Homework – 1 - Probability

Keval Sompura

1. a. Probability Jerry is at the bank P(J) = 0.2

   Probability Susan is at the bank P(S) = 0.3

   $P(J \cap S) = 0.08$

   Probability that Jerry was also at the bank, given that Susan was at the bank last Monday = $P(J \mid S) = P(J \cap S) / P(J) = 0.08/0.3 = 0.2667$

   Answer: **26.67%**

   b. Probability that susan is not at the bank P(S') = 0.7

   Next, we need to find the probability that Jerry was at the bank **and** Susan was not.

   $P(J \cap S') = P(J) - P(J \cap S) = 0.20 - 0.08 = 0.12$

   Now, using the conditional probability formula:

   $P(J|S') = P(J \cap S')/P(S')$

   $P(J|S') = 0.12/0.7 = 0.1714$

   So, the probability that Jerry was at the bank given that Susan was not there is approximately **17.14%.**

   c. Probability that atleast one of them is at the bank:

   $P(J \cup S) = P(J) + P(S) - P(J \cap S) = 0.20 + 0.30 - 0.08 = 0.42$

   Probability that both of them were there given that atleast one of them is there:

   $P(J \cap S|J \cup S) = P(J \cup S)P(J \cap S) = 0.08 / 0.42 = 0.1905$

   So, the probability that both Jerry and Susan were at the bank given that at least one of them was there is approximately **19.05%**.

2. Given :
   Harold's probability of getting a "B": P(H) = 0.8
   Sharon's probability of getting a "B": P(S) = 0.9

The probability of **at least one** of them getting a "B" : $P(H \cup S)=91\%=0.91$

Probability that both of them gets a B = $P(H \cap S)=P(H)+P(S)-P(H \cup S)$

$\boldsymbol{P(H \cap S)} = 0.80+0.90-0.91=\mathbf{0.79}$

a. Probability that only Harold gets a B = $P(\text{Only Harold})=P(H)-P(H \cap S)$

$= 0.80-0.79 = \mathbf{0.01}$

So, the probability that **only Harold** gets a "B" is **1%**.

b. Probability that only Shanon gets a B
   $P(\text{Only Sharon})=P(S)-P(H \cap S)$
   $P(\text{Only Sharon})=0.90-0.79=0.11$

So, the probability that only Sharon gets a "B" is 11%.

c. The probability that **both won't** get a "B" is
   $P(\text{neither}) = 1-P(H \cup S)$
   $= 1 - 0.91 = 0.09$

So, the probability that both wont get a B is 9%.

3. $P(J)=20\%=0.20$ (probability that Jerry is at the bank).
   $P(S)=30\%=0.30$ (probability that Susan is at the bank).
   $P(J \cap S)=8\%=0.8$ (probability that both Jerry and Susan are at the bank).

   Two events J and S are independent if: $P(J \cap S)=P(J)\cdot P(S)$
   $\boldsymbol{P(J)\cdot P(S)=0.20\cdot 0.30=0.06}$
   Since $P(J \cap S)=0.8$ is **not equal** to $P(J)\cdot P(S)=0.06$, the events are **not independent**.

4. a. Total outcomes when rolling two dice = 36
   Let A be the event that the sum is 6.
   The possible outcomes for a sum of 6 are:
   (1,5),(2,4),(3,3),(4,2),(5,1)
   $P(A) = 5/36$

   Let B be the event that the second die shows 5.

The second die can show 5 in 6 possible outcomes:
(1,5),(2,5),(3,5),(4,5),(5,5),(6,5)
P(B) = 6/36 = 1/6
Probability of A∩B: The sum is 6 and the second die shows 5.
The only outcome that satisfies both conditions is (1,5). Thus:
$P(A \cap B)=1/36$
For A and B to be independent:
$P(A \cap B)=P(A) \cdot P(B)$
Since 1/36 != 5/216, both the events are not independent

b. let C be the event that the sum is 7.
The possible outcomes for a sum of 7 are:
(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)
P(c) = 6/36 = 1/6

Let D be the event the first dice shows 5.
The first die can show 5 in 6 possible outcomes:
(5,1),(5,2),(5,3),(5,4),(5,5),(5,6)
P(D) = 6/36 = 1/6

Probability of C∩D: The sum is 7 and the first die shows 5.
The only outcome that satisfies both conditions is (5,2). Thus:
$P(C \cap D) = 1/36$

For C and D to be independent:
$P(C \cap D)=P(C) \cdot P(D)$

Since 1/36 = 1/36, the events are **independent**.

5. Let T,A,N as the events that the company drills in **Texas, Alaska, and New Jersey**, respectively.
   O as the event of **finding oil**.

   $P(T)=0.60$, $P(A)=0.30$, $P(N)=0.10$
   $P(O|T)=0.30$, $P(O|A)=0.20$, $P(O|N)=0.10$
   a. Probability of finding oil: $P(O)=P(O|T)P(T)+P(O|A)P(A)+P(O|N)P(N)$
      $P(O)= (0.30\times0.60)+(0.20\times0.30)+(0.10\times0.10) = 0.25$
      Therefore probability of finding oil is 25%.

b. We need to find P(T|O), using **Bayes' Theorem**:
   $P(T|O) = P(O|T) \times P(T)/P(O)$
   $P(T|O) = 0.30 \times 0.6 / 0.25 = 0.72$
   Therefore if the company found oil, there is a **72%** probability that they drilled in Texas

6. a. Probability that a passenger did not survive
   P(Not Survived) = Total Not survived / total = 1490 / 2201 = 0.677

   Therefore, the probability that a passenger did not survive is 67.7%

   b. Probability that a passenger was staying in first class

   P(first class) = total first class / total = 325 / 2201 = 0.148

   Therefore the probability is approximately 14.8%

   c. Probability that the person survived was staying in first class
   P(Survive first class) = total survived first class / total = 203/711 = 0.285
   Therefore probability is approximately 28.5%

   d. Two events A and B are independent if: $P(A \cap B) = P(A) \cdot P(B)$
   P(Survived) = 0.323
   P(1^{st} class) = 0.148
   P(Survived and first class) = 203/2201 = 0.092

   P(Survived) x P(1st class) = 0.048

   Since 0.092 != 0.048, the events are not independent.

   e. Given that a passenger survived, what is the probability that the passenger was staying in the first class and the passenger was a child?
   This is a conditional probability. We use the formula:
   P(1st Class and Child | Survived) =
   P(first class and child and survived)/P(survived)

   P(first class and child and survived) = 6/2201

P(survived) = 711/2201

P(1st Class and Child | Survived) = 6/711 = 0.0084

The probability is approximately 0.84%

f. Given that a passenger survived, what is the probability that the passenger was an adult?
   P(Adult | Survived) = P(Adult and survived)/P(Survived)
   P(Adult and Survived) = 654 / 2201
   P(Survived) = 711/2201
   P(Adult | Survived) = 654/711 = 0.92

   The probability is approximately 92%

g. Given that a passenger survived, are age and staying in the first class independent?
   To check independence, we compare:
   P(1st Class and Adult|Survived) and P(1st Class|Survived)·P(Adult|Survived)

   P(1st Class and Adult|Survived) = 197/711 = 0.277
   P(1st Class|Survived) = 203/711 = 0.285
   P(1st Class|Survived) = 654/711 = 0.920

P(1st Class|Survived)·*P*(Adult|Survived) = 0.285 x 0.920 = 0.262

Since 0.277 != 0.262, age and staying in the first class are **not independent** given survival.

7. **Total AI-generated documents:** 1000
   Total human-generated documents: 1000
   Misclassified human-generated documents (FP): 70
   Misclassified AI-generated documents (FN): 30

   Therefore confusion matrix is:

	Predicted Ai	Predicted Human	Total
Actual Ai	970	30	1000
Actual Human	70	930	1000
Total	1040	960	2000

True Positive (TP): AI documents correctly classified as AI = **970**

False Negative (FN): AI documents misclassified as Human = **30**

False Positive (FP): Human documents misclassified as AI = **70**

True Negative (TN): Human documents correctly classified as Human = **930**

Accuracy = (TP + TN) / (TP + TN + FP + FN)
= 970+930 / 2000 = 1900 / 2000 = 0.95
Therefore Accuracy is 95%

Precision = TP / (TP + FP)
= 970 / 970 + 70 = 0.9327

Therefore precision is 93.27%

Recall = TP / TP + FN
= 970 / 970+30 = 0.97
There recall is 97%

F1 = 2x (Precision x Recall)/ (Precision + Recall)
= 2 x (0.9327×0.97) / (0.9327 - 0.97)
= 0.9517

Therefore F1 is 95.17%