随机模拟方法与应用导论作业五

陈稼霖 45875852

2019-10-03

5.5 (Exploring percentages of full-time faculty)

The variable `Full.time` in the college dataset (see Example 5.3) contains the percentage of faculty who are hired full-time in the group of National Universities.

a. Using the `hist` function, construct a histogram of the full-time percentages and comment on the shape of the distribution.

b. Use the froot and flog transformations to reexpress the full-time percentages. Construct histograms of the collection of froots and the collection of flogs. Is either transformation successful in making the full-time percentages approximately symmetric?

c. For data that is approximately normally distributed, about 68% of the data fall within one standard deviation of the mean. Assuming you have found a transformation in part (b) that makes the full-time percentages approximately normal, find an interval that contains roughly 68% of the data on the new scale.

a. 首先读取文件`college.txt`

```r
dat = read.table('college.txt',head = TRUE,sep = '\t')
```
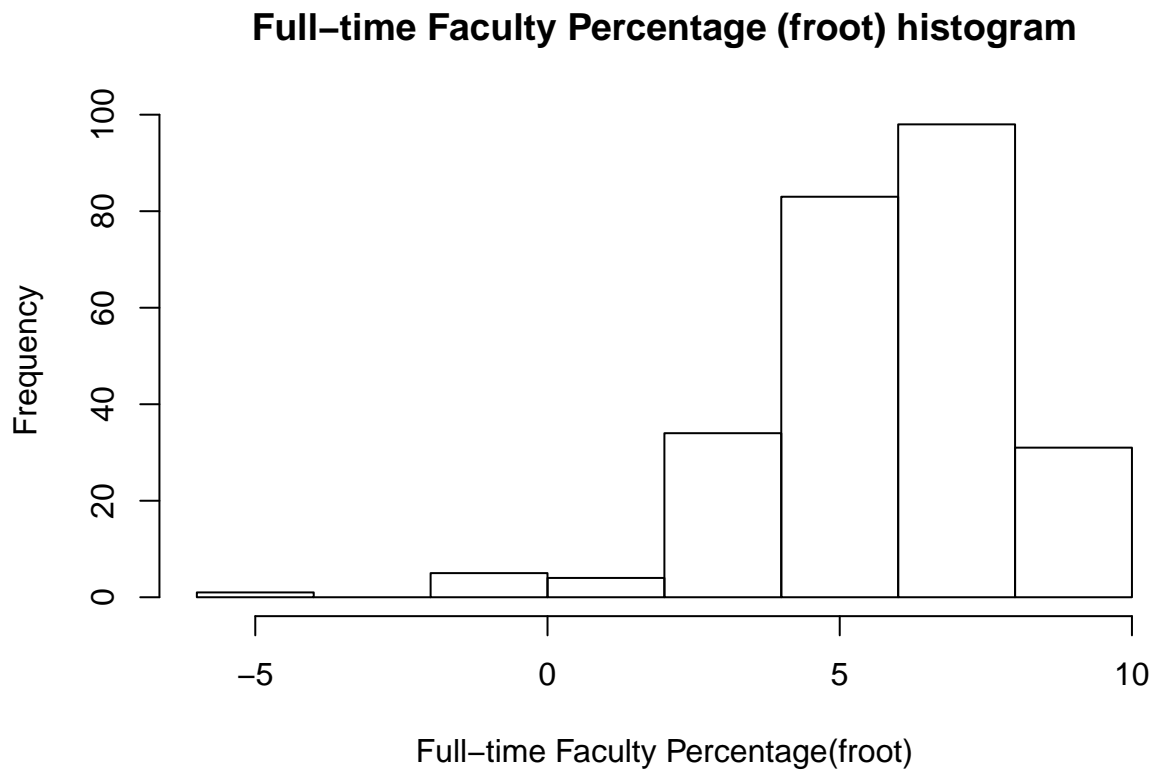
然后依据变量`Full.time`绘制全职教职工百分比的直方图

```r
hist(dat$Full.time,xlab = 'Full-time Faculty Percentage',
    main = 'Full-time Faculty Percentage')
```

Full–time Faculty Percentage
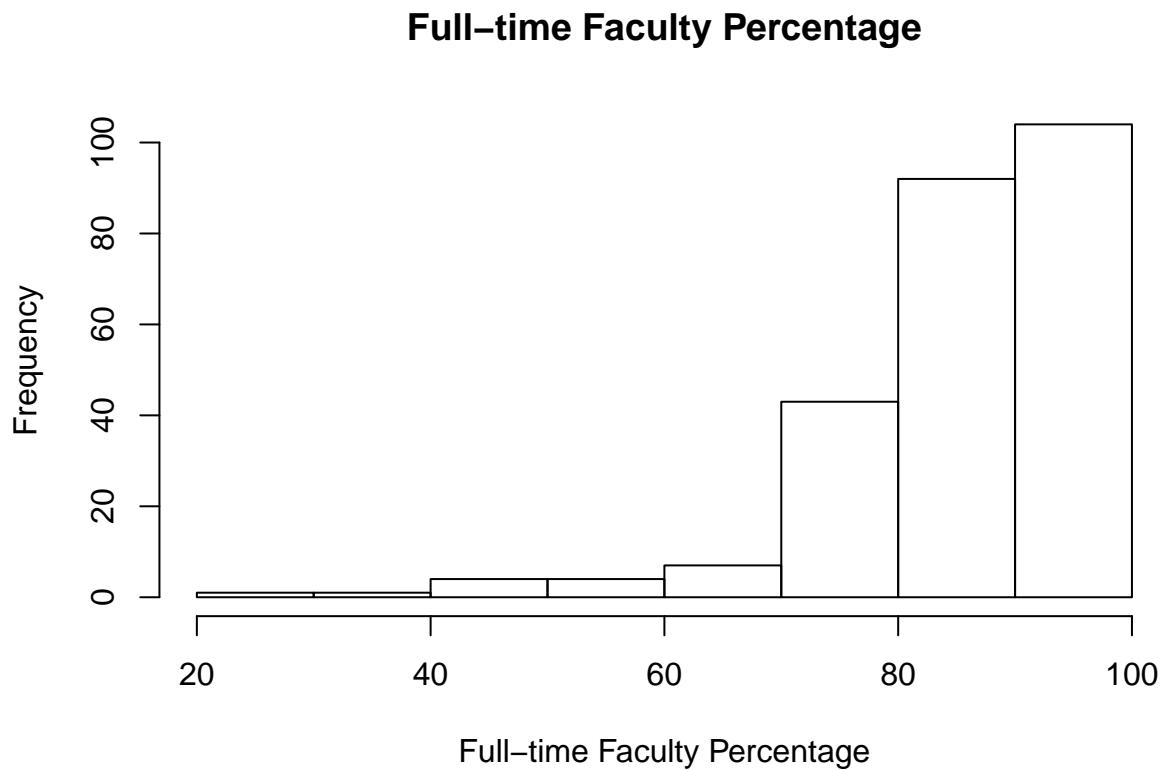
评价：由图可见，绝大多数大学的全职教职工百分比在80%以上，直方图形状右偏(right-skewed)。

    b. 用froot转换重新表示全职教职工百分比并绘制其直方图

```
froot = sqrt(dat$Full.time) - sqrt(100 - dat$Full.time)
hist(froot,xlab = 'Full-time Faculty Percentage(froot)',
     main = 'Full-time Faculty Percentage (froot) histogram')
```

Full−time Faculty Percentage (froot) histogram

Full−time Faculty Percentage(froot)

用flog转换重新表示全职教职工百分比并绘制其分布直方图

```r
flog = log(dat$Full.time) - log(100 - dat$Full.time)
hist(flog,xlab = 'Full-time Faculty Percentage (flog)',
     main = 'Full-time Faculty Percentage (flog) histogram')
```

Full−time Faculty Percentage (flog) histogram

Full−time Faculty Percentage (flog)
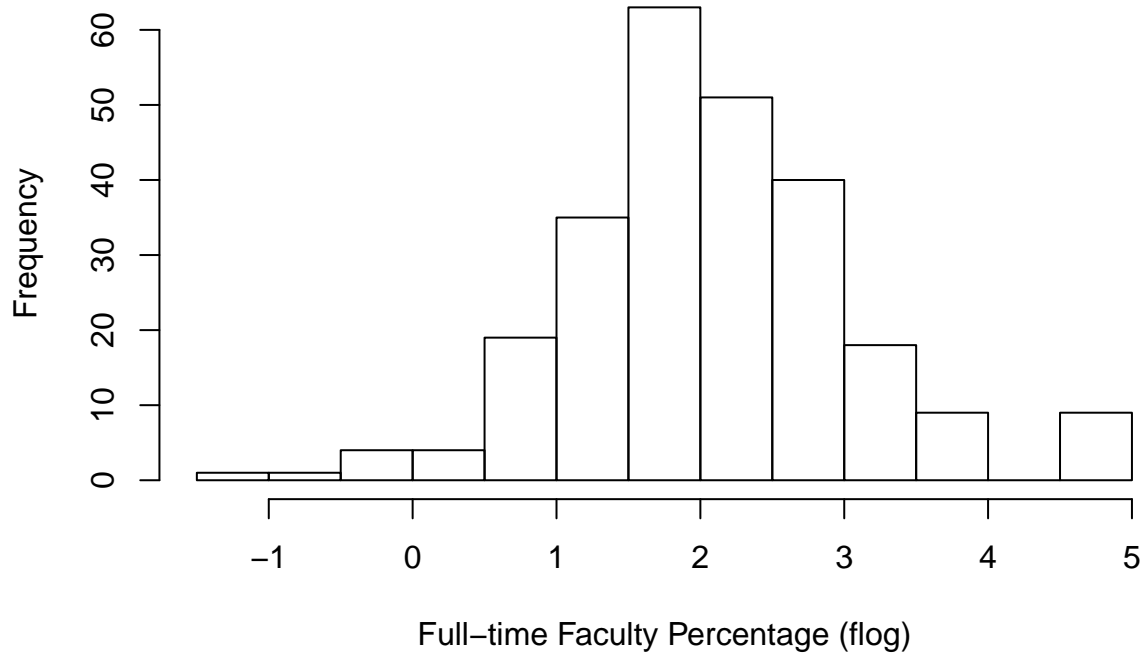
两种转换都使得全职教职工百分比的直方图更加对称。

c. 对于froot转换后得到的直方图，约68%的数据集中的区间为$(4, 8]$；对于flog转换后得到的直方图，约68%的数据集中的区间为$(1, 3]$。