Assignment 05

Zhen Liu

2022-10-28

```
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──
## ✔ ggplot2 3.3.6      ✔ purrr   0.3.4
## ✔ tibble  3.1.8      ✔ dplyr   1.0.10
## ✔ tidyr   1.2.1      ✔ stringr 1.4.1
## ✔ readr   2.1.3      ✔ forcats 0.5.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

Conditional probability, Bayes rule and independence

Bayes theorem

Q1

$$P(A) = 0.9$$
$$P(B|A) = 0.8$$
$$P(B^c|A^c) = 0.75$$

$$P(B) = P(B|A) * P(A) + P(B|A^c) * P(A^c) = P(B|A) * P(A) + [1 - P(B^c|A^c)] * [1 - P(A)] = 0.8 * 0.9 + 0.25 * 0.1 = 0.745$$

$$P(A|B) = \frac{P(A) * P(B|A)}{P(B)} = 0.9 * 0.8/0.745 = 0.966$$

Conditional probabilities

Q1

1. $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$
$$A \subseteq B$$
$$P(A \cap B) = P(A)$$
$$P(A|B) = \frac{P(A)}{P(B)}$$

$$P \subseteq \emptyset$$

2. $$P(A|B) = 0$$
$$P(A) = 0$$

3. $$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$
$$P(B) = 0$$

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

5.

$$P(A \cap B \cap C) = P(A \cap (B \cap C)) = P(A|(B \cap C))P(B \cap C)$$
$$= P(A|(B \cap C))(P(B|C)P(C))$$
$$= P(A|(B \cap C) * P(B|C) * P(C)$$

6.

$$P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B|C) * P(C)}$$
$$= \frac{P(B|A \cap C) * P(A|C) * P(C)}{P(B|C) * P(C)}$$
$$= \frac{P(B|A \cap C) * P(A|C)}{P(B|C)}$$

Q2

$$P(A|B) = 0.3$$
$$P(A|B^c) = 0.1$$
$$P(B) = 0.2$$
$$P(A^c) = 1 - P(A)$$
$$P(A) = P(A|B) * P(B) + P(A|B^c) * P(B^c)$$
$$= 0.3 * 0.2 + 0.1 * 0.8 = 0.14$$
$$P(A^c) = 0.86$$

Mutual independence and pair-wise independent

Q1

$$A := \{(1,0,1),(1,1,0)\}, B := \{(0,1,1),(1,1,0)\}$$
$$C := \{(0,1,1),(1,0,1)\}$$
$$P(A) = P(B) = P(C) = 1/2$$
$$P(A \cap B) = P(A) * P(B) = 1/4$$
$$P(A \cap C) = P(A) * P(C) = 1/4$$
$$P(C \cap B) = P(C) * P(B) = 1/4$$
$$P(A \cap B \cap C) = 0$$

it is pariwise-independent

The Monty hall problem

Q1 (optional)

Random variables and discrete random variables

Expectation and variance

Q1

$$Cov(X,Y) := E[(X - \overline{X}) * (Y - \overline{Y})]$$
$$Cov(X,Y) := E(XY) - E(X)E(Y)$$
$$= E(X)E(Y) - E(X)E(Y) = 0$$

Distributions

Q1

1.

$$P_X(x) = \begin{cases} \alpha & X = 3 \\ \beta & X = 10 \\ 1 - (\alpha + \beta) & X = 0 \\ 0 & \text{otherwise} \end{cases}$$

2. $$E(X) := \sum_{x \in \{0,3,10\}} x * p_X(x) = 3\alpha + 10\beta$$

3. $$Var(X) := E[(X - E(X))^2] = \alpha * (3 - (3\alpha + 10\beta))^2 + \beta * (10 - (3\alpha + 10\beta))^2 + (1 - (\alpha + \beta)) * (3\alpha + 10\beta)^2$$

4. $$SD = \sqrt{Var(X)}$$

Q2

1. $$P_X(X) = P(X \in S) = P(X \in S \cap \{0,3,10\}) = \frac{1}{3} \sum_{x \in \{0,3,10\}} 1_S(X)$$

2. $$F_X(X) = \begin{cases} \alpha & X = 3 \\ \beta & X = 10 \\ 1 - (\alpha + \beta) & X = 0 \end{cases}$$

Q3

$$Var(Y) := E[(Y - E(Y))^2] = \frac{1}{n^2} \sum (\alpha * (3 - (3\alpha + 10\beta))^2 + \beta * (10 - (3\alpha + 10\beta))^2 + (1 - (\alpha + \beta)) * (3\alpha + 10\beta)^2)$$

Q4

```
rmultinom(2, 7, prob=c(0.5, 0.2, 0.3))
```

```
##      [,1] [,2]
## [1,]    3    3
## [2,]    0    2
## [3,]    4    2
```

```
a<-rmultinom(5, 3, prob=c(0.5, 0.2, 0.3))
b<-a[1,]*0+a[2,]*3+a[3,]*10
```

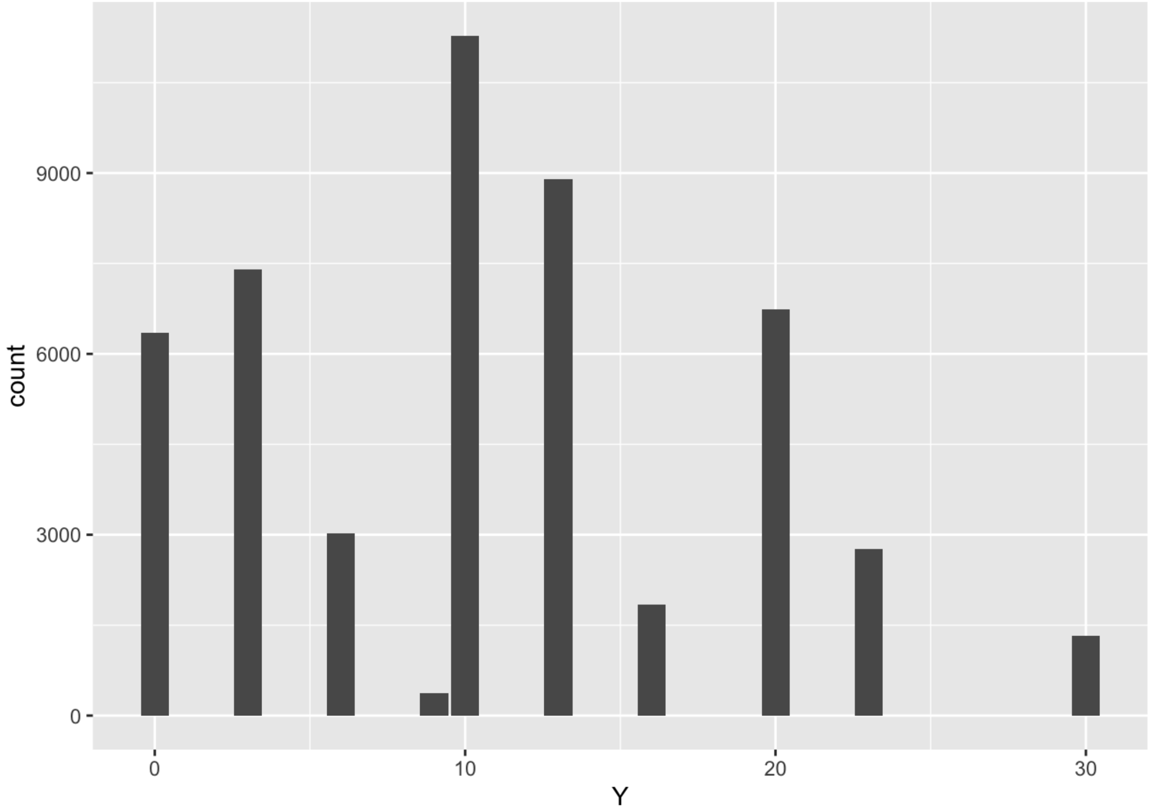
```
c<-data.frame(b)
c
```

```
##    b
## 1  3
## 2 13
## 3 13
## 4 10
## 5 20
```

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

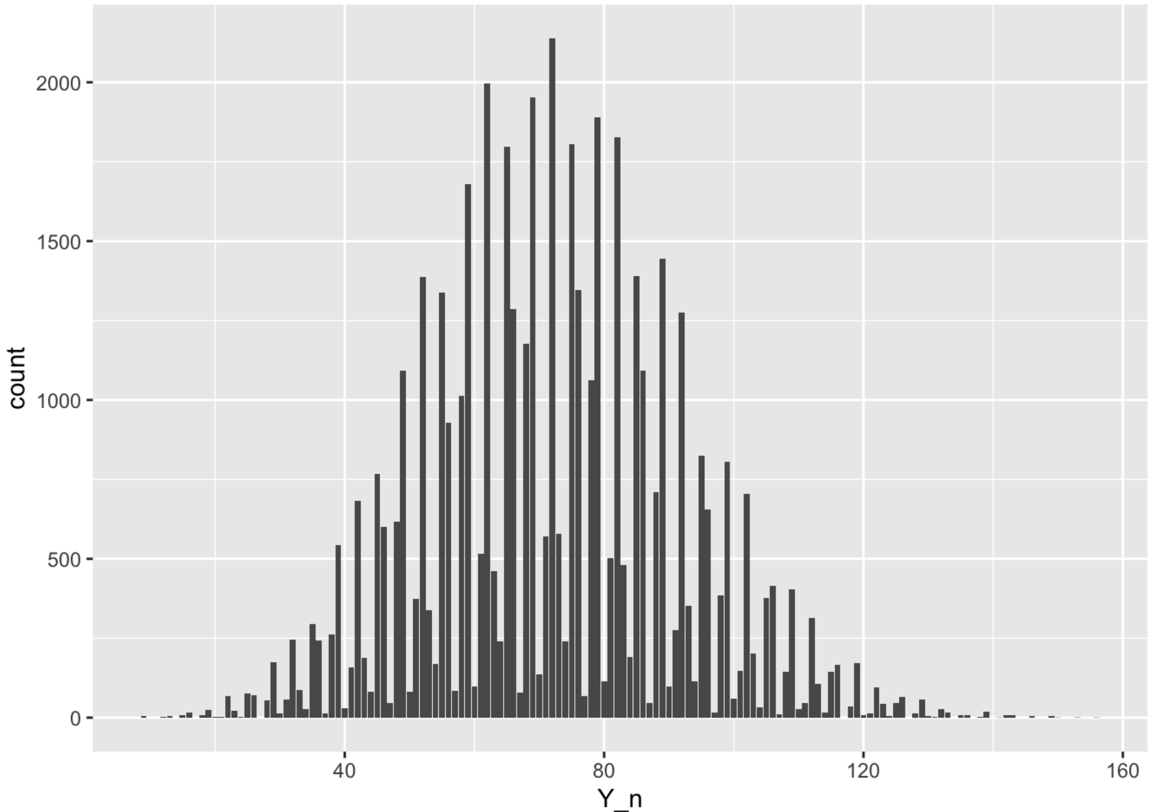
```
Y<-samples_Xi[1,]*0+samples_Xi[2,]*3+samples_Xi[3,]*10
samples_Y = data.frame(Y)
```

```
ggplot(samples_Y,aes(Y))+geom_bar()
```




```
samples_Xi_n<-rmultinom(50000, 20, prob=c(0.5, 0.2, 0.3))
```

```
Y_n<-samples_Xi_n[1,]*0+samples_Xi_n[2,]*3+samples_Xi_n[3,]*10
samples_Y_n = data.frame(Y_n)
ggplot(samples_Y_n,aes(Y_n))+geom_bar()
```



maximum value is 2300, minumum value is 0, range is 0~150

```
samples_Xi_t<-rmultinom(50000, 2000, prob=c(0.5, 0.2, 0.3))
Y_t<-samples_Xi_t[1,]*0+samples_Xi_t[2,]*3+samples_Xi_t[3,]*10
samples_Y_t = data.frame(Y_t)
ggplot(samples_Y_t,aes(Y_t))+geom_bar()
```