HW01_ Probability_0

Sri Jay Adarsh Gogineni
CWID: 20025484

1.1

$P(A)$ = Jerry at bank = 20% = 0.2
$P(B)$ = Susan at bank = 30% = 0.3
$P(A \text{ and } B)$ = Both at bank = 8% = 0.08

a) $P(A|B) \rightarrow$ probability that Jerry at bank given Susan at bank

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{0.08}{0.3} = .267 \rightarrow \boxed{26.7\%}$$

b) $P(\bar{B})$ = Susan not at bank = $1 - P(B) = 0.70$

$P(A|\bar{B}) \rightarrow$ Jerry at bank given Susan not at bank

$$P(A|\bar{B}) = \frac{P(A \text{ and } \bar{B})}{P(\bar{B})} = \frac{P(A) - P(A \cap B)}{P(\bar{B})} = \frac{0.2 - 0.08}{0.7} = 0.1714 = \boxed{17.14\%}$$

c) $P(A \cap B)$ = both at bank
$P(A \cup B)$ = at least one at bank

$P((A \cap B) | (A \cup B))$ = both at bank given at least one at bank

$$= \frac{P((A \cap B) \cap (A \cup B))}{P(A \cup B)} = \frac{P(A \cap B)}{P(A \cup B)} = \frac{0.08}{P(A) + P(B) - P(A \cap B)}$$

$$= \frac{0.08}{0.2 + 0.3 - 0.08} = \frac{0.08}{.42}$$

$$= 0.1904 = \boxed{19.04\%}$$

1.2

$P(H)$ = Harold's chances of getting "B" = 80% = 0.8
$P(S)$ = Sharon's chances of getting "B" = 90% = 0.9
$P(H \cup S)$ = At least one getting "B" = 91% = 0.91

a) $P(H \cap \bar{S})$ = only Harold gets "B"

$$P(H \cap \bar{S}) = \underbrace{P(H)}_{\text{Harold gets B}} - \underbrace{P(H \cap S)}_{\text{Both get B}}$$

$$= P(H) - [P(H) + P(S) - P(H \cup S)]$$

$$0.8 - [0.8 + 0.9 - 0.91] = 0.01 = \boxed{1\%}$$

b) $P(S \cap \bar{H})$ = only Sharon gets "B"

$$P(S \cap \bar{H}) = P(S) - P(H \cap S)$$

$$= P(S) - [P(H) + P(S) - P(H \cup S)]$$

$$= 0.9 - [0.8 + 0.9 - 0.91] = 0.11 = \boxed{11\%}$$

c) $P(\overline{H \cup S})$ = both don't get "B"

$$= 1 - \underbrace{P(H \cup S)}_{\text{at least one gets "B"}} = 1 - 0.91 = 0.09 = \boxed{9\%}$$

1.3

$$P(J) = 20\% = 0.2$$
$$P(S) = 30\% = 0.3$$
$$P(J \cap S) = 8\% = 0.08$$

a) Are "Jerry at bank" and "Susan at bank" Independent?

↳ Two events $P(J)$ & $P(S)$ are independent if $P(J \cap S) = P(J) \cdot P(S)$

↳ $P(J) \cdot P(S) = (0.2)(0.3) = 0.06 \neq \underbrace{0.08}_{P(J \cap S)}$

$\boxed{\text{Since } P(J \cap S) \neq P(J) \cdot P(S) \text{, events Jerry at the bank and Susan at bank are NOT independent}}$

1.4

- Rolls 2 dice

a) Are "the sum is 6" and "second die shows 5" independent?

$P(A) = \text{sum is } 6 = \frac{5}{36}$

$P(B) = \text{second die shows } 5 = \frac{6}{36} = \frac{1}{6}$

$P(A \cap B) = \frac{1}{36}$

$P(A)P(B) = \frac{5}{36} \times \frac{1}{6} = \frac{5}{216}$

$P(A \cap B) \neq P(A)P(B)$

↳ They are dependent events

b) Are the events "the sum is 7" and "first die show 5" independent?

$P(A) = \text{Sum is } 7 = \frac{6}{36} = \frac{1}{6}$

$P(B) = \text{first die show } 5 = \frac{6}{36} = \frac{1}{6}$

$P(A \cap B) = \frac{1}{36}$

$P(A)P(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

$P(A \cap B) = P(A)P(B)$

↳ Independent Events

1.5

$P(Tx) = 60\%$ $\quad P(oil \mid Tx) = 30\%$

$P(Ak) = 30\%$ $\quad P(oil \mid Ak) = 20\%$

$P(Nj) = 10\%$ $\quad P(oil \mid Nj) = 10\%$

1.) Probability of Finding Oil?

$$P(oil) = P(oil \mid Tx) \cdot P(Tx) + P(oil \mid Ak) \cdot P(Ak) + P(oil \mid Nj) \cdot P(Nj)$$
$$= (30\%)(60\%) + (20\%)(30\%) + (10\%)(10\%)$$

$$P(oil) = 18\% + 6\% + 1\% = \boxed{25\%}$$

2.) Company drilled & found oil. Probability that they drilled in Tx?

$$P(Tx \mid oil) \cdot P(oil) = P(oil \mid Tx) \cdot P(Tx)$$

$$P(Tx \mid oil) = \frac{P(oil \mid Tx)\, P(Tx)}{P(oil)} = \frac{(30\%)(60\%)}{(25\%)} = \boxed{72\%}$$

1.6

- What is the probability that a passenger did not survive?

$$P(\text{Not Survived}) = \frac{1490}{2201} = \boxed{67.69\%}$$

- What is the probability that a passenger was staying in the first class?

$$P(1^{st}\text{class}) = \frac{325}{2201} = \boxed{14.76\%}$$

- Given that a passenger survived, what is the probability that the passenger was staying in the first class?

$$P(1^{st} \mid S) = \frac{P(1^{st} \cap S)}{P(S)} = \frac{203}{711} = \boxed{28.55\%}$$

- Are survival and staying in the first class independent?

$$P(S) = \frac{711}{2201} \qquad P(1^{st}) = \frac{325}{2201}$$

$$P(S) \cdot P(1^{st}) = \frac{711}{2201} \times \frac{325}{2201} = 4.77\%$$

$$P(S) \cap P(1^{st}) = \frac{203}{2201} = 9.22\%$$

$\neq$ $\boxed{\text{Not Independent Events}}$

- Given that a passenger survived, what is the probability that the passenger was staying in the first class and the passenger was a child?

$$P(1^{st} \cap C \mid \text{Survived}) = \frac{P(1^{st} \cap C \cap \text{Survived})}{P(\text{survived})} = \frac{6}{711} = \boxed{0.84\%}$$

- Given that a passenger survived, what is the probability that the passenger was an adult?

$$P(\text{Adult} \mid \text{Survived}) = \frac{P(\text{Adult} \cap \text{Survived})}{P(\text{survived})} = \frac{654}{711} = \boxed{91.98\%}$$

- Given that a passenger survived, are age and staying in the first class independent?

$$P(\text{Adult} \cup \text{Child} \mid \text{Survived}) = \frac{711}{711} \qquad P(1^{st} \mid \text{Survived}) = \frac{203}{711}$$

$$\hookrightarrow P(A \cup C \mid S) \times P(1^{st} \mid S) = \frac{711}{711} \times \frac{203}{711} = 28.55\%$$

$$P((A \cup C \mid S) \cap (1^{st} \mid S)) = \frac{P((A \cup C) \cap 1^{st}) \mid S)}{P(S)} = \frac{203}{711} = 28.55\%$$

$=$ $\boxed{\text{Events are Independent}}$

A developer claims that her app can distinguish AI-generated documents from human-generated ones. To assess its performance, we have submitted 1000 AI-generated and 1000 human-generated documents to the app.

- The app misclassified 70 human-generated documents as AI-generated
- and 30 AI generated documents as human- generated.

Build the confusion matrix for the above app and calculate the following:

Accuracy, precision, recall and F1

	Predicted AI	Predicted Human	Total
Actual AI	970 TP	30 FP	1000
Actual Human	70 FN	930 TN	1000
Total	1000	1000	

1. $\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} = \frac{970+930}{2000} = \frac{1900}{2000} = \boxed{95\%}$

2. $\text{Precision} = \frac{TP}{TP+FP} = \frac{970}{970+30} = \frac{970}{1000} = \boxed{97\%}$

3. $\text{Recall} = \frac{TP}{TP+FN} = \frac{970}{970+70} = \frac{970}{1040} = \boxed{93.26\%}$

4. $F1: \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times 0.97 \times 0.9326}{0.97 + 0.9326} = \boxed{95.09\%}$