ECS 132 Term Project

Steven Alvarado, Russell Chien, and Ruth Hailu

University of California, Davis

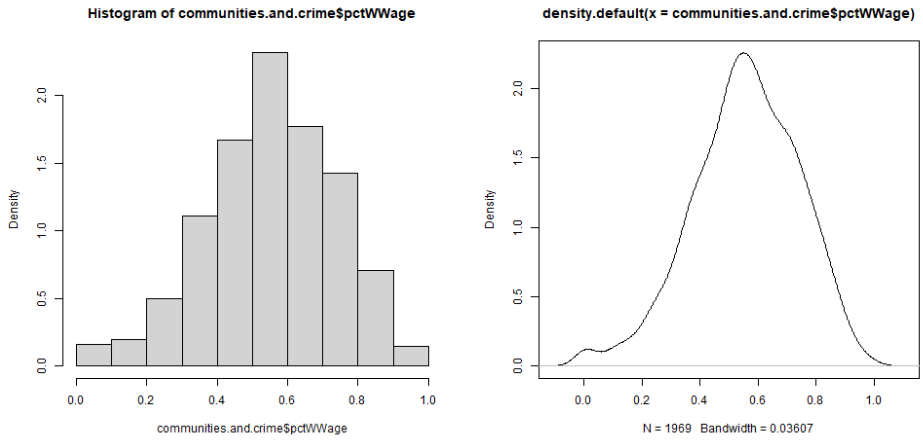
June 2023

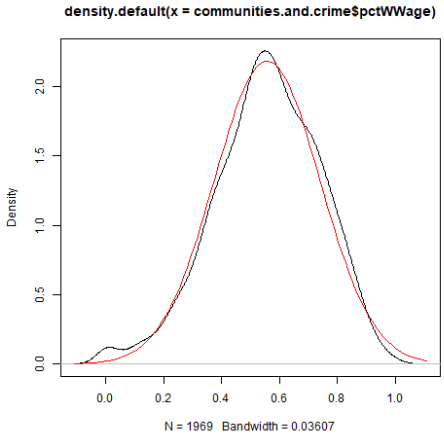
Chapter 1

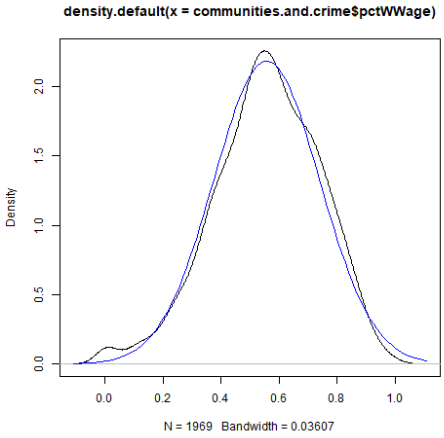
The Normal Family

1.1 Communities and Crime: pctWWage

Our group observed that the variable **pctWWage** of the Communities and Crime dataset seemed well-approximated by the normal family of continuous distributions. According to the UCI Machine Learning Repository, **pctWWage** is described as the percentage of households within the United States with wage or salary income in 1989.




density.default(x = communities.and.crime$pctWWage)
Density
0.0
0.5
1.0
1.5
2.0
0.0
0.2
0.4
0.6
0.8
1.0
N = 1969 Bandwidth = 0.03607

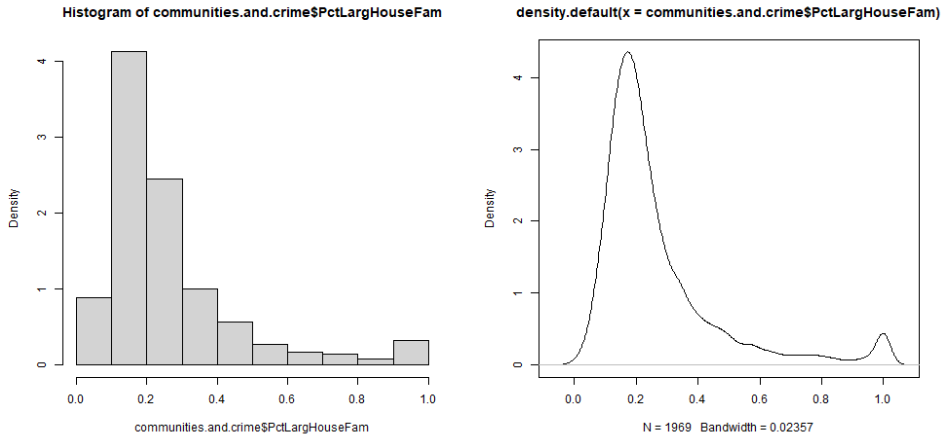

density.default(x = communities.and.crime$pctWWage)
Density
0.0
0.5
1.0
1.5
2.0
0.0
0.2
0.4
0.6
0.8
1.0
N = 1969 Bandwidth = 0.03607

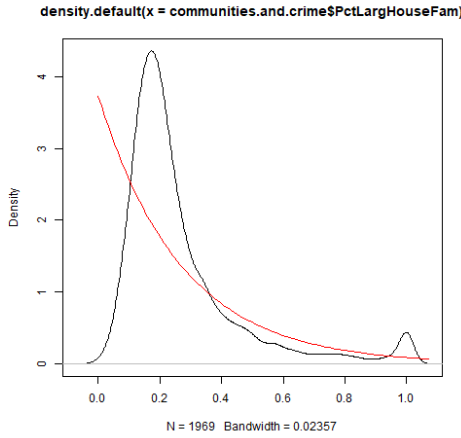
Chapter 2

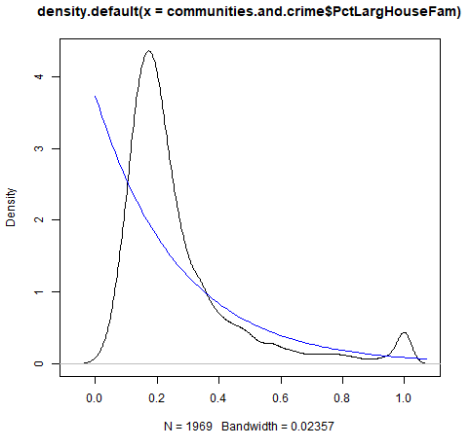
The Exponential Family

2.1 Communities and Crime: PctLargHouseFam

For the exponential family of continuous distributions, we observed that the variable **PctLargHouseFam** was a suitable approximation. According to the UCI Machine Learning Repository, **PctLargHouseFam** is described as the percentage of family households with six or more family members.




density.default(x = communities.and.crime$PctLargHouseFam)
Density
0
1
2
3
4
0.0
0.2
0.4
0.6
0.8
1.0
N = 1969 Bandwidth = 0.02357

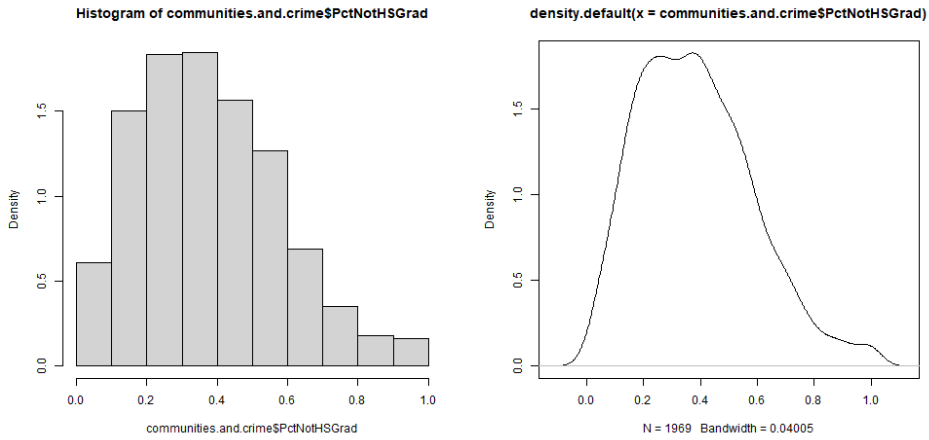

density.default(x = communities.and.crime$PctLargHouseFam)
Density
0
1
2
3
4
0.0
0.2
0.4
0.6
0.8
1.0
N = 1969 Bandwidth = 0.02357

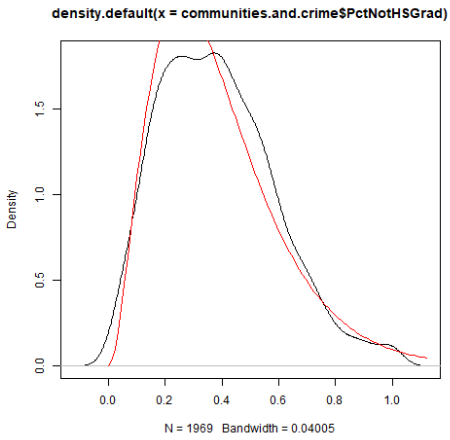
Chapter 3

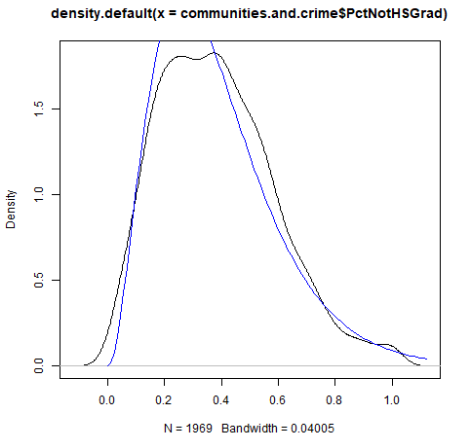
The Gamma Family

3.1 Communities and Crime: PctNotHsGrad

We observed that the variable **PctNotHsGrad** of the Communities and Crime dataset seemed well-approximated by the gamma family of continuous distributions. According to the UCI Machine Learning Repository, **PctNotHsGrad** is described as the percentage of people 25 and over that are not high school graduates.




density.default(x = communities.and.crime$PctNotHSGrad)
Density
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0
N = 1969 Bandwidth = 0.04005

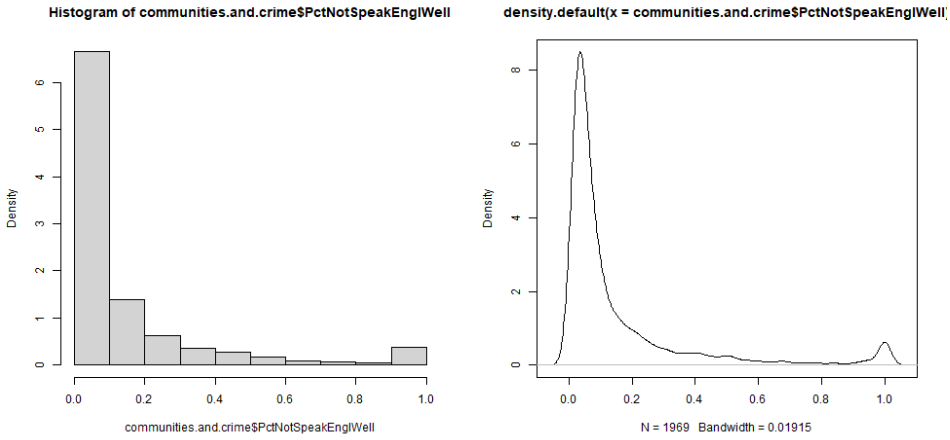

density.default(x = communities.and.crime$PctNotHSGrad)
Density
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0
N = 1969 Bandwidth = 0.04005

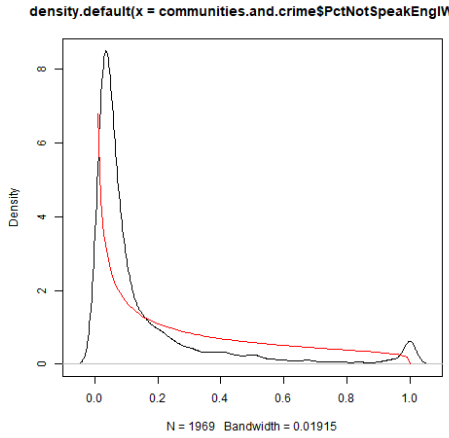
Chapter 4

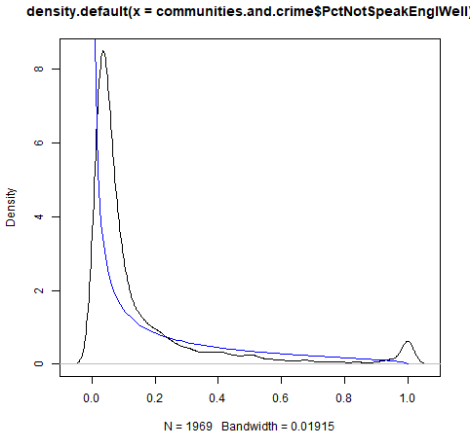
The Beta Family

4.1 Communities and Crime: PctNotSpeakEnglWell

For the beta family of continuous distributions, we observed that the variable **PctNotSpeakEnglWell** was a suitable approximation. According to the UCI Machine Learning Repository, **PctNotSpeakEnglWell** is described as the percentage of people who do not speak English well.




density.default(x = communities.and.crime$PctNotSpeakEnglWell
Density
0
2
4
6
8
0.0
0.2
0.4
0.6
0.8
1.0
N = 1969 Bandwidth = 0.01915


density.default(x = communities.and.crime$PctNotSpeakEnglWell
Density
0
2
4
6
8
0.0
0.2
0.4
0.6
0.8
1.0
N = 1969 Bandwidth = 0.01915