GENDER INEQUALITY ASSIGNMENT

CISM-I

ADITYA DUBEY

ENROLL – 382616

Loading the Dataset:

First, we load the dataset into R using the code

```
# Loading Responses Data
responses <- read.csv("Responses.csv",header =TRUE,skip=26)
head(responses,n=3)

# Extracting the required columns and changing their names
columns <-
c("Time","fst_gend","B2C","G2C","B2B","G2B","BB3C","GG3C","BG3C","GG3B","Age","Pinco
de","Gender")
responses <- responses[,c(1,2,7,11,15,20,25,29,33,37,42:44)]
names(responses) <- columns
head(responses,n=3)
```

Generating Probabilities required to generate the hypothetical population:

```
# Creating a Function to convert all columns to tables and then to data frame
# We will use this to easily assign probabilities to each option selected by respondents

myconvert <- function(x){
  x <- as.data.frame(table(x))
}

# Creating lists with converted data frames
list1 <- apply(responses[,c(2,5,6,10)],MARGIN = 2,FUN=myconvert)
list2 <- apply(responses[,c(3,4,7:9)],MARGIN = 2,FUN=myconvert)

# Assigning lower and upper probabilities to each option
for (i in (1:length(list1))){
  list1[[i]]$LowProb <- c(20,65,35,50,50,0,80)/100
  list1[[i]]$UppProb <- c(35,80,50,65,50,20,100)/100
}

for (i in (1:length(list2))){
  list2[[i]]$LowProb <- c(60,20,40,0,80,0)/100
  list2[[i]]$UppProb <- c(80,40,60,0,100,20)/100
}
```

```
# Now we will generate random probabilities for all respondents and take mean to get an overall probability
myprob <- data.frame(sl=1:341)

# Generating random probabilities for elements in list1 with 7 options
for (i in (1:length(list1))){
  probs <- numeric()
  for (j in (1:7)){
    probs <- append(probs,runif(list1[[i]]$Freq[j],list1[[i]]$LowProb[j],list1[[i]]$UppProb[j]))
  }
  myprob <- cbind(myprob,probs)
}

# Generating random probabilities for elements in list2 with 6 options
for (i in (1:length(list2))){
  probs <- numeric()
  for (j in (1:6)){
    probs <- append(probs,runif(list2[[i]]$Freq[j],list2[[i]]$LowProb[j],list2[[i]]$UppProb[j]))
  }
  myprob <- cbind(myprob,probs)
}
names(myprob) <- c("SL",names(list1),names(list2))
head(myprob)
# Storing the different probabilities which will be mean of the random probabilities in respective column

prob_B1 <- mean(myprob$fst_gend)        # Probability of first child to be a boy
prob_G1 <- 1-prob_B1                    # Probability of first child to be a girl
prob_B2C <- mean(myprob$B2C)            # Probability of going for 2nd child if 1st is a boy
prob_G2C <- mean(myprob$G2C)            # Probability of going for 2nd child if 1st is a girl
prob_B2B <- mean(myprob$B2B)            # Probability of 2nd child to be a boy if 1st is a boy
prob_G2B <- mean(myprob$G2B)            # Probability of 2nd child to be a boy if 1st is a girl
prob_BB3C <- mean(myprob$BB3C)          # Probability of going for 3rd child if first two are both boys
prob_BB3B <- 0.3                        # Probability of 3rd child to be a boy if first two are both boys
prob_BG3C <- mean(myprob$BG3C)          # Probability of going for 3rd child if first two are boy and girl
prob_BG3B <- 0.5                        # Probability of 3rd child to be a boy if first two are boy and girl
prob_GG3C <- mean(myprob$GG3C)          # Probability of going for 3rd child if first two are both girls
prob_GG3B <- mean(myprob$GG3B)          # Probability of 3rd child to be a boy if first two are both girls
```

FOR 100 FAMILIES:

```
# Generating population for 100 families
tot_fam <- 100
child1 <- sample(c("B","G"),100,prob = c(prob_B1,1-prob_B1),replace = T)
tab1 <- as.data.frame(table(child1))
child_1B2 <-
sample(c("B","G"),round(tab1$Freq[tab1$child1=="B"]*prob_B2C),prob=c(prob_B2B,1-
prob_B2B),replace = T)
tab2 <- as.data.frame(table(child_1B2))
child_1G2 <-
sample(c("B","G"),round(tab1$Freq[tab1$child1=="G"]*prob_G2C),prob=c(prob_G2B,1-
prob_G2B),replace = T)
tab3 <- as.data.frame(table(child_1G2))
child_BB3 <-
sample(c("B","G"),round(tab2$Freq[tab2$child_1B2=="B"]*prob_BB3C),prob=c(prob_BB3B,
1-prob_BB3B),replace=T)
tab4 <- as.data.frame(table(child_BB3))
child_GG3 <-
sample(c("B","G"),round(tab3$Freq[tab3$child_1G2=="G"]*prob_GG3C),prob=c(prob_GG3B
,1-prob_GG3B),replace=T)
tab5 <- as.data.frame(table(child_GG3))
child_BG3 <-
sample(c("B","G"),round(tab2$Freq[tab2$child_1B2=="G"]*prob_BG3C),prob=c(prob_BG3B,
1-prob_BG3B),replace=T)
tab6 <- as.data.frame(table(child_BG3))
child_GB3 <-
sample(c("B","G"),round(tab3$Freq[tab3$child_1G2=="B"]*prob_BG3C),prob=c(prob_BG3B,
1-prob_BG3B),replace=T)
tab7 <- as.data.frame(table(child_GB3))

final_tab <- cbind(tab1,tab2$Freq,tab3$Freq,tab4$Freq,tab5$Freq,tab6$Freq,tab7$Freq)
population <- data.frame(gender=final_tab$child1,total = rowSums(final_tab[,c(2:8)]))
sex_ratio <-
(population$total[population$gender=="B"]/population$total[population$gender=="G"])
prop_boys <- population$total[population$gender=="B"]/sum(population$total)*100
prop_girls <- population$total[population$gender=="G"]/sum(population$total)*100
```

FOR 1,00,00,000 (1 Crore) Families

```
# Generating population for 10000000 families
tot_fam <- 10000000
child1 <- sample(c("B","G"),10000000,prob = c(prob_B1,1-prob_B1),replace = T)
tab1 <- as.data.frame(table(child1))
child_1B2 <-
sample(c("B","G"),round(tab1$Freq[tab1$child1=="B"]*prob_B2C),prob=c(prob_B2B,1-
prob_B2B),replace = T)
tab2 <- as.data.frame(table(child_1B2))
child_1G2 <-
sample(c("B","G"),round(tab1$Freq[tab1$child1=="G"]*prob_G2C),prob=c(prob_G2B,1-
prob_G2B),replace = T)
tab3 <- as.data.frame(table(child_1G2))
child_BB3 <-
sample(c("B","G"),round(tab2$Freq[tab2$child_1B2=="B"]*prob_BB3C),prob=c(prob_BB3B,
1-prob_BB3B),replace=T)
tab4 <- as.data.frame(table(child_BB3))
child_GG3 <-
sample(c("B","G"),round(tab3$Freq[tab3$child_1G2=="G"]*prob_GG3C),prob=c(prob_GG3B
,1-prob_GG3B),replace=T)
tab5 <- as.data.frame(table(child_GG3))
child_BG3 <-
sample(c("B","G"),round(tab2$Freq[tab2$child_1B2=="G"]*prob_BG3C),prob=c(prob_BG3B,
1-prob_BG3B),replace=T)
tab6 <- as.data.frame(table(child_BG3))
child_GB3 <-
sample(c("B","G"),round(tab3$Freq[tab3$child_1G2=="B"]*prob_BG3C),prob=c(prob_BG3B,
1-prob_BG3B),replace=T)
tab7 <- as.data.frame(table(child_GB3))

final_tab1 <- cbind(tab1,tab2$Freq,tab3$Freq,tab4$Freq,tab5$Freq,tab6$Freq,tab7$Freq)
population1 <- data.frame(gender=final_tab1$child1,total = rowSums(final_tab1[,c(2:8)]))
sex_ratio1 <-
(population1$total[population1$gender=="B"]/population1$total[population1$gender=="G
"])
prop_boys1 <- population1$total[population1$gender=="B"]/sum(population1$total)*100
prop_girls1 <- population1$total[population1$gender=="G"]/sum(population1$total)*100
```

RESULT:

	100 Families		1 Crore Families	
Population	B = 106	G = 69	B = 10800799	G = 6175328
Proportion	B = 60.57	G = 39.43	B = 63.62	G = 36.38
Sex ratio	1.54		1.75	

CONCLUSION:

Based on the results of **responses of 341 individuals**, the hypothetical population generated had a sex ratio of –

1. **1.54**, male to female, when **started with 100 families**.
2. **1.75,** male to female, when **started with 1 Crore (1,00,00,000) families.**

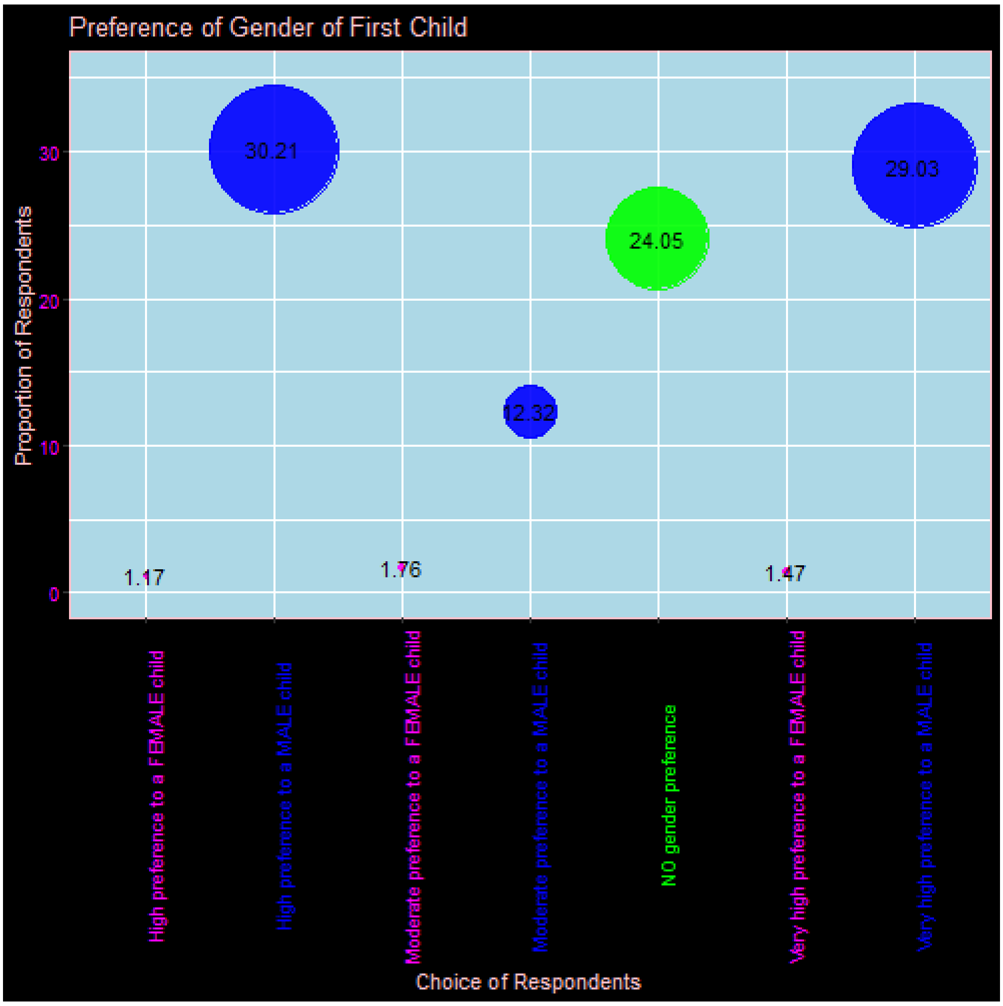
Thus, in both the cases, we found that the number of males were much greater than the number of females in the generated population, leading to gender inequality.

SHORT ANALYSIS OF RESPONSES:

PLOT1: Preference of gender in case of 1st child:

```
q1 <- list1[[1]]
q1$Freq <-(q1$Freq/sum(q1$Freq))*100
library(ggplot2)
ggplot(data=q1 ,aes(x=levels(as.factor(x)),y=Freq))+

geom_point(size=q1$Freq,colour=c("magenta","blue","magenta","blue","green","magenta"
,"blue"),
          alpha=0.9)+ylim(0,35)+
  theme(axis.text.x = element_text(colour=c("magenta","blue","magenta","blue",
                                  "green","magenta","blue"),angle=90),
      axis.text.y = element_text(colour="magenta"),plot.title = element_text(colour="pink"),
      axis.title.x =element_text(colour="pink"), axis.title.y=element_text(colour="pink"))+
  geom_text(aes(label = round(Freq,2)))+labs(x="Choice of Respondents",y="Proportion of
Respondents")+
  ggtitle("Preference of Gender of First Child")+
  theme(panel.background = element_rect(fill = 'light blue', colour = 'pink'),
      plot.background = element_rect(fill="black"))
```

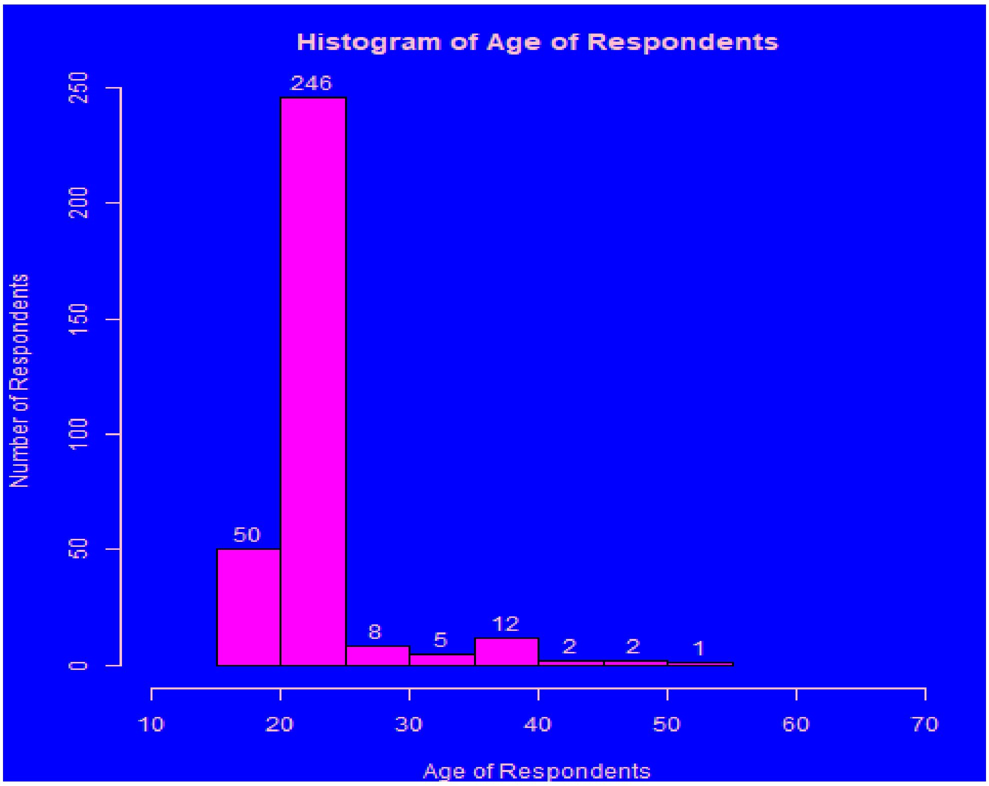


Interpretation: From the Plot, we can clearly see that-

1. **Approximately 71%** of respondents are of the opinion that in case of a first child our society gives more **preference to the birth of a male child**.
2. What is shocking, is that only a minute **4.4%** respondents have said that more **preference is given towards a female child**.

Plot2: Histogram of Age of Respondents:

```
# Histogram of Age of Respondents
par(bg="blue",fg="pink")
hist(responses[responses$Age<=70 &
responses$Age>=10,"Age"],xlim=c(10,70),col=c("magenta"),border="black",
   labels=TRUE,main="Histogram of Age of Respondents",xlab="Age of
Respondents",ylab="Number of Respondents",
   col.main="pink",col.axis="pink",col.lab="pink")
```



Plot3: Distribution of respondents based on gender:

```
# Distribution of Respondents Based upon Gender
par(bg="blue",fg="pink")
barplot(table(responses$Gender)[c(2,3)],col=c("magenta","light blue"),ylim=c(0,200),xlab =
"Respondent Gender",
      ylab="Number of Respondents",main="Distribution of Respondents based on
Gender",col.main="pink",
      col.axis="pink",col.lab="pink")
```