PHYS 139 Assignment 3

Yuntong Zhou

November 2022

Problem 1 and 2. Gamma(4,1)

The analytical form for a gamma probability distribution function is,

$$f(k,\theta) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}$$

Therefore, Gamma(4,1) is,

$$f(4,1) = \frac{1}{\Gamma(4)\theta^4} x^3 e^{-x}$$

The mean and variance of the analytical form is,

$$\bar{x} = k\theta = 1 \times 4 = 4$$
$$\sigma^2 = k\theta^2 = 4 \times 1 = 4$$

Steal Wilson-Hilferty transformation to calculate the mean,

$$median = \nu(k, \theta = 1) = k(1 - \frac{1}{9k})^3$$

k=4,

$$\nu = 3.676$$

Therefore, the theoretical mean over median ratio is approximately,

$$\frac{4}{3.676} = 1.0882$$

The jackknife sampling and bootstrap sampling methods are summarized in *problem12.py*. A sample output is saved in *output.txt*. The sample distribution and invisible distribution are shown below,

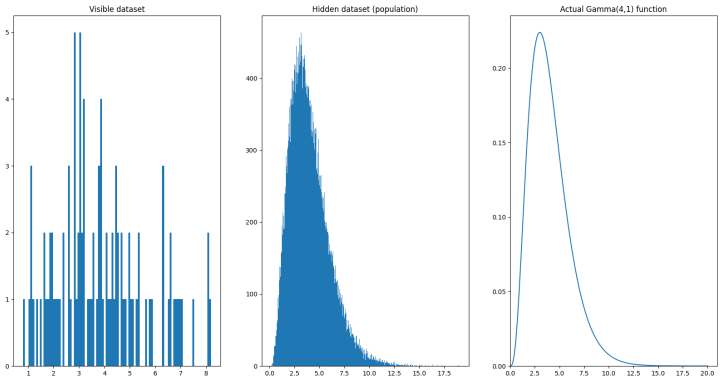

Figure 1: Sample distribution (n=100) vs. Invisible distribution

```
The mean of the sample with size of 100 is:  3.858005424532575
The variance of the sample is:  3.1265716073473926
======================Problem 1======================
_____________________Using Bootstrap sampling_____________________________
The average ratio of mean over median is  1.0659494252210748
The standard error in ratio of mean over median is  0.043317429047680114
_____________________________________________________________________________
_____________________Using Invisible distribution sampling_____________________
The average ratio of mean over median is  1.0943697889606734
The standard error of mean over median is  0.043681289639969534
======================Problem 2======================
_____________________Using Jackknife sampling_____________________________
The average ratio of mean over median is  1.055253893966224
The standard error of mean over median is  0.18269654774536837
_____________________________________________________________________________
_____________________Using Invisible distribution sampling_____________________
The average ratio of mean over median is  1.0943859292689706
The standard error of mean over median is  0.04171964033295947
```

Figure 2: Sample output for problem 1 and problem 2

The mean over ratio statistic is very close to the theoretical value for bootstrap sampling. Jackknife sampling is slightly worse but still very close.

Problem 3

The solution is illustrated in *problem3.py*. The expected mean of the mixture of two normal distributions is,

$$\overline{x} = 0.3 \times 1 + 0.7 \times 4 \approx 3.1$$

The mean for the dataset (size=100) and the approximated mean for the invisible dataset is,

```
Dataset statistics:
The mean of the sample dataset is  3.037141734170151
The mean of the invisible dataset is  3.1016645218974617
```

Figure 3: Sample output mean

It is difficult to obtain an closed analytical form of median for the mixture distribution, though the median of the invisible dataset can be approximated.

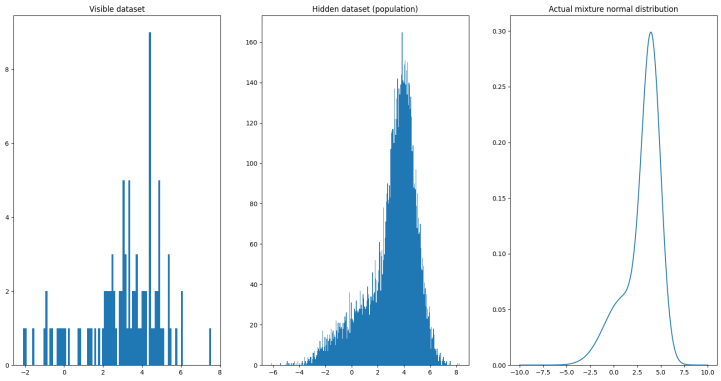

Figure 4: Sample distribution vs. invisible distribution

The same jackknife and bootstrap methods from *python12.py* can be used in this problem. The result is saved in *output2.txt*