Stat 222 - Exercise 1

kclazaro

2023-02-21

Exercise. Refer to the height-weight data set in the Lesson 2 Folder in the Class Drive. Assume the data is taken from certain population through random sampling. Let X,Y be the height and weight, respectively, of a randomly selected person in the population. Find the observed value of the following statistics in Excel or R.

1. $\overline{X}$
2. S_Y^2 (sample variance of Y)
3. $Y_{(2)}$
4. $M'_{X,2}$ (second sample raw moment of X)

Using the data set, we get the following summary statistics.

```
#Load the 'readxl' library to be able to parse Excel files. Then we assign the variable name 'weight_he
library(readxl)
weight_height <- read_excel("Downloads/weight-height.xlsx")
summary(weight_height)
```

```
##     Gender              Height          Weight
##  Length:10000       Min.   :54.26   Min.   : 64.7
##  Class :character   1st Qu.:63.51   1st Qu.:135.8
##  Mode  :character   Median :66.32   Median :161.2
##                     Mean   :66.37   Mean   :161.4
##                     3rd Qu.:69.17   3rd Qu.:187.2
##                     Max.   :79.00   Max.   :270.0
```

To get the mean of the height column, we use the following code:

```
# Get the mean of the Height column by column name
xbar <- mean(weight_height$Height)
```

This means that the mean height of the data set is...

```
## [1] 66.36756
```

For the sample variance of Y, we use the following code:

```
# Get the variance of the Weight column by column name
var_y <- var(weight_height$Weight)
```

Using this, the variance of the weight column is...

Statistics	Answer
$\overline{X}$	66.368
S^2_Y	1030.952
$Y_{(2)}$	68.983
$M{'}_{X,2}$	4419.455

```
## [1] 1030.952
```

In order to obtain the second lowest value in the Weight column, we will use the following code:

```
# Sort the Weight column from lowest to highest, and we assign it to the 'weight_order' variable.
weight_order <- sort(weight_height$Weight)
```

From this, it's easy to obtain the second smallest value, which is...

```
## [1] 68.98253
```

Lastly, to get the second sample raw moment of height, we can use the moments package.

```
# Load the moments package then get the second raw moment by specifying the order.
library(moments)
second_raw_moment_height <- moment(weight_height$Height, order=2)
```

Which will then give us the second sample raw moment for the height column...

```
## [1] 4419.455
```

In summary, the following are the observed values for the statistics: