Assignment 5

Vishal

2022-10-26

1. Conditional probability, Bayes rule and independence

1.1 (Q1)

We have,

$\mathbb{P}(A) = 0.9$
$\mathbb{P}(B|A) = 0.8$
$\mathbb{P}(B^{\complement}|A^{\complement}) = 0.75$

And we know that,
$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(A^{\complement} \cap B) \Rightarrow \mathbb{P}(B|A).\mathbb{P}(A) + \mathbb{P}(B|A^{\complement}).\mathbb{P}(A^{\complement})$
$\mathbb{P}(B) = \mathbb{P}(B|A).\mathbb{P}(A) + \mathbb{P}(B|A^{\complement}).\mathbb{P}(A^{\complement})$
$\Rightarrow 0.8 * 0.9 + [1 - \mathbb{P}(B^{\complement}|A^{\complement})] * [1 - \mathbb{P}(A)]$
$\Rightarrow 0.8 * 0.9 + [0.25] * 0.1$
$\Rightarrow 0.745$

Using Bayes theorem,
$\mathbb{P}(A|B) = \frac{\mathbb{P}(A).\mathbb{P}(B|A)}{\mathbb{P}(B)}$
$\Rightarrow \frac{0.9*0.8}{0.745}$
$\Rightarrow 0.9666443$

Therefore, the probability of rain is **0.966443** given the forecast of rain.

1.2 (Q1)

1. $\mathbb{P}(A|B) = \frac{\mathbb{P}(A).\mathbb{P}(B|A)}{\mathbb{P}(B)}$
   If $\mathbb{P}(B \backslash A) = \phi$, the probability becomes indeterminate.

2. $\mathbb{P}(A|B) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$
   Hence, if $\mathbb{P}(B \cap A) = 0$, $\mathbb{P}(A|B) = 0$.

3. Since, $B \in A$, probability becomes
   $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$
   $\Rightarrow \frac{\mathbb{P}(A)}{\mathbb{P}(A)} \Rightarrow 1$
   If $\mathbb{P}(B \backslash A) = 0$, then probability is 0.

4. $\mathbb{P}(A \cap \Omega) = \mathbb{P}(A) = \mathbb{P}(A).1 = \mathbb{P}(A)\mathbb{P}(\Omega)$.

5. $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \cap (B \cap C))$
$\Rightarrow \mathbb{P}(A|(B \cap C))\mathbb{P}(B \cap C)$
$\Rightarrow \mathbb{P}(A|(B \cap C))(\mathbb{P}(C|A)\mathbb{P}(B))$
$\Rightarrow \mathbb{P}(C)\mathbb{P}(B|C)\mathbb{P}(A|B \cap C)$

6. We know that,
$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$
Here, if we condition everything on C, we get,
$\mathbb{P}(A|B \cap C) = \frac{\mathbb{P}(A \cap B|C)}{\mathbb{P}(B|C)}$
$\Rightarrow \mathbb{P}(A|(B \cap C)) = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(B \cap C)}$
$\Rightarrow \frac{\mathbb{P}(B|(A \cap C)).\mathbb{P}(A|C)}{\mathbb{P}(B|C)}$

1.2 (Q2)

Let's say,
$\mathbb{P}(W) = 0.2$ i.e. probability of wind
$\mathbb{P}(F|W) = 0.3$ i.e. probability of flight when there's a wind
$\mathbb{P}(F^{\mathsf{C}}|W^{\mathsf{C}}) = 0.1$ i.e. probability of flight being cancelled when there's no wind
$\mathbb{P}(F^{\mathsf{C}}) =?$ Probability of flight not being cancelled

We know that,
$\mathbb{P}(F^{\mathsf{C}}|W^{\mathsf{C}}) = \frac{\mathbb{P}(F^{\mathsf{C}}) - \mathbb{P}(F^{\mathsf{C}}|W) \cdot \mathbb{P}(W)}{\mathbb{P}(W^{\mathsf{C}})}$
$0.1 = \frac{\mathbb{P}(F^{\mathsf{C}}) - [1 - \mathbb{P}(F|W)] \cdot 0.2}{0.8}$
$0.08 = \mathbb{P}(F^{\mathsf{C}}) - [0.14]$
$\mathbb{P}(F^{\mathsf{C}}) = 0.22$

Therefore, the probability of flight being not cancelled is **0.22**

1.3 Mututal independence and pair-wise independent (Q1)

$(A \cap B) = \{(1, 1, 0)\}$
And probability of $(A \cap B) = 1/4$
$\mathbb{P}(A) = 2/4$ & $\mathbb{P}(B) = 2/4$
Therefore,
$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \Rightarrow 1/4$ $(A \cap C) = \{(1, 0, 1)\}$
$\mathbb{P}(A) = 2/4$ & $\mathbb{P}(C) = 2/4$
And probability of $(A \cap C) = 1/4$
Therefore,
$\mathbb{P}(A \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(C)$ $(C \cap B) = \{(0, 1, 1)\}$
And probability of $(C \cap B) = 1/4$
$\mathbb{P}(C) = 2/4$ & $\mathbb{P}(B) = 2/4$
Therefore,
$\mathbb{P}(C \cap B) = \mathbb{P}(C) \cdot \mathbb{P}(B)$
$A \cap B \cap C$ denotes the intersection of 3 sets, A, B and C. i.e. common elements that occur in all three sets.
$P(A \cap B \cap C) = P\big((A \cap B) \cap C\big)$

$\Rightarrow P\big(C \mid (A \cap B)\big)P(A \cap B)$

$\Rightarrow P\big(C \mid (A \cap B)\big)\Big(P(B \mid A)P(A)\Big)$

$\Rightarrow P(A)P(B \mid A)P(C \mid A \cap B)$
Yes, events A, B and C are mutually independent since their outcomes do not depend on the outcome on the other outcomes.

1.4 (Q1)

Yes, the contestant is improving their chances of winning the car if they switch their initial choice.
If the contestant decides to stick to the first choice, probability of winning a car is 1/3 whereas, if they decide to switch, probability becomes 2/3.

$\mathbb{P}(A_3|B_1 \cap C_2) = \frac{2}{3}$
Yes, contestant should switch their choice to increase their chances of winning a car!

2. Random variables and discrete random variables

2.1 (Q1)

We know that,
$E(XY) = E(X)E(Y)$ when X is independent of Y and $Cov(X,Y) = E(XY) - E(X)E(Y)$
Therefore, $Cov(X,Y) = 0$ is obtained when X is independent of Y.

2.2 (Q1)

1. Probability Mass Function
$p_X(X) = \mathbb{P}(X = x) for x = 0, 3, 10$
We have,
$p_X(0) = \mathbb{P}(X = 0) = 0$
$p_X(3) = \mathbb{P}(X = 3) = \alpha$
$p_X(10) = \mathbb{P}(X = 10) = \beta$

2. Expectation
$E(X) = 3\alpha + 10\beta$

3. Variance
$Var(X) = \mathbb{E}[(X - \mathbb{E}(X))^2]$
$\mathbb{E}(X) = 3\alpha + 10\beta$
and $\mathbb{E}(X^2) = 9\alpha^2 + 60\alpha\beta + 100\beta^2$
Therefore
$Var(X) = 3\alpha + 10\beta - 9\alpha^2 + 60\alpha\beta + 100\beta^2$

4. Standard Deviation
$\sigma = \sqrt{\mathbb{E}(X^2) - [\mathbb{E}(X)^2]}$
$\mathbb{E}(X) = 3\alpha + 10\beta$

and $\mathbb{E}(X^2) = 9\alpha^2 + 60\alpha\beta + 100\beta^2$
$\mathbb{E}(X)^2 = 9\alpha + 100\beta$
$\sigma = \sqrt{3\alpha + 10\beta - (9\alpha^2 + 60\alpha\beta + 100\beta^2)}$

2.2 (Q2)

1. Distribution $p_X(X)$
$p_X(0) = \mathbb{P}(X = 0) = 0$
$p_X(3) = \mathbb{P}(X = 3) = \alpha$
$p_X(10) = \mathbb{P}(X = 10) = \beta$

2. Distribution function $F_X(X)$
$F_X = \left\{ 3\alpha, forx = 3 \right.$
$10\beta, forx = 10$
$\left. 0, otherwise \right\}$

2.2 (Q3)

Consider the rolling of a fair six-sided die, with X the number on the uppermost face. We know that the p_f of X is

$p_X(x) = \frac{1}{6}, x = 1, 2, 3, 4, 5, 6$
$Var(X) = \mathbb{E}[(X - \mathbb{E}(X))^2]$

2.2 (Q4)

A random variable is called discrete when it can take countable number of values. And it is said random when the sum of probabilities is 1.
So, in the given example, the probabilities given are 0.2 and 0.3 and it is also provided that $\alpha + \beta <= 1$.
Hence, Y is a discrete.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
samples_Xi <- rmultinom(50000, 3, prob = c(0.5, 0.2, 0.3))

temp <- as.data.frame((samples_Xi))

## Y = X1 + X2 + .... + Xn
```

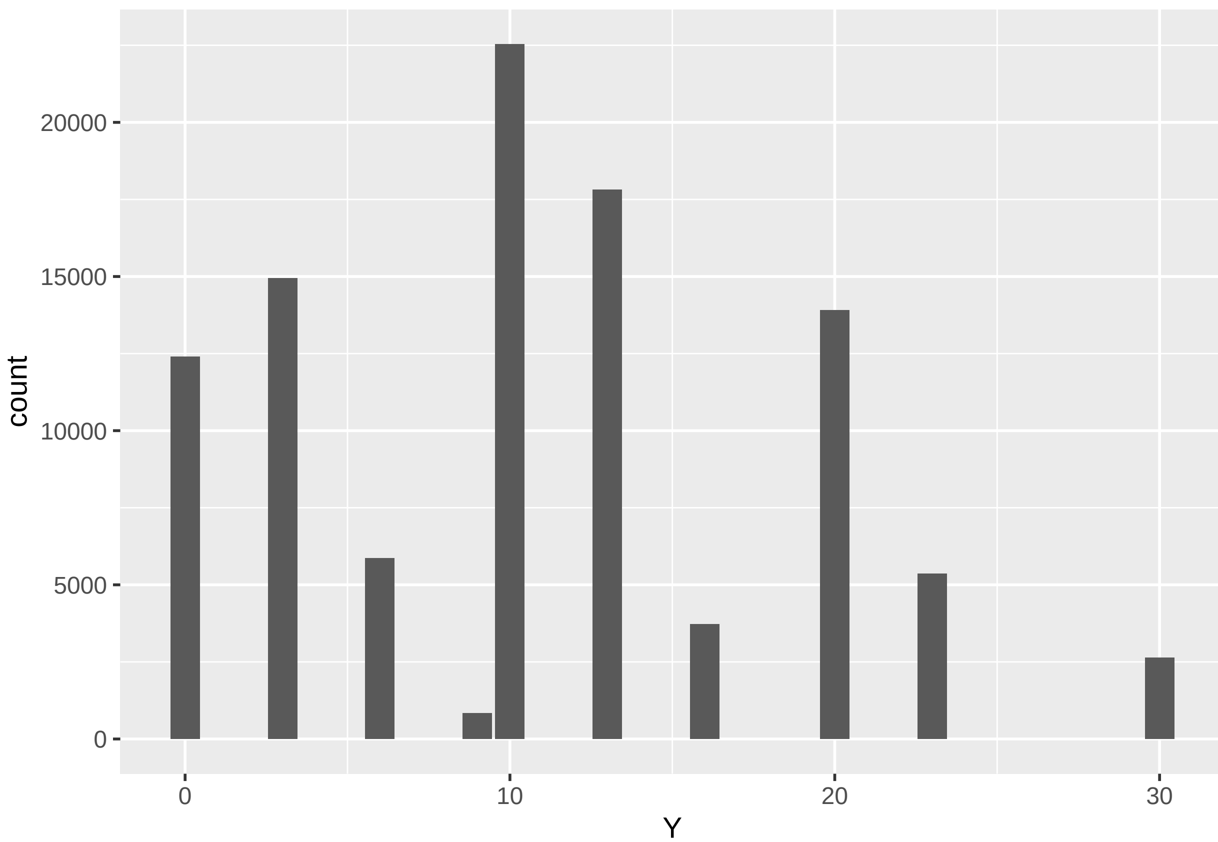
```
firstRow <- temp[1, ] * 0
secondRow <- temp[2, ] * 3
thirdRow <- temp[3, ] * 10

# Creating data frame from the vectors
temp2 <- rbind(firstRow, secondRow, thirdRow)

# Calculate sum of all columns to create a df of Ys
samples_Y <- temp2 %>%
  mutate(Total = select(., V1:V50000)) %>%
  colSums(na.rm = T)

samples_Y <- data.frame(samples_Y)
```

```
colnames(samples_Y) <- "Y"
samples_Y %>%
  ggplot(aes(x = Y)) +
  geom_bar()
```




```
# n = 20
samples_Xi <- rmultinom(50000, 20, prob = c(0.5, 0.2, 0.3))

temp <- as.data.frame(samples_Xi)
firstRow <- temp[1, ] * 0
secondRow <- temp[2, ] * 3
```

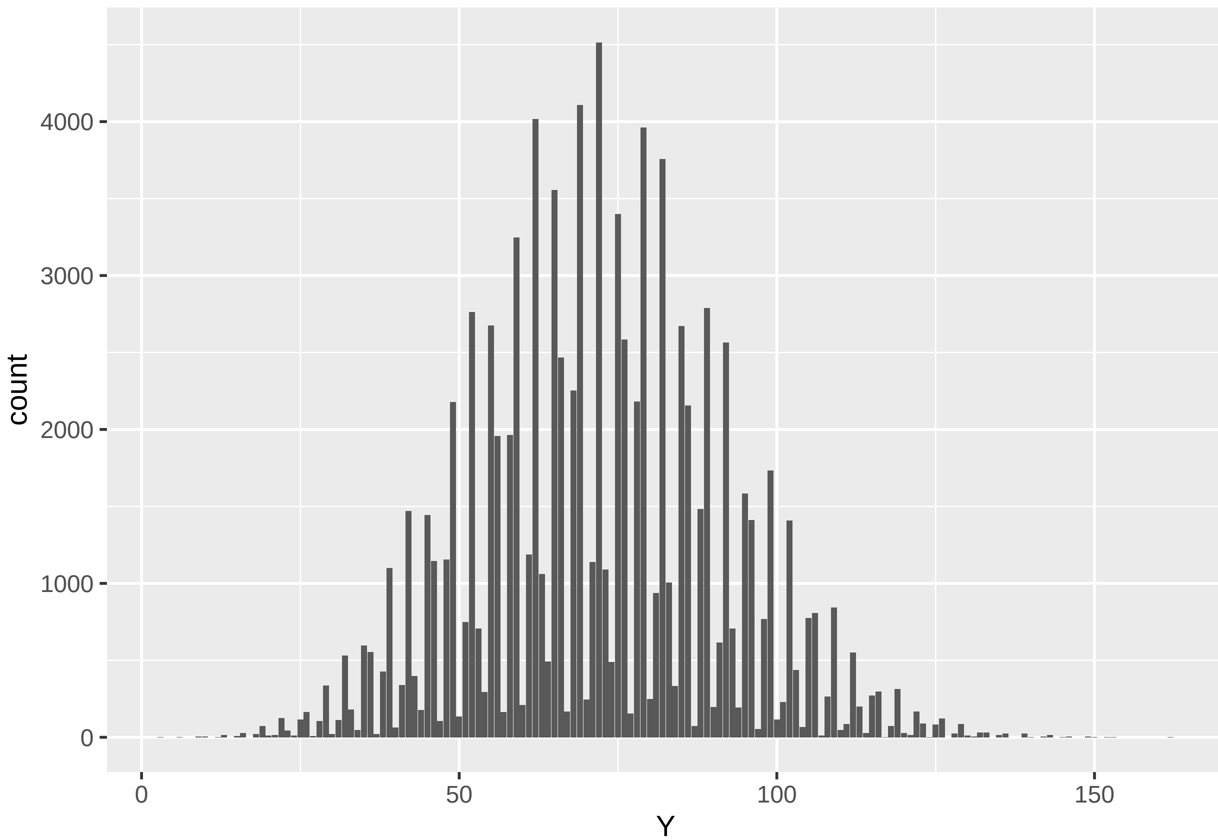
```r
thirdRow <- temp[3, ] * 10

# Creating data frame from the vectors
temp2 <- rbind(firstRow, secondRow, thirdRow)

# Calculate sum of all columns to create a df of Ys
samples_Y <- temp2 %>%
  mutate(Total = select(., V1:V50000)) %>%
  colSums(na.rm = T)

samples_Y <- data.frame(samples_Y)
colnames(samples_Y) <- "Y"
```

```r
samples_Y %>%
  ggplot(aes(x = Y)) +
  geom_bar()
```



Here, we can see that minimum value of sample is 6 and the maximum value is 156.
Sample range is [6, 156].

```r
# n = 2000
samples_Xi <- rmultinom(50000, 2000, prob = c(0.5, 0.2, 0.3))

temp <- as.data.frame(samples_Xi)
firstRow <- temp[1, ] * 0
```

```
secondRow <- temp[2, ] * 3
thirdRow <- temp[3, ] * 10

# Creating data frame from the vectors
temp2 <- rbind(firstRow, secondRow, thirdRow)

# Calculate sum of all columns to create a df of Ys
samples_Y <- temp2 %>%
  mutate(Total = select(., V1:V50000)) %>%
  colSums(na.rm = T)

samples_Y <- data.frame(samples_Y)
colnames(samples_Y) <- "Y"
```

```
samples_Y %>%
  ggplot(aes(x = Y)) +
  geom_bar()
```