS)$ – see Fig. 2 where we can use S and $\neg RS$ interchangeably.

(c) [10 pts] Compute the following probability:

$$P(WetGrass \wedge RainySeason \wedge \neg Rain \wedge \neg Sprinkler)$$

Answer:
Finally, the probability $P(WG \wedge RS \wedge \neg R \wedge \neg S) = P(WG \mid RS \wedge \neg R \wedge \neg S)P(\neg R \mid RS \wedge \neg S)P(RS \wedge \neg S)$.

Now, since $RS \equiv \neg S$, the above is equal to $P(WG \mid RS \wedge \neg R)P(\neg R \mid RS)P(RS) = 0.1 \times (1 - 0.9) \times 0.7 = 0.007$.