Assignment 5

Yujie Wang

2022-10-26

```
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()
```

1. Conditional probability, Bayes rule and independence

1.1 Bayes theorem

1.1 (Q1)

Questions:

Let A be the event that it rains next week and B the event that the weather forecaster predicts that there will be rain next week.

Let's suppose that the probability of rain next week is $P(A) = 0.9$.

Suppose also that the conditional probability that there is a forecast of rain, given that it really does rain, is $P(B|A) = 0.8$.

On the other hand, the conditional probability that there is a forecast of dry weather, given that there really isn't any rain is $P(B^c|A^c) = 0.75$.

Now suppose that there is a forecast of rain. What is the conditional probability of rain, given the forecast of rain $P(A|B)$?

Answers:

Based on the preconditions, we know that:

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

Thus:

$P(B|A^c) = 1 - P(B^c|A^c) = 0.25$
$P(A^c) = 1 - 0.9 = 0.1$

$P(B) = P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)$
$= 0.8 \times 0.9 + 0.25 \times 0.1$
$= 0.745$
$P(A|B) = \dfrac{P(A) \cdot P(B|A)}{P(B)} = \dfrac{0.9 \times 0.8}{0.745}$
≈ 0.9664

1.2 Conditional probabilities

1.2 (Q1)

Questions:

Suppose we have a probability space Ω, E, P.

1. Suppose that $A, B \in E$ and $A \cap B$ and $P(B) \neq 0$. Give an expression for $P(A|B)$ in terms of $P(A)$ and $P(B)$. What about when $P(B \setminus A) = \emptyset$?
2. Suppose that $A, B \in E$ with $A \cap B = \emptyset$. Give an expression for $P(A|B)$. What about when $P(A \cap B) = \emptyset$?
3. Suppose that $A, B, \in E$ with $B \subseteq A$. Give an expression for $P(A|B)$. What about when $P(B \setminus A) = 0$?
4. Suppose that $A \in E$. Give an expression for $P(A|\Omega)$ in terms of $P(A)$?
5. Show that given three events $A, B, C \in E$ we have $P(A \cap B \cap C) = P(A|B \cap C) \cdot P(B|C) \cdot P(C)$.
6. Show that given three events $A, B, C \in E$ and $P(B \cap C) \neq 0$, we have $P(A|B \cap C) = P(B|A \cap C) \cdot P(A|C) \cdot P(B|C)$.

My answers:

1.

(1) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{P(A)}{P(B)}$

(2) $P(B \setminus A) = P(B) - P(A \cap B) = \emptyset$

2.

(1) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{P([p\emptyset)}{P(B)} = 0$

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

3.

(1) $P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{P(B)}{P(B)} = 1$

(2) $P(B \setminus A) = P(B) - P(A \cap B) = 0$
$P(B) = P(A \cap B)$

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

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

5.

$$Based\,on\,the\,formula\,of\,conditional\,probability$$
$$P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)}$$
$$P(B|C) = \frac{P(B \cap C)}{P(C)}$$
$$Thus, P(B \cap C) = P(B|C) \cdot P(C)$$
$$P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(C) \cdot P(B|C)}$$
$$Thus, P(A \cap B \cap C) = P(A|B \cap C) \cdot P(B|C) \cdot P(C)$$

6.

$$Based\,on\,the\,formula\,of\,conditional\,probability:$$
$$P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)}......(1)$$
$$P(B|A \cap C) = \frac{P(A \cap B \cap C)}{P(A \cap C)}......(2)$$
$$P(B|C) = \frac{P(B \cap C)}{P(C)}......(3)$$
$$P(A|C) = \frac{P(A \cap C)}{P(C)}......(4)$$
$$According\,to\,(2), P(A \cap B \cap C) = P(B|A \cap C) \cdot P(A \cap C)......(5)$$
$$According\,to\,(3), P(B \cap C) = P(B|C) \cdot P(C)......(6)$$
$$According\,to\,(4), P(A \cap C) = P(A|C) \cdot P(C)......(7)$$
$$So, based\,on\,the\,rules\,(1),(5),(6),(7):$$
$$P(A|B \cap C) = \frac{P(B|A \cap C) \cdot P(A|C)}{P(B|C)}$$

1.2 (Q2)

Question:

1. If it is windy, then the probability of the flight being cancelled is 0.3.
2. If it is not windy, then the probability of the flight being cancelled is 0.1.

My answer:

Let A is the event that it is windy

Let B is the event that the flight is not cancelled

Then

$P(A) = 0.2, P(A^c) = 0.8$
$P(B^c|A) = 0.3, P(B^c|A^c) = 0.1$

$So\, P(B|A) = 0.7, P(B|A^c) = 0.9$

$So\, P(B) = P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)$
$= 0.7 \times 0.2 + 0.9 \times 0.8$
$= 0.14 + 0.72$
$= 0.86$

1.3 Mutual independence and pair-wise independent

1.3 (Q1)

Questions:

Consider a simple probability space (Ω, E, P) with $\Omega = \{(0,0,0),(0,1,1),(1,0,1),(1,1,0)\}$. Since (Ω, E, P) is a simple probability space containing four elements we have

$P(\{(0,0,0)\}) = P(\{(0,1,1)\}) = P(\{(1,0,1)\}) = P(\{(1,1,0)\}) = 1/4.$

Consider the events $A := \{(1,0,1),(1,1,0)\}$, $B := \{(0,1,1),(1,1,0)\}$ and $C := \{(0,1,1),(1,0,1)\}$.

Verify that $P(A \cap B) = P(A) \cdot P(B)$, $P(A \cap C) = P(A) \cdot P(C)$ and $P(C \cap B) = P(C) \cdot P(B)$. Hence, we deduce that the events A, B, C are pair-wise independent.

What is $A \cap B \cap C$? What is $P(A \cap B \cap C)$? Are the events A, B, C mutually independent?

My answers:

1. Verify:

$A = \{(1,0,1)\} \cup \{(1,1,0)\}, P(A) = P(\{(1,0,1)\}) + P(\{(1,1,0)\}) - P(\{(1,0,1)\} \cap \{(1,1,0)\}) = \frac{1}{2}$

Hence:

$P(B) = \frac{1}{2}, P(C) = \frac{1}{2}$

$P(A \cap B) = P(\{(1,1,0)\}) = \frac{1}{4}$

$P(A \cap C) = P(\{(1,0,1)\}) = \frac{1}{4}$

$P(C \cap B) = P(\{(0,1,1)\}) = \frac{1}{4}$

Hence

$P(A \cap B) = P(A) \cdot P(B) = \frac{1}{4}$

$P(A \cap C) = P(A) \cdot P(C) = \frac{1}{4}$

$P(C \cap B) = P(C) \cdot P(B) = \frac{1}{4}$

Hence, the events A,B,C are pair-wise independent.

2. $A \cap B \cap C$

$A \cap B \cap C = \emptyset$

3. $P(A \cap B \cap C)$

$P(A \cap B \cap C) = P(\emptyset) = 0 \neq P(A) \cdot P(B) \cdot P(C)$

Hence, the events A,B,C aren't mutually independent.

1.4 The Monty hall problem (*)

1.4 (Q1)

Consider the following game: At a game show there are three seemingly identical doors. Behind one of the doors is a car, and behind the remaining two is a goat.

1. The contestant of the game first gets to choose one of the three doors. The host then opens one of the other two doors to reveal a goat.
2. The contestant now gets a chance to either (a) switch their choice to the other unopened door or (b) stick to their original choice.
3. The host then opens the door corresponding to the contestant's final choice. They get to keep whatever is behind their final choice of door.

Question: does the contestant improve their chances of winning the car if they switch their choice?

Answer:Yes

$P(car) = 1/3, P(goat) = 2/3$

$P(not\ change\ and\ win) = 1 \times \frac{1}{3} + 0 \times \frac{2}{3} = \frac{1}{3}$

$P(chang\ and\ win) = 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3}$

For clarity, we make the following assumptions:

1. The car is assigned to one of the doors at random with equal probability for each door.
2. The assignment of the car and the initial choice of the contestant are independent.
3. Once the contestant makes their initial choice, the host always opens a door which (a) has a goat behind it and (b) is not the contestant's initial choice. If there is more than one such door (i.e. when the contestant's initial choice corresponds to the door with a car behind it) the host chooses at random from the two possibilities with equal probability.

To formalise our problem we introduce the following events for i = 1, 2, 3:

- A_i denotes the event that car is placed behind the i-th door;
- B_i denotes the event that contestant initially chooses the i-th door;
- C_i denotes the event that the host opens the i-th door to reveal a goat. Consider a situation in which the contestant initially selects the first door (B_1) and then the host opens the second door to reveal a goat (C_2). What is $P(A_3|B_1 \cap C_2)$?

Answer:

$P(A_3|B_1 \cap C_2) = \frac{\frac{1}{12}}{\frac{1}{6}} = \frac{1}{2}$

What does this suggest about a good strategy? Should we switch choices?

Yes,when we switch the choices, the probability of success become double

2. Random variables and discrete random variables

2.1 Expectation and variance

2.1 (Q1)

Question:

Suppose that we have random variables X and Y. Recall that the covariance between X and Y is defined by

$Cov(X, Y) = E[(X - \bar{X}) \cdot (Y - \bar{Y})]$

where $\bar{X}$ and $\bar{Y}$ are the expectations of X and Y , respectively.

Now, suppose X and Y are independent. Apply the linearity of expectation and the equivalent condition for independent random variables described above, to prove that $Cov(X, Y) = 0$. To apply the above condition for independent random variables, you may choose proper functions $f1, f2,$ ůůů when necessary.

Proof:

$Cov(X, Y) = E[(X - E(X)) \cdot (Y - E(Y))]$
$= E[XY - Y \cdot E(X) - X \cdot E(Y) + E(X) \cdot E(Y)]$
$= E(XY) - E(Y)E(X) - E(X)E(Y) + E(X)E(Y)$
$= E(XY) - E(X)E(Y)$

Based on the rules of "Equivalent conditions for independent random variables" :

if X,Y are independent,then :

$E(XY) = E(X)E(Y)$

Thus: $Cov(X, Y) = E(XY) - E(X)E(Y)$
$= E(XY) - E(XY) = 0$

2.2 Distributions

Suppose that $\alpha, \beta \in [0, 1]$ with $\alpha + \beta \leq 1$ and let X be a discrete random variable with distribution supported on {0, 3, 10}. Suppose that $P(X = 3) = \alpha$ and $P(X = 10) = \beta$ and $P(X \notin \{0, 3, 10\}) = 0$.

2.2 (Q1)

Questions:

Expectation and variance of a discrete random variable

1. What is the probability mass function p_X for X?
2. What is the expectation of X?
3. What is the variance of X?
4. What is the standard deviation of X?

Answers:

1. Probability mass function pX:

$$p_X(x) = \begin{cases} 1-\alpha-\beta & \text{if x=0,} \\ \alpha & \text{if x=3,} \\ \beta & \text{if x=10,} \\ 0 & \text{otherwise.} \end{cases}$$

2. Expectation:

$E(X) = \sum_{x \in R} x \cdot p_X x$
$= \alpha \times 3 + \beta \times 10 + (1-\alpha-\beta) \times 0 + \sum_{x \in R \setminus 0,3,10} x \cdot 0$
$= 3\alpha + 10\beta$

3. Variance:

$Var(X) = E[(X - E[X])^2]$
$= (1-\alpha-\beta) \cdot (3\alpha + 10\beta)^2 + \alpha \cdot (3 - 3\alpha - 10\beta)^2 + \beta \cdot (10 - 3\alpha - 10\beta)^2$
$= 100\beta + 9\alpha - (3\alpha + 10\beta)^2$

4. Standard deviation:

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

2.2 (Q2)

Questions:

Distribution and distribution function.

Recall that the distribution of a random variable maps a subset S of R to a real number.

1. Write down the distribution P_X of X. You can use the indicator function 1_S to "indicate" if a number is in the set S.
2. Write down the distribution F_X of X

Answers:

1. $P_X(S)$:

$P_X(S) = \alpha \cdot 1_S(3) + \beta \cdot 1_S(10) + (1 - \alpha - \beta) \cdot 1_S(0)$

2. $F_X(A)$:

$$F_X(S) = \begin{cases} 0 & \text{if } x < 0, \\ 1 - \alpha - \beta & \text{if } 0 \leq x < 3, \\ 1 - \beta & \text{if } 3 \leq x < 10, \\ \beta & \text{if } 10 \leq x \end{cases}$$

2.2 (Q3)

Question:

Define a new random variance $Y = X_1 + X_2 + ... + X_n$ where $X_1, X_2, ..., X_n$ are independent random variables, each of which has the same distribution as the random variable X.

Derive an expression for the variance of Y. You can use the conclusion from Question 2.1 (Q1).

Answer:

$Var(Y) = Var(X_1 + ... + X_n) = Var(X_1) + ... + Var(X_n) = n \cdot (100\beta + 9\alpha - (3\alpha + 10\beta)^2)$

2.2 (Q4)

Explore the distribution of the sum of n independent random variables when n is large.

Let Y be defined above, and additionally, let $\alpha = 0.2$ and $\beta = 0.3$. Recall that X_1, X_2, ůůů, X_n are discrete random variables. Explain why the random variable Y is discrete.

Now let's explore the probability mass function of Y using R programming, and see its behaviour when n is large.

(Step 1). First, Y can be viewed as a generalization of the binomial random variables. A binomial random variable can be written as the sum of a sequence of Bernoulli random variables (which take values from {0, 1}). Here Y can be written as the sum of Xi, which takes values from {0, 3, 10} (hence the distributions of Xi are known as generalized Bernoulli distributions). In R, we can obtain samples of Y by using the function rmultinom().

Now let n = 3 and generate 50000 samples of X_1, X_2, X_3 using the rmultinom() function. Store the samples in an object called samples_Xi. Then based on samples_Xi, create 50000 samples of Y and store it in a data-frame called samples_Y, consisting of a single column called Y.

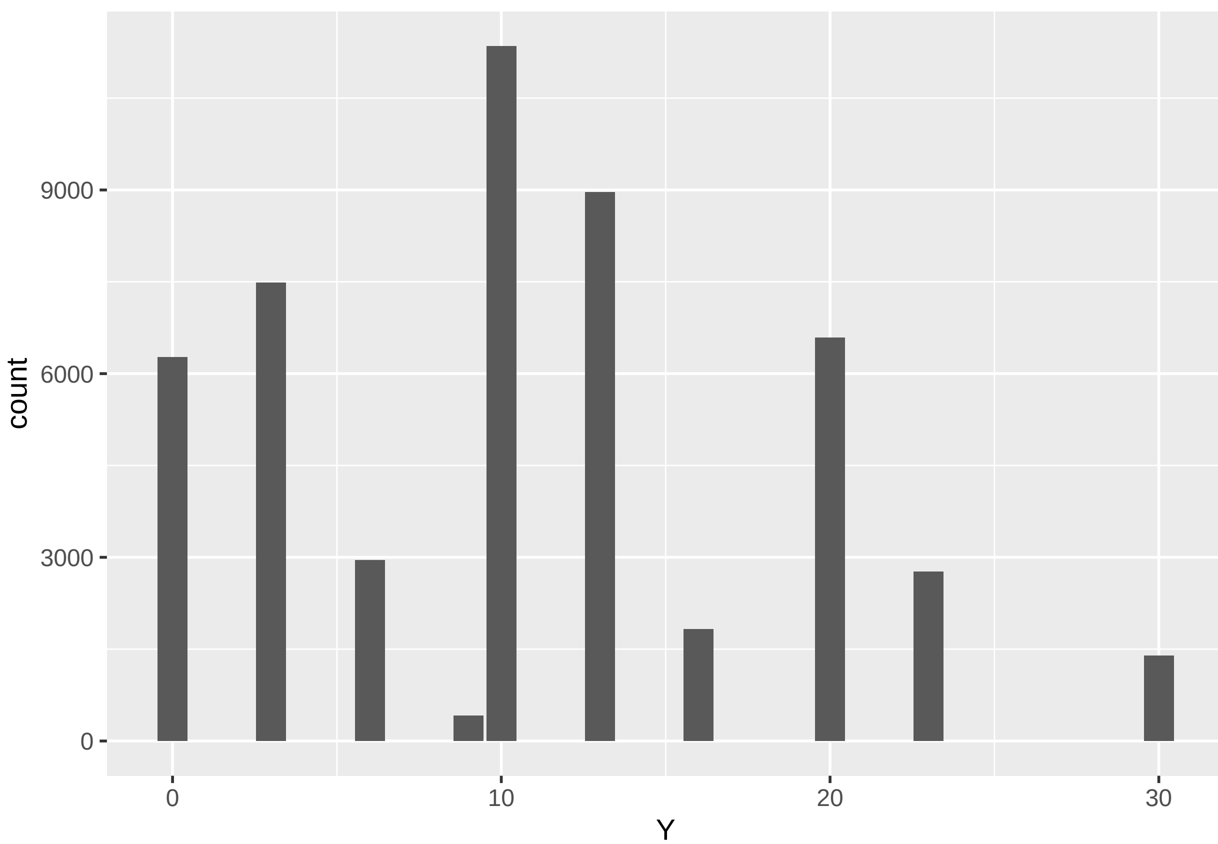
```
samples_Xi<-rmultinom(50000,3,prob=c(0.5,0.2,0.3))
```

```
tem_Y<-matrix(c(0,3,10),nrow=1)
samples_Y<-data.frame(Y=t(tem_Y%*%samples_Xi))
samples_Y%>%head(10)
```

```
##     Y
## 1  10
## 2  20
## 3   3
## 4  20
## 5  10
## 6  10
## 7  23
## 8  10
## 9  16
## 10 13
```

(Step 2) Use the **ggplot** and **geom_bar()** function to create a bar plot for the samples of Y . Notice that Y takes values from a subset of [0, 30]

```
ggplot(data=samples_Y,aes(x=Y))+geom_bar()
```



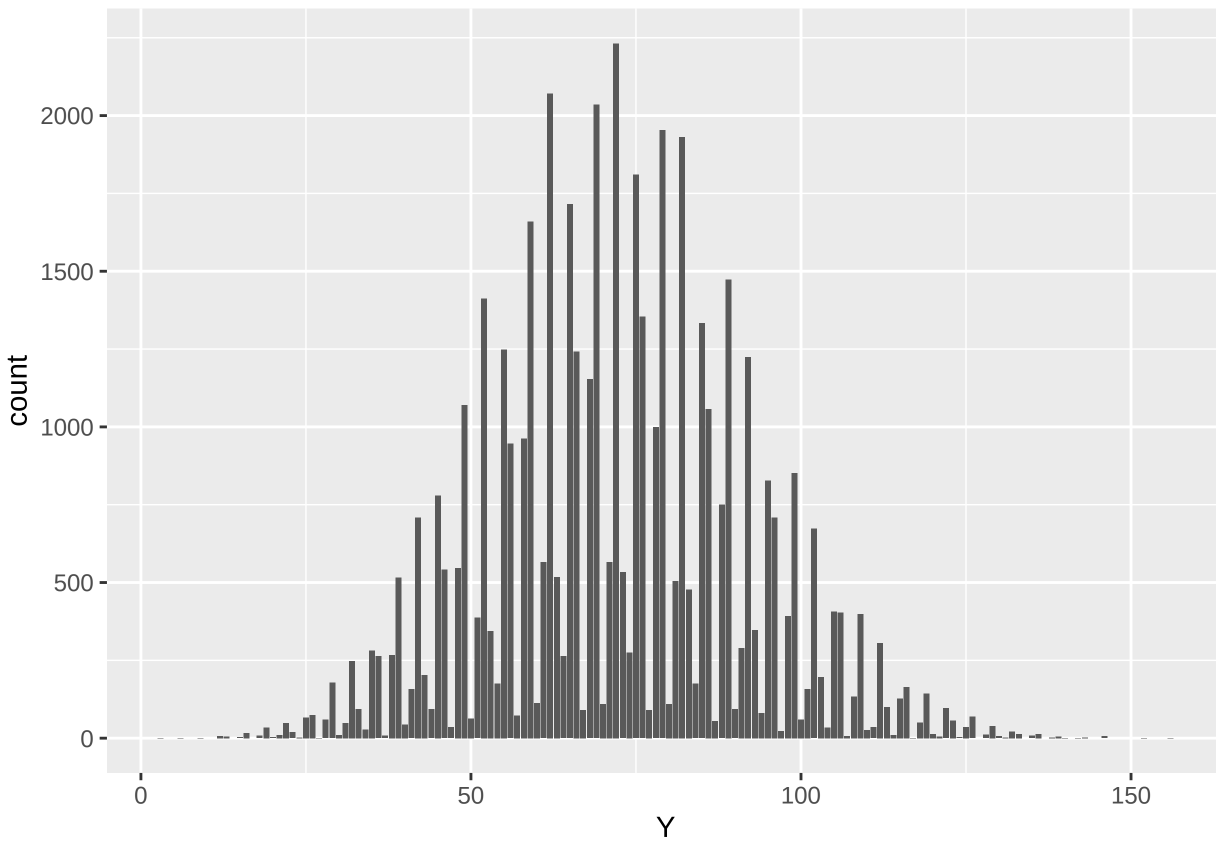
(Step 3) Now, increase the values of n by setting n = 20, and repeat Step 1 and Step 2 to create a new bar plot for the samples of Y . What are the maximum and minimum values of the samples? What is the sample range?

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

```
samples_Y<-data.frame(Y=t(tem_Y%*%samples_Xi))
samples_Y%>%head(10)
```

```
##        Y
## 1   110
## 2    52
## 3   101
## 4    75
## 5    79
## 6    55
## 7    82
## 8    56
## 9    72
## 10  102
```

```
ggplot(data=samples_Y,aes(x=Y))+geom_bar()
```



Answers:

the maximum value of the sample is 200,and the minimumvalue of the sample is 0.The sample range is [0,200].

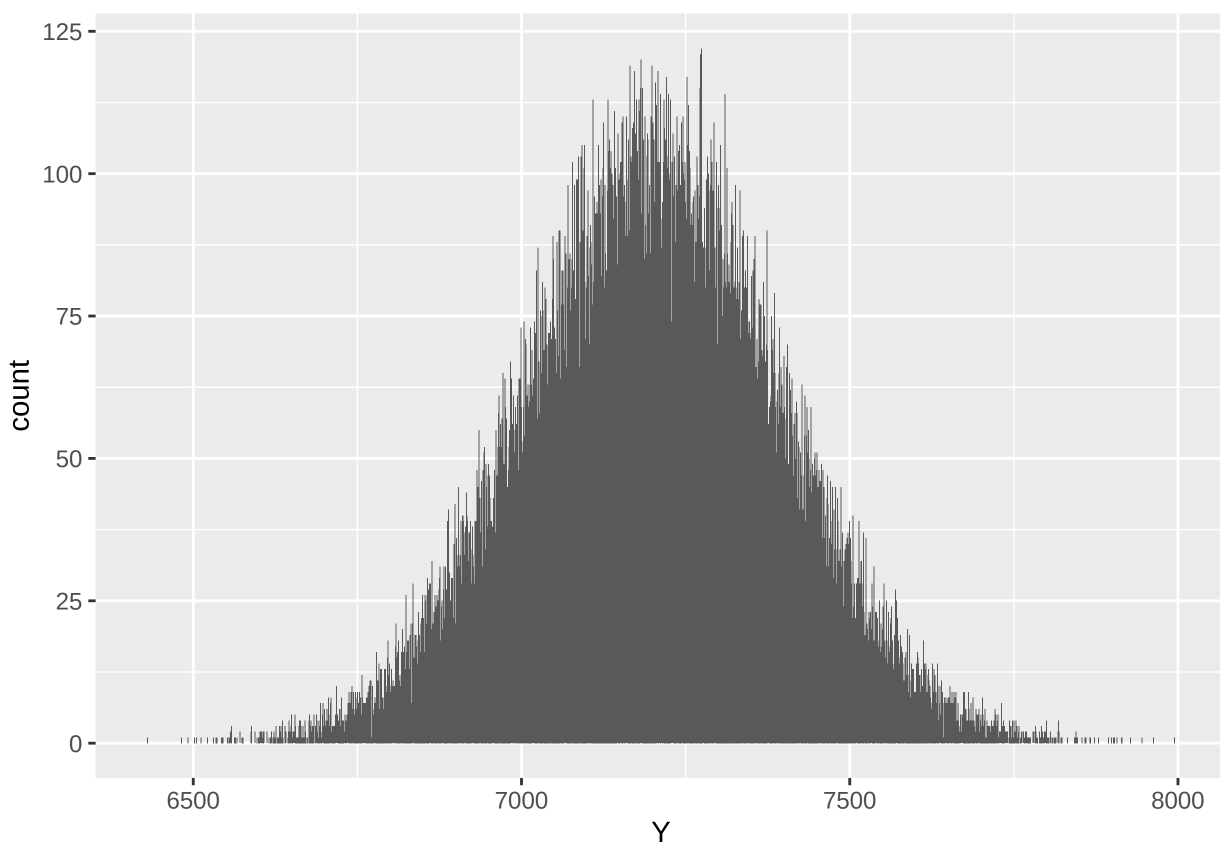
(Step 4) Next, increase n to 2000 and do the plot again.

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

```
samples_Y<-data.frame(Y=t(tem_Y%*%samples_Xi))
samples_Y%>%head(10)
```

```
##        Y
## 1  7260
## 2  7117
## 3  7593
## 4  7042
## 5  7104
## 6  7507
## 7  7180
## 8  7071
## 9  7351
## 10 7116
```

```
ggplot(data=samples_Y,aes(x=Y))+geom_bar()
```



Notice that as we increase n, the distribution of Y tends to follow a bell-like shape. While Y is a discrete random variable (no matter how large n is), the distribution looks closer to the distribution of a continuous random variable, which is known as a Gaussian random variable.