

Homework 03  
Joseph Blubaugh  
[jblubau1@tamu.edu](mailto:jblubau1@tamu.edu)  
STAT 641-720

1. a.

$$\begin{aligned} S(y) &= P[Y > y] = 1 - F(y) \\ &= 1 - 1 - \exp\left[-\left(\frac{y-\theta}{\alpha}\right)^\gamma\right] \\ &= -\exp\left[-\left(\frac{y-\theta}{\alpha}\right)^\gamma\right] \end{aligned}$$

b.

$$\begin{aligned} h(y) &= \frac{f(y)}{S(y)} = \frac{\frac{\gamma}{\alpha^\gamma}(y-\theta)^{\gamma-1}\exp\left[-\left(\frac{y-\theta}{\alpha}\right)^\gamma\right]}{-\exp\left[-\left(\frac{y-\theta}{\alpha}\right)^\gamma\right]} \\ &= -\frac{\gamma}{\alpha^\gamma}(y-\theta)^{\gamma-1} \end{aligned}$$

2. a.  $Q(.25) = Y_{((51-1).25+1)} = Y_{(13.5)} = 2.61$

b.  $Q(.5) = Y_{((51-1).5+1)} = Y_{26} = 5$

c.  $Q(.75) = Y_{((51-1).75+1)} = Y_{38.5} = 10.445$

3. a) Combining the Kernel Density function with the pdf, our function to estimate  $f(3)$  and  $f(6)$  are:

$$f(y) = \frac{1}{51(3)} \sum_{i=1}^{51} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(y-Y_i)^2}{2(3)^2}\right]$$

[1] "f(3) = 0.0718"

[1] "f(16) = 0.0225"

b) With a bin width of 5, minimum value of .42

bin1 = [.42, 5.42)

bin2 = [5.42, 10.42)

bin3 = [10.42, 15.42)

bin4 = [15.42, 20.42)

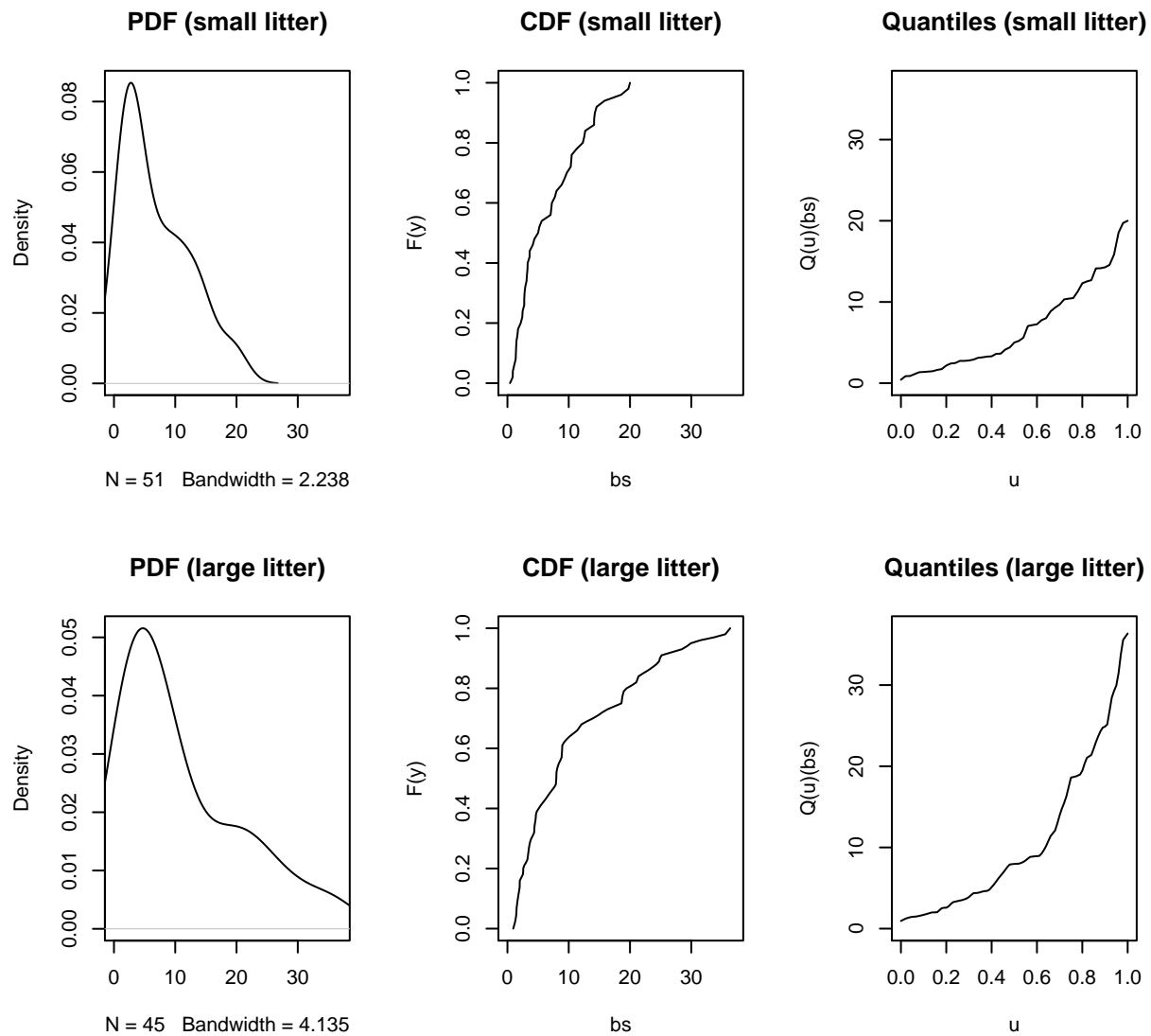
[1] "f(3) = 0.0824"

[1] "f(16) = 0.0157"

c) The smallest contribution at  $y=16$ ,  $f(16)$  is the smallest value from 16 which is .42

d) The largest contribution at  $y=16$ ,  $f(16)$  is the closest value to 16 which is 15.84

4. a)



b) Both distributions are right skewed, but the small litter distribution has more area near the peak of the distribution while the large litter has less area under the peak and tails off to the right.

c) Based on the two distributions my conclusion is that larger litters tend to have larger brain sizes but there is a lot more variation in brain size for large litters than there are for small litters.

5.

- 1) E, each function can be found from any of the other functions assuming they exist
- 2) B, observations far from the center of the distribution will have large bin widths and will not be representative of the concentration of data in the bin
- 3) D, this will take into account the concentration of points in the bin
- 4) B, kernel density is better than relative frequency because relative frequency only considers the observations in specific bins

- 5) B, mean and variance alone will not tell you the shape of the graph
- 6) D, skewness measures symmetry and kurtosis measures concentration in the tails or lack thereof which would result in more area in the peak of the distribution
- 7) E, both C and D are correct
- 8) B,  $(\mu, \sigma)$  are better estimators when the distribution is pretty close to normal
- 9) E, MAD is preferred to SIQR for very skewed distributions or heavy tails
- 10) A, since the expected value of  $E(t) = 0$ , the mean and variance are equal to the expected values of  $\mu$  and  $\sigma^2$