

Statistics 624: Statistical Computing

Homework # 7

Due Friday, 5 Oct., 5pm

In this assignment, please complete, organize, and turn in **electronically** to `stat624@stat.byu.edu`.

This assignment compares two ‘algorithms’ for computing the sample variance, s^2 .

1. By definition, the sample variance is

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Because this formula requires calculating \bar{x} before calculating s^2 , this program requires reading all the data into memory.

Write a C program with the following features:

- read in a vector from a file
- compute \bar{x}
- compute $\sum_{i=1}^n (x_i - \bar{x})^2$
- compute s^2
- write the results in a file that contains:

```
data read from filename
```

```
n = ***
```

```
xbar = ***
```

```
s^2 = ***
```

```
sum(xi-xbar)^2 = ***
```

2. It is common in textbooks to see a ‘computational form’ for the sample variance given by

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right]$$

This formula avoids requiring all the values are stored in memory since as each observation is read from the file the two sums can be ‘updated.’

Write a C program with the following features:

- read data from a file
- after each row of the data file is read, add the values x_i and x_i^2 to their respective sums.
- after the last observation is read, compute \bar{x} and s^2
- write the results in a file that contains:

```
data read from filename
```

```
n = ***
```

```
xbar = ***
```

```
s^2 = ***
```

```
sum xi = ***
```

```
sum xi^2 = ***
```

3. Test your code on the 25 observations in `vartest1.txt`. Are the results for each program the same? How does the result compare to the computed value in R?
4. Test your code on the 25 observations in `vartest2.txt`. Are the results for each program the same? How does the result compare to the computed value in R?