Day 3 of #100daysofmathandstats: Estimates of Location

By Harsh Kathiriya

Mean

The most basic estimate of location is the mean, or *average* value.

$$\text{Mean} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Trimmed mean

A variation of the mean is a trimmed mean, which you calculate by dropping a fixed number of sorted values at each end and then taking an average of the remaining values

$$\text{Trimmed mean} = \bar{x} = \frac{\sum_{i=p+1}^{n-p} x_{(i)}}{n - 2p}$$

Weighted mean

Another type of mean is a weighted mean, which is calculated by multiplying each observation with corresponding weight.

$$\text{Weighted mean} = \bar{x}_w = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

R implementation

```
> data <- datasets::mtcars
> # mean
> base::mean(data$mpg)
[1] 20.09062
> # trimmed_mean
> base::mean(data$mpg, trim = 0.2)
[1] 19.22
> # weighted mean
> stats::weighted.mean(data$mpg, data$cyl)
[1] 18.65455
> # median
> stats::median(data$mpg)
[1] 19.2
> # weighted median
> # uncomment the below line if you don't have spatstat library installed
> # install.packages('spatstat')
> library(spatstat.geom)
> spatstat.geom::weighted.median(data$mpg, data$cyl)
[1] 17.55
> # percentile
> stats::quantile(data$mpg)
    0%     25%     50%     75%    100%
10.400 15.425 19.200 22.800 33.900
> # 25th 50th and 95h percentile
> stats::quantile(data$mpg, c(.25,.5,.95))
   25%     50%     95%
15.425 19.200 31.300
> # since 50th percentile is equal to median following shoud be TRUE
> stats::quantile(data$mpg, c(.5)) == stats::median(data$mpg)
 50%
TRUE
>
```

Python implementation

```python
1  import pandas as pd
2  data = pd.read_csv('/content/mtcars.csv')
3  # mean
4  print('Mean: ',data['mpg'].mean())
5
6  # trimmed mean from scipy
7  import scipy.stats as sc
8  print('Trimmed Mean: ',sc.trim_mean(data['mpg'], 0.1))
9
10 # median
11 print('Median', data['mpg'].median())
12
13 import numpy as np
14 print('Weighted Mean: ',np.average(data['mpg'], weights = data['cyl']))
15
16 #pip install wquantiles
17 import wquantiles
18 # weighted median
19 print('Weighted Median: ',wquantiles.median(data['mpg'], weights = data['cyl']))
```

```
Mean:   20.090624999999996
Trimmed Mean:   19.696153846153848
Median 19.2
Weighted Mean:   18.654545454545453
Weighted Median:   17.8
```

Thank you

Don't forget to post your questions and feedbacks on the post.