Course : CS513 Knowledge Discovery and Data Mining
First Name: John
Last Name : Rizzo
Student ID: 20036833
Purpose : Homework 1 Probability Assignment

Problem 1.1

Jerry and Susan have a joint bank account.
Jerry goes to the bank 20% of the days.
Susan goes there 30% of the days.
Together they are at the bank 8% of the days.
$P(B_j) = .20$
$P(B_s) = .30$
$P(B_j \cap B_s) = .08$

a

$P(B_j|B_s) = \frac{P(B_j \cap B_s)}{P(B_s)} = \frac{.08}{.30} = 0.2667 = 27\%$

b

$P(B_s') = 1 - P(B_s) = 1 - .30 = .70$
The probability that Jerry goes to the bank and not Susan
$P(B_j \cap B_s') = P(B_j) - P(B_j \cap B_s) = .20 - .08 = .12$
The probability that Jerry was at the bank given Susan didn't go
$P(B_j|B_s') = \frac{P(B_j \cap B_s')}{P(B_s')} = \frac{.12}{.70} = 0.1714 = 17.14\%$

c

At least one
$P(B_j \cup B_s) = P(B_j) + P(B_s) - P(B_j \cap B_s) = .20 + .30 - .08 = .42 = 42\%$
Probability that both go is $P(B_j \cap B_s) = .08 = 8\%$

Problem 1.2

Harold and Sharon are studying for a test.
Harold's chances of getting a "B", $P(B_h)$, are 80%.
Sharon's chances of getting a "B", $P(B_s)$ are 90%.
The probability of at least one of them getting a "B", $P(B_h \cup B_s)$, is 91%.
$P(B_h) = .80$
$P(B_s) = .90$
$P(B_h \cup B_s) = .91$
$P(B_h \cup B_s) = P(B_h) + P(B_s) - P(B_h \cap B_s)$
$.91 = .80 + .90 - P(B_h \cap B_s) = 0.79 = 79\%$
$P(B_h \cap B_s) = .80 + .90 - .91 = .79 = 79\%$

a $P(B_h) - P(B_h \cap B_s) = .80 - .79 = .01 = 1\%$

b $P(B_s) - P(B_h \cap B_s) = .90 - .79 = .11 = 11\%$

c $1 - P(B_h' \cup B_s') = 1 - .91 = .09 = 9\%$

Problem 1.3

The events are dependent events because $P(B_j \cap B_s) \neq P(B_j)P(B_s)$

$P(B_j) = .2 = 20\%$

$P(B_s) = .3 = 30\%$

$P(B_j \cap B_s) = 0.08 = 8\%$

$P(B_j)P(B_s) = 0.2 \times 0.3 = 0.06 = 6\%$

$0.08 \neq 0.06$

Problem 1.4

a

The events are dependent events because $P(A \cap B) \neq P(A)P(B)$

There are 36 possible combinations of 2 six sided die.

P(A) = sum is 6

P(B) = second die is 5

P(A) - Combinations include (5,1) (4,2) (3,3) (2,4) (1,5) so probability is $\frac{5}{36}$

P(B) - Combinations include (1,5) (2,5) (3,5) (4,5) (5,5) (6,5) so probability is $\frac{6}{36} = \frac{1}{6}$

$P(A \cap B) = \frac{1}{36}$ because the only possible option is (1,5)

Solving

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

$P(A \cap B) = \frac{1}{36}$ or if multiplied by 6 $\frac{6}{216}$

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

$\frac{1}{36} \neq \frac{6}{216}$ so they dependent

b

P(A) = The sum is 7

There are 36 possible combinations of two die

P(A) Possible combinations are (1,6) (2,5) (3,4) (4,3) (5,2) (6,1)

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

P(B) = the first die shows 5.

P(B) possible combinations include (5,1) (5,2) (5,3) (5,4) (5,5) (5,6)

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

$P(A \cap B) = \frac{1}{36}$ since only (5,2) satisfies the criteria

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

$\frac{1}{36} = \frac{1}{36}$

Since $P(A \cap B) = P(A)P(B)$ then they are independent.

Problem 1.5

Probability of Finding Choosing a State

Texas: $P(TX) = 0.6$

New Jersey: $P(NJ) = 0.1$

Alaska: $P(AK) = 0.3$ since the total must be 1

Probability of Finding Oil

Texas: $P(Oil|TX) = 0.3$

New Jersey: $P(Oil|NJ) = 0.1$

Alaska: $P(Oil|AK) = 0.2$

1

Using the law of Total Probability

$P(Oil) = P(TX)P(Oil|TX) + P(NJ)P(Oil|NJ) + P(AK)P(Oil|AK)$

$P(Oil) = (0.6)(0.3) + (0.1)(0.1) + (0.3)(0.2)$
$P(Oil) = 0.18 + 0.01 + 0.06 = 0.25 = 25\%$

2

Using Bayes' Theorem
$P(TX|Oil) = \frac{P(Oil|TX)P(TX)}{P(Oil)} = \frac{(0.3)(0.6)}{0.25} = \frac{0.18}{0.25} = 0.72 = 72\%$

Problem 1.6

Total Passengers: 2201

The probability that a passenger did not survive.
$\frac{1490}{2201} = 0.68 = 68\%$

The probability that a passenger was staying in the first class.
$\frac{325}{2201} = 0.15 = 15\%$

Given that a passenger survived, what is the probability that the passenger was staying in the first class.
$\frac{203}{711} = .29 = 29\%$

Are survival and staying in the first class independent?
$P(A) = P(Survival) = \frac{711}{2201} = 0.32 = 32\%$
$P(B) = P(1^{st} class) = \frac{325}{2201} = 0.15 = 15\%$
$P(A \cap B) = .29 = 29\%$
$P(A \cap B) \neq P(A)P(B)$
$29\% \neq 47\%$ therefor they are dependent.

Given that a passenger survived what is the probability that the passenger was staying in first class and the passenger was a child?
$\frac{6}{203} = 0.03 = 3\%$

Given that a passenger survived, what is the probability that the passenger was an adult?
$\frac{654}{711} = 0.92 = 92\%$

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

Problem 1.7

	AI Generated (Positive)	Human Generated Negative	Total
AI Generated (Positive)	970	70	1040
Human Generated (Negative)	30	930	960
	1000	1000	2000

	Actual Positive	Actual Negative
Predicted Positive	TP	FP
Predicted Negative	FN	TN

Accuracy $= \frac{TP+TN}{TP+FP+FN+TN} = \frac{970+930}{970+70+30+930} = \frac{1900}{2000} = 0.95 = 95\%$
Precision $= \frac{TP}{TP+FP} = \frac{970}{970+70} = \frac{970}{1040} = 0.93 = 93\%$
Recall $= \frac{TP}{TP+FN} = \frac{970}{970+30} = \frac{970}{1000} = 0.97 = 97\%$
F1 $= \frac{2 \times Precision \times Recall}{Precision+Recall} = \frac{2(0.93)(0.97)}{(0.93)+(0.97)} = \frac{1.8042}{1.9} = 0.95 = 95\%$