Lab2

Jiahuan Li

2023-01-25

```r
afdat <-
  read.csv("d:/Users/Lijh/Desktop/710 R & statistics/lab2/AfrPlots.csv",
           header = TRUE)
library(ggplot2)
```

Problem 1

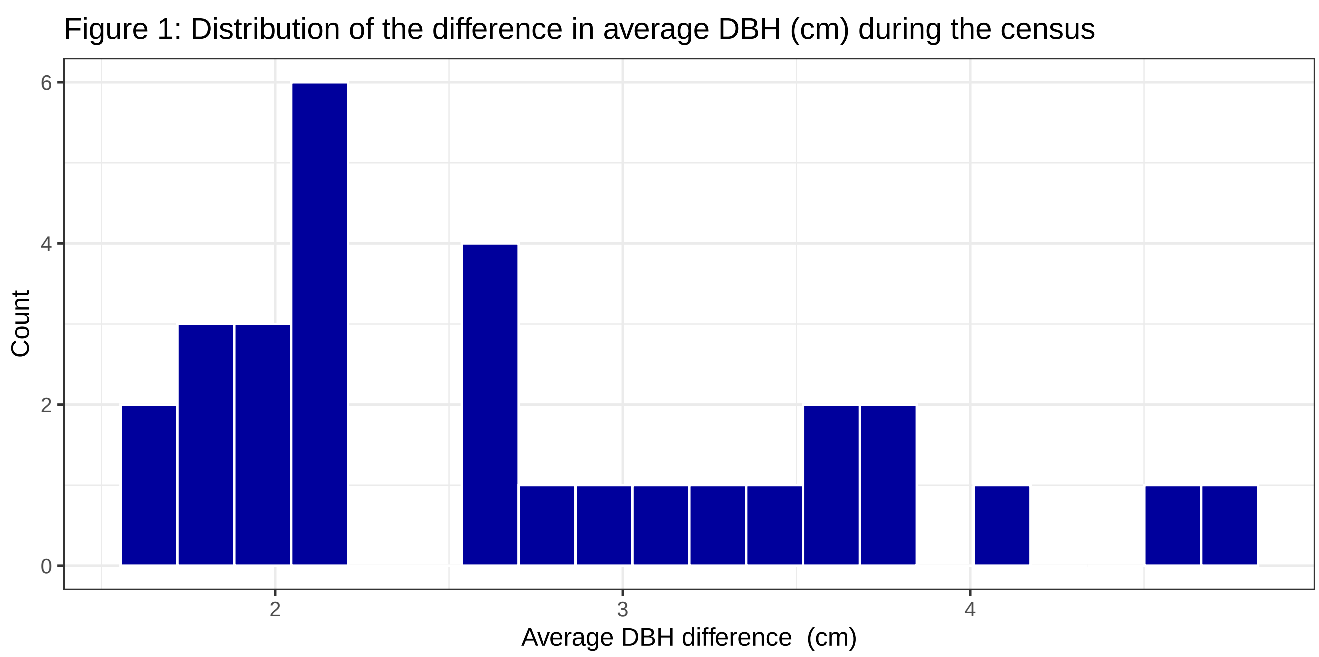
a

```r


dukeblue <- "#00009C"
fig1 = ggplot(afdat, aes(x = MeanGr))  +
  geom_histogram(fill = dukeblue, colour = "white", bins = 20) +
  xlab(expression(paste("Average DBH difference", "  (cm)"))) +
  ylab("Count") +
  ggtitle('Figure 1: Distribution of the difference in average DBH (cm) during the census') +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme_bw()

fig1
```

histogram

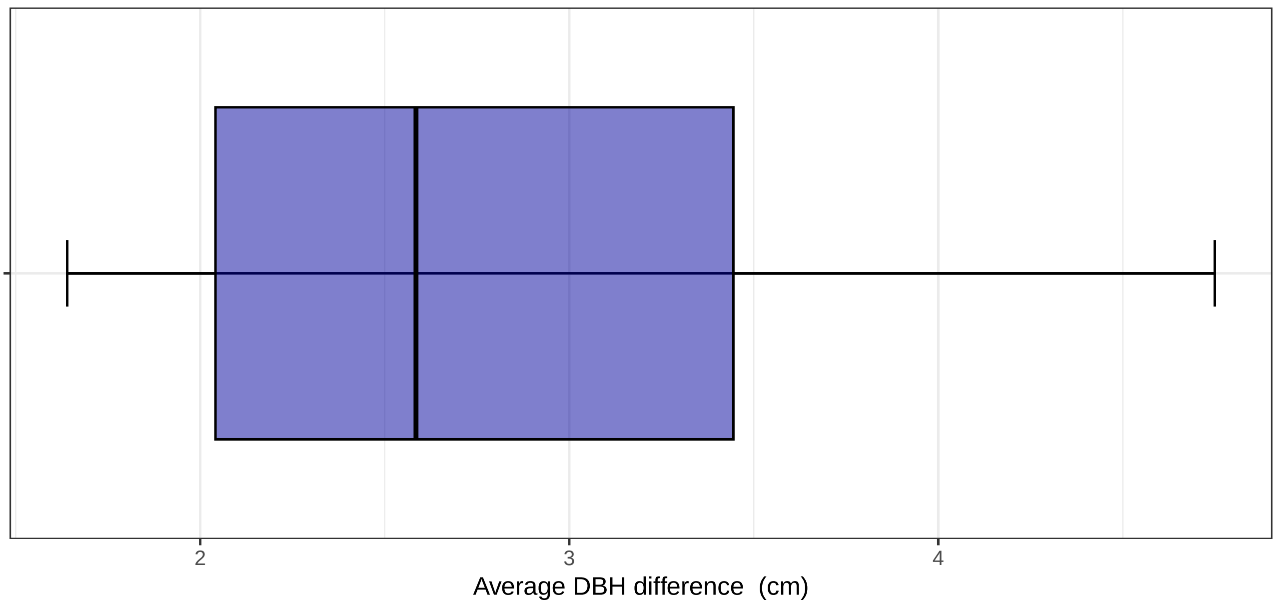


```r
fig2 = ggplot(afdat, aes(x = MeanGr, y = "")) +
  stat_boxplot(geom = "errorbar", width = 0.15, color = 1) +
  geom_boxplot(fill = dukeblue, alpha = 0.5, color = 1, outlier.colour = 2) +
  xlab(expression(paste("Average DBH difference", "  (cm)"))) +
  ylab("") +
  ggtitle('Figure 2: Boxplot of the distribution of the average DBH (cm) difference') +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme_bw()

fig2
```

boxplot

Figure 2: Boxplot of the distribution of the average DBH (cm) difference



b

```r
kurtosis <- function(y) {
    n <- length(y)
    kurt <- 1 / (n * sd(y) ^ 4) * sum((y - mean(y, na.rm = TRUE)) ^ 4) - 3
    kurt
}

MeanGr1 = afdat$MeanGr[!is.na(afdat$MeanGr)]
answer = kurtosis(MeanGr1)
print(paste("answer = ", answer))
```

```
## [1] "answer =  -0.728889486347993"
```

c

```
library (scales)

Mean = mean(afdat$MeanGr, na.rm = TRUE)
Median = median(afdat$MeanGr, na.rm = TRUE)
Sd = sd(afdat$MeanGr, na.rm = TRUE)
COV = percent(Sd / Mean, accuracy = .01)

print(paste("Mean =", Mean))
```

```
## [1] "Mean = 2.70576275873333"
```

```
print(paste("Median =", Median))
```

```
## [1] "Median = 2.58466596"
```

```
print(paste("Sd =", Sd))
```

```
## [1] "Sd = 0.894471691835576"
```

```
print(paste("COV =", COV))
```

```
## [1] "COV = 33.06%"
```

Problem 2

a

```
p1 = sum(dbinom(6:20,20,0.5))
p2 = pbinom(q = 5, 20, 0.5, lower.tail = FALSE)

print(paste("Answer from dbinom =", p1))
```

```
## [1] "Answer from dbinom = 0.979305267333984"
```

```
print(paste("Answer from pbinom =", p2))
```

```
## [1] "Answer from pbinom = 0.979305267333984"
```

b

```
pbi = function (x, n, p) {
  bi = factorial(n)/(factorial(x)*factorial(n-x))*p^x*(1-p)^(n-x)
  bi
}

p3 = sum(pbi(6:20,20,0.5))
print(paste("Answer from equation =", p3))
```

```
## [1] "Answer from equation = 0.979305267333984"
```

Problem 3

```
p_17_correct = pbinom(16, 20, 0.25, lower.tail = FALSE)
p_17_correct1 = sum(dbinom(17:20, 20, 0.25))

print(paste("Probability of answering 17 or more answers correctly =", p_17_correct1))
```

```
## [1] "Probability of answering 17 or more answers correctly = 2.96049620374107e-08"
```

Problem 4

a

```
p_9 = dpois(9, 4)
Poiss = function (x, l) {
  pois = l^x/factorial(x)*exp(-l)
  pois
}
q4_1 = Poiss(9,4)

print(paste("P(X = 9) =", q4_1))
```

```
## [1] "P(X = 9) = 0.0132311916910503"
```

b

```
P_9_13 = sum(dpois(9:13,4))
P_9_13_1 = ppois(13,4) - ppois(8,4)
P_9_infinite = ppois(8, 4, lower.tail = F)
P_9_infinite_1 = 1 - sum(dpois(0:8, 4))
```

Answer for P(X >= 9) using dpois() is `P_9_13` with limitation of the maximum number as 13;

Similarly, the answer for P(X >= 9) using ppois() is also `P_9_13_1` with limitation of the maximum number as 13.