**STATISTICS 641 - ASSIGNMENT 3**

**DUE DATE: NOON (CDT), MONDAY, SEPTEMBER 27, 2021**

Name Jack Rodan

Email Address jrodan1@tamu.edu

**Please TYPE your name and email address. Often we have difficulty in reading the handwritten names and email addresses. Make this cover sheet the first page of your Solutions.**

STAT 641 Assign #3: H.O. 4 & 5 / Chps 1 & 4 in book.

1.) Let $Y$ have a double exponential distribution, that is $Y$ has a pdf & cdf in the following form w/ parameters $\theta, \beta > 0$.

$$f_Y(y) = (1/2\beta)\, e^{-\left(\left|\frac{y-\theta}{\beta}\right|\right)} \quad ; -\infty < y < \infty$$

$$F_Y(y) = \begin{cases} \frac{1}{2} e^{-\left(\frac{\theta - y}{\beta}\right)} & ; y < \theta \\ 1 - \frac{1}{2} e^{-(y-\theta/\beta)} & ; y \geq \theta \end{cases}$$

(a) Derive the quantile function for $Y$.

$$Q(y) = F^{-1}(y) \Rightarrow \text{for } y<\theta: \; y = \frac{1}{2} e^{-\left(\frac{\theta - Q(y)}{\beta}\right)} \iff 2y = e^{-\left(\frac{\theta - Q(y)}{\beta}\right)}$$

$$\iff \ln(2y) = \frac{-\theta + Q(y)}{\beta} \iff Q(y) = \beta \ln(2y) + \theta \; ; y < \theta$$

$$\text{for } y \geq \theta: \quad y = 1 - \tfrac{1}{2} e^{-(Q(y) - \theta)/\beta} \iff 2 - 2y = e^{-(Q(y) - \theta)/\beta}$$

$$\iff \beta \ln(2 - 2y) = -Q(y) + \theta \iff \theta - \beta \ln(2 - 2y) = Q(y); \; y \geq \theta$$

$$\boxed{Q(y) = \begin{cases} \beta \ln(2y) + \theta & ; y < \theta \\ \theta - \beta \ln(2 - 2y) & ; y \geq \theta \end{cases}}$$

(b) Derive the survival function for $Y$ (Pg 46 H.O. 3)*

$$S(y) = 1 - F(y)$$

$$\boxed{S(y) = \begin{cases} 1 - \frac{1}{2} e^{-(\theta - y)/\beta} & ; y < \theta \\ \frac{1}{2} e^{-(y-\theta)/\beta} & ; y \geq \theta \end{cases}}$$

(c) Derive the Hazard function for $Y$.

$$h(y) = \frac{f(t)}{S(t)}$$

$$\text{for } y<\theta: \quad h(y) = \frac{\frac{1}{2\beta} e^{-\left(\left|\frac{y-\theta}{\beta}\right|\right)}}{1 - \frac{1}{2} e^{-(\theta - y)/\beta}} = \boxed{\frac{e^{-\left(\left|\frac{y-\theta}{\beta}\right|\right)}}{\beta\left(2 - e^{-(\theta - y)/\beta}\right)}}$$

$$\text{for } y \geq \theta: \quad h(y) = \frac{(1/2\beta)\, e^{-\left(\left|\frac{y-\theta}{\beta}\right|\right)}}{\frac{1}{2} e^{-(y-\theta)/\beta}} = \boxed{\frac{1}{\beta}}$$

$$\boxed{h(y) = \begin{cases} \dfrac{e^{\left(\frac{y-\theta}{\beta}\right)}}{\beta\left(2 - e^{-(\theta - y)/\beta}\right)} & ; y < \theta \\ 1/\beta & ; y \geq \theta \end{cases}}$$

2.) (Done in R.) Calculate the estimates of the Quartiles of
$Q(0.25), Q(0.5), Q(0.75)$ for just the large Litter size.

** In R **: quantile(LLitterSize, na.rm = True).

$Q(0.25) = 3.3525$
$Q(0.50) = 7.9300$
$Q(0.75) = 16.6525$

3) Using the data frame from (2) for just the Large Litter Size, we want to estimate the pdf $f(y)$ for the relative brain weights of the 44 species of mammal.
The kernal density estimate is given by:

$$\hat{f}(y) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{y - y_i}{h}\right).$$

& we use the gaussian kernel and a bandwidth of $h = 3$.

(a) Estimate $f(3)$ & $f(16)$ using the kernel density estimator.
** computed in R **

$$f(3) = \frac{1}{44(3)} \sum_{i=1}^{44} \frac{1}{\sqrt{2\pi}} e^{-0.5\left(\frac{3 - y_i}{3}\right)^2} \approx 0.05977$$

$$f(16) = \frac{1}{44(3)} \sum_{i=1}^{44} \frac{1}{\sqrt{2\pi}} e^{-0.5\left(\frac{16 - y_i}{3}\right)^2} \approx 0.0167$$

(b)* Done in R.*
$f(3) \approx 0.06$ ; $f(16) \approx 0.02$

(c) & (d): *Done in R*
(c) 35.45 (d) 16

4) + Done in R +
(b) For large: Range: [0.94, 35.45]; Location: [$\mu$ = 10.392, median = 7.93]
Shape: distribution is bi-modal w/ the largest occurring crest $x = 4$ and the other mode occurring around $x = 20$. The dist is also skewed right.
For Small: Range [0.42, 20]; location [$\mu$ = 6.886, median = 5]
Shape: The distribution is unimodal & skewed right.
(c) Litter size seems to be positively correlated w/ Brain weight

5.) Select the letter of the best answer for each question.

(1) E (All the functions can be derived given any one of the options).

(2) D (see H.O.4 pg 22)

(3) A

(4) D (H.O.4 pg 27)

(5) B (H.O.4 pg 47)

(6) B (see wiki page for Kernel Density Estimation [Bandwidth Selection])

(7) D

(8) D (H.O.5 pg 14)

(9) C (H.O.5 pg 22 (bottom))

(10) D (H.O.5 pg 27 (top) + Log normal not symmetric)

(11) E (H.O.5 pg 14)

(12) E (H.O.5 pg 27 ③)

(13) B (H.O.5 pg 32)