FA-5 R 7.1

Espiritu, Joseph Raphael M.

2025-03-05

Exercises 7.1

Problem Statement 1:

An email message can travel through one of three server routes. The percentage of errors in each of the servers and the percentage of messages that travel through each route are shown in the following table. Assume that the servers are independent.

	Percentage of Messages	Percentage of Errors
Server 1	40	1
Server 2	25	2
Server 3	35	1.5

(a) What is the probability of receiving an email containing an error?

(b) What is the probability that a message will arrive without error?

(c) If a message arrives without error, what is the probability that it was sent through server 1?

```
p_server <- c(0.40, 0.25, 0.35)  # Probability of each server
p_error  <- c(0.01, 0.02, 0.015) # Error probability for each server

# Compute total probability of error
p_email_error <- sum(p_server * p_error)

p_email_error
```

```
## [1] 0.01425
```

```
p_email_no_error <- 1 - p_email_error
p_email_no_error
```

```
## [1] 0.98575
```

```
p_no_error_given_s1 <- 1 - p_error[1]  # P(E^c | S1)
```

```r
# Compute P(S1 | E^c)
p_s1_given_no_error <- (p_no_error_given_s1 * p_server[1]) / p_email_no_error

p_s1_given_no_error
```

```
## [1] 0.4017246
```

Problem Statement 2:

9. A software company surveyed managers to determine the probability that they would buy a new graphics package that includes three-dimensional graphics. About 20% of office managers were certain that they would not buy the package, 70% claimed that they would buy, and the others were undecided. Of those who said that they would not buy the package, only 10% said that they were interested in upgrading their computer hardware. Of those interested in buying the graphics package, 40% were also interested in upgrading their computer hardware. Of the undecided, 20% were interested in upgrading their computer hardware.

   Let A denote the intention of not buying, B the intention of buying, C the undecided, and G the intention of upgrading the computer hardware.

   (a) Calculate the probability that a manager chosen at random will not upgrade the computer hardware ($P(\overline{G})$).

   (b) Explain what is meant by the posterior probability of B given G, $P(B|G)$.

   (c) Construct a tree diagram and use it to calculate the following probabilities: $P(G)$, $P(B|G)$, $P(B|\overline{G})$, $P(C|G)$, $P(\overline{C}|\overline{G})$.

```r
library(DiagrammeR)
```

```
## Warning: package 'DiagrammeR' was built under R version 4.4.3
```

```r
library(DiagrammeRsvg)
```

```
## Warning: package 'DiagrammeRsvg' was built under R version 4.4.3
```

```r
library(rsvg)
```

```
## Warning: package 'rsvg' was built under R version 4.4.3
## Linking to librsvg 2.57.0
```

```r
library(knitr)

p_A <- 0.20  # Not buying
p_B <- 0.70  # Buying
p_C <- 0.10  # Undecided

p_G_given_A <- 0.10  # P(G | A)
p_G_given_B <- 0.40  # P(G | B)
```

```r
p_G_given_C <- 0.20  # P(G | C)

# Compute P(G)
p_G <- (p_G_given_A * p_A) + (p_G_given_B * p_B) + (p_G_given_C * p_C)

p_G
```

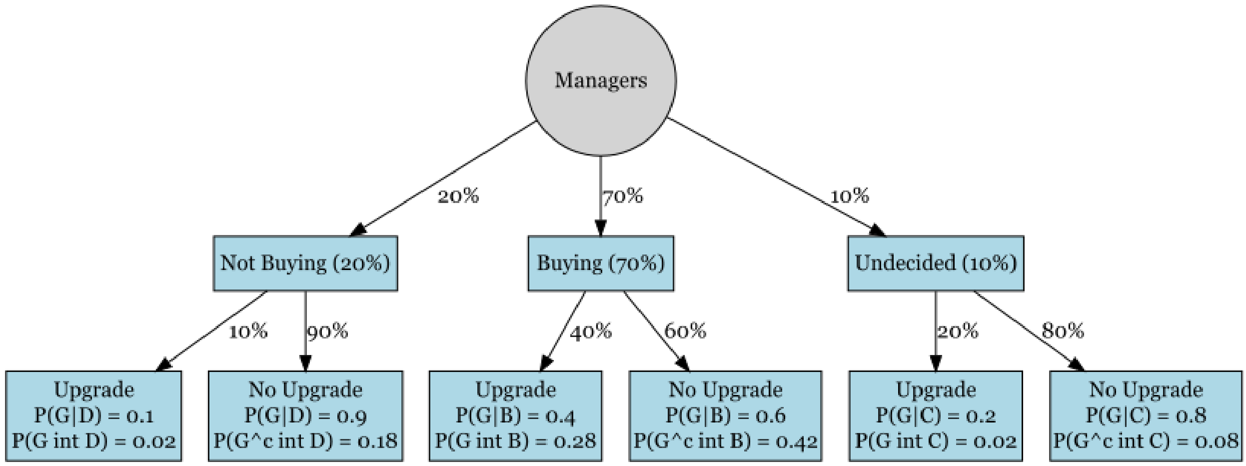
```
## [1] 0.32
```

```r
p_B_given_G <- (p_G_given_B * p_B) / p_G
p_B_given_G
```

```
## [1] 0.875
```

```r
diagram <- grViz("
digraph Tree {
  node [shape=box, style=filled, fillcolor=lightblue]

  Start [label='Managers', shape=circle, fillcolor=lightgray]
  NotBuying [label='Not Buying (20%)']
  Buying [label='Buying (70%)']
  Undecided [label='Undecided (10%)']
  Upgrade_NotBuying [label='Upgrade \\nP(G|D) = 0.1\\nP(G int D) = 0.02']
  NoUpgrade_NotBuying [label='No Upgrade \\nP(G|D) = 0.9\\nP(G^c int D) = 0.18']
  Upgrade_Buying [label='Upgrade \\nP(G|B) = 0.4\\nP(G int B) = 0.28']
  NoUpgrade_Buying [label='No Upgrade \\nP(G|B) = 0.6\\nP(G^c int B) = 0.42']
  Upgrade_Undecided [label='Upgrade \\nP(G|C) = 0.2\\nP(G int C) = 0.02']
  NoUpgrade_Undecided [label='No Upgrade \\nP(G|C) = 0.8\\nP(G^c int C) = 0.08']

  Start -> NotBuying [label=' 20%']
  Start -> Buying [label=' 70%']
  Start -> Undecided [label=' 10%']

  NotBuying -> Upgrade_NotBuying [label=' 10%']
  NotBuying -> NoUpgrade_NotBuying [label=' 90%']

  Buying -> Upgrade_Buying [label=' 40%']
  Buying -> NoUpgrade_Buying [label=' 60%']

  Undecided -> Upgrade_Undecided [label=' 20%']
  Undecided -> NoUpgrade_Undecided [label=' 80%']
}
")

# Check if output format is HTML
if (knitr::is_html_output()) {
  diagram  # Render in HTML
} else {
  # Convert to image for non-HTML formats
  svg_file <- "tree_diagram.svg"
  png_file <- "tree_diagram.png"

  cat(export_svg(diagram), file = svg_file)
  rsvg_png(svg_file, png_file)
```

```r
  knitr::include_graphics("tree_diagram.png")
}
```



Problem Statement 3:

A **malicious spyware** can infect a computer system through either the **Internet** or through **email**. The probabilities for the **entry points** and the **detection rates** are as follows:

- **70%** of spyware infections occur through the **Internet**.
- **30%** of spyware infections occur through **Email**.

If the spyware enters via the **Internet**, the anti-virus detector will detect it with a probability of **60%**. If it enters via **Email**, the anti-virus will detect it with a probability of **80%**.

These probabilities can be expressed as:

- **70%** of all spyware comes through the **Internet**: $P(\text{Internet}) = 0.70$
- **30%** of all spyware comes through **Email**: $P(\text{Email}) = 0.30$
- The probability of detection if the spyware is from the **Internet**: $P(\text{Detected}|\text{Internet}) = 0.60$
- The probability of detection if the spyware is from **Email**: $P(\text{Detected}|\text{Email}) = 0.80$
- The probability of undetected spyware from the **Internet**: $P(\text{Undetected}|\text{Internet}) = 0.40$
- The probability of undetected spyware from **Email**: $P(\text{Undetected}|\text{Email}) = 0.20$

Questions:

(a) What is the probability that this **spyware infects the system**?

(b) If the spyware is **detected**, what is the **probability that it came through the Internet**?

Solution

- The probability of detected spyware is:

$$P(\text{Detected}) = P(\text{Detected}|\text{Internet})P(\text{Internet}) + P(\text{Detected}|\text{Email})P(\text{Email})$$

```r
# Given Probabilities
P_Internet <- 0.70
P_Email <- 0.30
P_Detected_Internet <- 0.60
P_Detected_Email <- 0.80
P_Undetected_Internet <- 0.40
```

```r
P_Undetected_Email <- 0.20

# Solving for Detected Spyware
P_Detected <- P_Detected_Internet * P_Internet + P_Detected_Email * P_Email

# Show Value
P_Detected
```

```
## [1] 0.66
```

Using **law of total probability**, the probability of receiving **Undetected spyware** is:

$$P(\text{Undetected}) = P(\text{Undetected}|\text{Internet})P(\text{Internet}) + P(\text{Undetected}|\text{Email})P(\text{Email})$$

Using **Bayes' theorem**, the probability that a **spyware detected from the internet** is:

$$P(\text{Internet}|\text{Detected}) = \frac{P(\text{Detected}|\text{Internet})P(\text{Internet})}{P(\text{Detected})}$$

```r
# Solving for Spyware Undetected
P_Undetected <- P_Undetected_Internet * P_Internet + P_Undetected_Email * P_Email

# Solving for Spyware detected from Internet
P_Internet_Detected <- P_Detected_Internet * P_Internet / P_Detected

# Show Values
P_Undetected
```

```
## [1] 0.34
```

```r
P_Internet_Detected
```

```
## [1] 0.6363636
```

Thus, Undetected Spyware appears 34% and Detected spyware from the Internet appears around 63.64% in any computer system.