1. Let $X_1, X_2, \dots, X_n \sim f(x_i|\theta)$ where $f(x|\theta) = \frac{1}{2i\theta} I_{\{-i(1-\theta), i(\theta+1)\}}(x_i)$. Before we derive the distribution of the sample, we note that

$$\{w_i : -i(1-\theta) < w_i < i(\theta+1)\} = \{w_i/i : \theta - 1 < w_i/i < \theta + 1\}$$

From this we see that

$$I_{\{-i(\theta-1), i(\theta+1)\}}(w_i) = I_{\{\theta-1), \theta+1\}}(w_i/i)$$

Using this, we can write the joint distribution of the sample as below.

$$\begin{aligned}
f(\underline{x}|\theta) &= \prod_{i=1}^n \frac{1}{2i\theta} I_{\{-i(1-\theta), i(\theta+1)\}}(x_i) \\
&= \frac{1}{n!(2\theta)^n} \prod_{i=1}^n I_{\{\theta-1, \theta+1\}}(x_i/i) \\
&= \frac{1}{n!(2\theta)^n} \prod_{i=1}^n I_{\{\theta-1, \max(x_i/i)\}}(x_i/i) \prod_{i=1}^n I_{\{\min(x_i/i), \theta+1\}}(x_i/i) \\
&= \frac{1}{n!(2\theta)^n} I_{\{\theta-1, \max(x_i/i)\}}(\min(x_i/i)) I_{\{\min(x_i/i), \theta+1\}}(\max(x_i/i))
\end{aligned}$$

Now, by letting $Y_i = X_i/i$ we see

$$f(\underline{x}|\theta) = \frac{1}{n!(2\theta)^n} I_{\{\theta-1, Y_{(n)}\}}, (Y_{(1)}) I_{\{Y_{(1)}), \theta+1\}}(Y_{(n)})$$

Therefore, if we let $g(T(X), \theta) = f(\underline{x}|\theta) = \frac{1}{n!(2\theta)^n} I_{\{\theta-1, Y_{(n)}\}}, (Y_{(1)}) I_{\{Y_{(1)}), \theta+1\}}(Y_{(n)})$ and $h(x) = 1$, we see that

$$T(\underline{x}) = (Y_{(1)}, Y_{(n)}) = (\min x_i/i, \max x_i/i)$$

is a sufficent statistic for θ.

2. (a) Let $\underline{X}$ and $\underline{Y}$ be samples from $f(z|\theta) = e^{-(z-\theta)} I_{(\theta,\infty)}(z)$. The distribution of $\underline{X}$ is given by

$$\begin{aligned}
f(\underline{x}|\theta) &= \prod_{i=1}^n e^{-(x_i-\theta)} I_{(\theta,\infty)}(x_i) \\
&= \exp\{-\sum_{i=1}^n (x_i - \theta)\} \prod_{i=1}^n I_{(\theta,\infty)}(x_i) \\
&= \exp\{n\theta - \sum_{i=1}^n x_i\} I_{(\theta,\infty)}(x_{(1)}) \\
&= \exp\{n\theta\} \exp\{-\sum_{i=1}^n x_i\} I_{(\theta,\infty)}(x_{(1)})
\end{aligned}$$

Using this, we see that the ratio of the distribution of $\underline{X}$ and $\underline{Y}$ is given by

$$\begin{aligned}\frac{f(\underline{x}|\theta)}{f(\underline{y}|\theta)} &= \frac{\exp\{n\theta\}\exp\{-\sum_{i=1}^n x_i\}I_{(\theta,\infty)}(x_{(1)})}{\exp\{n\theta\}\exp\{-\sum_{i=1}^n y_i\}I_{(\theta,\infty)}(y_{(1)})}\\ &= \exp\{\sum_{i=1}^n (y_i - x_i)\}\frac{I_{(\theta,\infty)}(x_{(1)})}{I_{(\theta,\infty)}(y_{(1)})}\end{aligned}$$

Here we see that this ratio is free from θ if and only if $x_{(1)} = y_{(1)}$. Therefore, we see that $X_{(1)}$, the first order statistic is a minimal sufficent statistic for θ.

(b) Following the same procedure as above, let $\underline{X}$ and $\underline{Y}$ be samples from

$$f(z|\theta) = \frac{\exp\{-(z-\theta)\}}{(1+\exp\{-(z-\theta)\})^2}$$

Finding the density of $\underline{X}$ we have

$$\begin{aligned}f(\underline{x}|\theta) &= \prod_{i=1}^n \frac{\exp\{-(x_i-\theta)\}}{(1+\exp\{-(x_i-\theta)\})^2}\\ &= \frac{\exp\{n\theta\}\exp\{-\sum_{i=1}^n x_i\}}{\prod_{i=1}^n(1+\exp\{-(x_i-\theta)\})^2}\end{aligned}$$

Now considering the ratio of the two distributions we see

$$\begin{aligned}\frac{f(\underline{x}|\theta)}{f(\underline{y}|\theta)} &= \frac{\exp\{n\theta\}\exp\{-\sum_{i=1}^n x_i\}}{\prod_{i=1}^n(1+\exp\{-(x_i-\theta)\})^2}\cdot\frac{\prod_{i=1}^n(1+\exp\{-(y_i-\theta)\})^2}{\exp\{n\theta\}\exp\{-\sum_{i=1}^n y_i\}}\\ &= \exp\{\sum_{i=1}^n(y_i-x_i)\}\left(\frac{\prod_{i=1}^n(1+\exp\{-(y_i-\theta)\})}{\prod_{i=1}^n(1+\exp\{-(x_i-\theta)\})}\right)^2\end{aligned}$$

In order this expression to be free from θ, we require that

$$\prod_{i=1}^n(1+\exp\{-(x_i-\theta)\})^2 = \prod_{i=1}^n(1+\exp\{-(y_i-\theta)\})^2$$

The only way this can occur is if $\underline{X} = \underline{Y}$ *up to permuation.* Thus the order statistics of $\underline{X}$ or simply the sample $\underline{X}$ serves as a sufficent statistic for θ. Here we see that *no data reduction occurs.*

3. Suppose $X_1, X_2 \sim f(x|\alpha) = \alpha x^{\alpha-1}e^{-x^\alpha}I_{(0,\infty)}(x)$ and $\alpha > 0$. Then we see that

$$\log(X_1) \sim g(y|\alpha) = \alpha(e^y)^{(\alpha-1)}e^{-(e^y)^\alpha}e^y = \alpha\exp\left\{y\alpha - e^{\alpha y}\right\}$$

Now, let $\psi(t) = \exp\{t - e^t\}$. Then $g(y|\alpha) = \frac{1}{1/\alpha}\psi(\frac{1}{1/\alpha}y)$. From this, we see that $g(y|\alpha)$ is a scale familiy with scale parameter $1/\alpha$. Therefore, we know there exists

$Y_1 = \frac{1}{\alpha}\log(X_1)$ and $Y_2 = \frac{1}{\alpha}\log(X_2)$ where $Y_i \sim \psi(t)$ *which is free from* α. This gives that

$$\frac{\log(X_1)}{\log(X_2)} = \frac{1/\alpha Y_1}{1/\alpha Y_2} = \frac{Y_1}{Y_2}$$

Recall that X_1 and X_2 were indepenent, so Y_1 and Y_2 are independent. Therefore, ther joing density $f(Y_1, 1/Y_2) = f(Y_1)f(1/Y_2)$. We know $Y_1 \sim \psi(t_1)$ and by the continuous mapping theorem $Y_2 \sim \frac{1}{\psi(t_2)}$. Thus

$$\frac{Y_1}{Y_2} \sim \frac{\psi(t_1)}{\psi(t_2)}$$

which is free from α. From here we see see the distribution function of $\frac{\log(X_1)}{\log(X_2)}$, $F(x)$ is given by

$$F(x) = \int_0^x \frac{\psi(t_1)}{\psi(t_2)}dt$$

The integrand is free from α and the limits to do not depend on α so $F(x)$ is free from θ. Therefore, we see that $\frac{\log(X_1)}{\log(X_2)}$ is an ancillary statistic.

4. If $X_1, \ldots, X_n \sim F(x-\theta)$ are from a location familiy, then there exists $Z_1, \ldots, Z_n \sim F(x)$ such that $X_i = Z_i + \theta$. From here we have

$M_X = median(X_1, \ldots, X_n) = median(Z_1+\theta, \ldots, X_n+\theta) = \theta + median(Z_1, \ldots, Z_n) = \theta + M_Z$

$$\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{n}\sum_{i=1}^n (Z_i + \theta) = \frac{1}{n}\sum_{i=1}^n Z_i + \frac{1}{n}\sum_{i=1}^n \theta = \overline{Z} + \theta$$

Therefore the statistic $M_X - \overline{X} = M_Z - \overline{Z}$. Now M_Z and $\overline{Z}$ are combinations of random variables from the same distribution $F(x)$ which is *free from* θ. Therefore, the distribution of $M_Z - \overline{Z}$ will also be free from θ. Hence, $M_X - \overline{X}$ is an ancillary statistic.

5. (a) Since, x is bounded above by θ and $f(x|\theta)$ is monotone with respect to θ, we conjecture that $X_{(n)}$ is a complete sufficent statistic for θ. We first check sufficency. Let $X_1, \ldots, X_n \sim f(x|\theta)$. Then the joint distribution of these random variables is given by

$$\begin{aligned} f(\underline{x}|\theta) &= \prod_{i=1}^n \frac{2x_i}{\theta^2} I_{(0,\theta)}(x_i) \\ &= \left(\frac{2}{\theta^2}\right)^n \prod_{i=1}^n x_i \prod_{i=1}^n I_{(0,\theta)}(x_i) \\ &= \left(\frac{2}{\theta^2}\right)^n I_{(0,\theta)}(x_{(n)}) \prod_{i=1}^n x_i \end{aligned}$$

Letting $g(T(x),\theta) = \left(\frac{2}{\theta^2}\right)^n I_{(0,\theta)}(x_{(n)})$ and $h(\underline{x}) = \prod_{i=1}^n x_i$ we see that $T(X) = x_{(n)}$ is a sufficent statistic for θ. For completeness, consider an aribtrary measureabl function $g(\cdot)$. We will analyze $\mathbb{E}_\theta(g(X_{(n)}))$ but will need the density of our statistic. Recall for the nth order statistic $f_{X_n}(y) = n[F_X(y)]^{(n-1)}f_X(y)$. Then

$$F_X(y) = \int_0^y \frac{2t}{\theta^2}dt = \frac{y^2}{\theta^2}$$

So we have

$$f_{X_{(n)}} = n\left[\frac{y^2}{\theta^2}\right]^{(n-1)}\frac{2y}{\theta^2} = \frac{2n}{\theta^{2n}}y^{2n-1}$$

Now for any $g(\cdot)$ measurable, we have

$$\begin{aligned}
\mathbb{E}_\theta(g(X_{(n)})) &= \int_0^\theta g(y)f_{X_{(n)}}(y)dy = 0\\
&= \int_0^\theta g(y)\frac{2n}{\theta^{2n}}y^{2n-1}dy\\
&\overset{Leibnitz}{=} g(\theta)\frac{2n}{\theta^{2n}}\theta^{2n-1} + \int_0^\theta \frac{\partial}{\partial\theta}g(y)\frac{2n}{\theta^{2n}}y^{2n-1}dy\\
&= g(\theta)\frac{2n}{\theta} + \int_0^\theta g(y)2ny^{2n-1}(\frac{-2n}{\theta^{2n+1}})dy\\
&= g(\theta)\frac{2n}{\theta} + \frac{-2n}{\theta}\int_0^\theta g(y)\frac{2n}{\theta^{2n}}y^{2n-1}dy\\
&= g(\theta)\frac{2n}{\theta} + \frac{-2n}{\theta}\mathbb{E}_\theta(g(X_{(n)}))\\
&= g(\theta)\frac{2n}{\theta}
\end{aligned}$$

From here we see that $g(\theta)\frac{2n}{\theta} = 0$ which implies $g(\theta) = 0$ for $\theta > 0$. Since $0 < x < 1$ we see that $g(x) = 0$ for all x. Therefore, $X_{(n)}$ is a complete sufficent statistic.

(b) Let $X_1, \dots, X_n \sim f(x|\theta) = \frac{\theta}{(1+x)^{(1+\theta)}}I_{(0,\infty)}(x)$. First notice that θ is one dimensional and

$$f(x|\theta) = \theta\exp\{-(1+\theta)\log(1+x)\}$$

is a *full* exponential family. Therefore, $T(\underline{X}) = \sum_{i=1}^n \log(1+x_i)$ is a sufficent statistic for θ. Moreover, since $\{(\theta-1) : \theta \in \mathbb{R}\} = \mathbb{R}$ is an open set we see that by $\sum_{i=1}^n \log(1+x_i)$ is a complete sufficent statistic for θ.

(c) Let $X_1, \dots, X_n \sim f(x|\theta) = \frac{(\log(\theta)\theta^x)}{\theta-1}I_{(0,1)}(x)$ for $\theta > 1$. Again notice that θ is one dimensional and

$$\begin{aligned}
f(x|\theta) &= \frac{(\log(\theta))\theta^x}{\theta - 1} \\
&= \exp\{\log(\log(\theta)) + x\log(\theta) - \log(\theta - 1)\} \\
&= \frac{\log(\theta)}{\theta - 1}\exp\{x\log(\theta)\}
\end{aligned}$$

is an exponential family. Therefore, $T(\underline{X}) = \sum_{i=1}^n x_i$ is a sufficent statistic for θ. Moreover, $\{\log(\theta) : \theta > 1\} = \mathbb{R}$ is an open set. Therefore, $\sum_{i=1}^n x_i$ is a complete sufficent statistic for θ.

6. (a) Let X be an observation from $f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|}I_{\{-1,0,1\}}(x)$ for $0 \leq \theta \leq 1$. If X were a complete sufficent statistic, then $\mathbb{E}_\theta(g(X)) = 0$ would imply $g(X) = 0$. Now, since the support of X is only three points, we have

$$\mathbb{E}_\theta(g(X)) = g(-1)\frac{\theta}{2} + g(0)(1-\theta) + g(1)\frac{\theta}{2} = 0$$

Clearly from this we see there are measureable functions such that $g(x) \neq =$ but satisfy this equation. For instance, if $g(-1) = 1 = -g(1)$ and $g(x) = 0$ otherwise then $\mathbb{E}_\theta(g(x)) = 0$ but $g(x) \not\equiv 0$. Hence X is not a complete sufficent statistic.

(b) To see why $|X|$ is a sufficent statistic consider the following

$$\begin{aligned}
f(\underline{x}|\theta) &= \left(\frac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|} \\
&= \exp\left\{|x|\log(\theta/2) + (1-|x|)\log(1-\theta)\right\} \\
&= \exp\left\{|x|\log(\theta/2) + \log(1-\theta) - |x|\log(1-\theta)\right\} \\
&= (1-\theta)\exp\left\{|x|\log(\theta/2) - |x|\log(1-\theta)\right\} \\
&= (1-\theta)\exp\left\{|x|(\log(\theta/2) - \log(1-\theta))\right\}
\end{aligned}$$

Therefore, $f(x|\theta)$ is an exponential family and since our sample is has $n = 1$ we have $T(\underline{X}) = \sum_{i=1}^1 |x_i| = |x|$. Thus $|x|$ is a sufficent statistic.
To see why it is complete, consider the following

$$\mathbb{E}_\theta(g(|x|)) = g(1)\theta + g(0)(1-\theta) = 0$$

Taking the derivative with respect to θ yeilds.

$$g(1) - g(0) = 0$$

hence $g(1) = -g(0)$. Plugging this into our original equation we have

$$
\begin{aligned}
-g(0)\theta + g(0)(1-\theta) &= 0\\
g(0) &= 2\theta g(0)
\end{aligned}
$$

This equality holds for all θ only when $g(0) = 0$. Thus, $g(1) = -g(0) = 0$. Thus $g(|x|) \equiv 0$ for all values of θ.

(c) Yes. Recall from part b we have

$$f(x|\theta) = (1-\theta)\exp\left\{|x|(\log(\theta/2) - \log(1-\theta))\right\}$$

Letting $h(x) > 1$, $c(\theta) = (1-\theta) \geq 0$, $w(\theta) = \log(\theta/2) - \log(1-\theta)$, and $t(x) = |x|$.

7. (a) Let $\underline{X}$ be a sample from $f(x|\theta)$. Then we have

$$f(\underline{x}|\theta) = \prod_{i=1}^{n} \theta x^{\theta-1} = \theta^n \left(\prod_{i=1}^{n} x_i\right)^{\theta-1}$$

So by Neyman - Fisher Factorization $\prod_{i=1}^n x_i$ is a sufficent statistic for θ, but $\sum_{i=1}^n x_i$ is not sufficent.

(b) We already showed that $\prod_{i=1}^n x_i$ is sufficent. All we must show now is that $\prod_{i=1}^n x_i$ is complete as well. First notice that

$$
\begin{aligned}
f(x|\theta) &= \theta x^{\theta-1}\\
&= \exp\left\{\log(\theta) + (\theta-1)\log(x)\right\}\\
&= \theta\exp\left\{(\theta-1)\log(x)\right\}
\end{aligned}
$$

So we see that $f(x|\theta)$ an exponential family. Thus, $\sum_{i=1}^n \log(x_i)$ is a complete statistic for θ. But notice that $\sum_{i=1}^n \log(x_i) = \log\left(\prod_{i=1}^n x_i\right)$ and since $\log(\cdot)$ is one to one, so $\prod_{i=1}^n x_i$ is also a complete statistic. Therefore, $\prod_{i=1}^n x_i$ is a complete and sufficent statistic.

8. (a) First we will show that $X_{(1)}$ is a sufficent statistic using Neyman-Fisher's factorization theorem.

$$
\begin{aligned}
f(\underline{x}|\mu) &= \prod_{i=1}^{n} e^{-(x_i-\mu)} I_{(\mu,\infty)}(x_i)\\
&= \exp\Big\{-\sum_{i=1}^{n}(x_i-\mu)\Big\}\prod_{i=1}^{n} I_{(\mu,\infty)}(x_i)\\
&= \exp\Big\{-\sum_{i=1}^{n} x_i + n\mu\Big\} I_{(\mu,\infty)}(x_{(1)})\\
&= \frac{\exp\{n\mu\}}{\exp\{\sum_{i=1}^n x_i\}} I_{(\mu,\infty)}(x_{(1)})
\end{aligned}
$$

Letting $h(\underline{x}) = \exp\{-\sum_{i=1}^n x_i\}$ and $g(T(x),\theta) = \exp n\mu I_{\mu,\infty}(X_{(1)})$. Therefore, we see that $X_{(1)}$ is a sufficent statistic for μ.

To see why it is complete, we will need to first find the density of $X_{(1)}$. First recall that for the first order statistic we have $f_{X_{(1)}}(y) = n\big[1 - F_X(y)\big]^{n-1} f_X(y)$. Here

$$F_X(y) = \int_\mu^y e^{-(t-\mu)}dt = 1 - e^{-(y-\mu)}$$

Then we see that

$$F_{X_{(1)}}(y) = n\big[1 - 1 + e^{-(y-\mu)}\big]^{n-1}e^{-(y-\mu)} = ne^{-n(y-\mu)}$$

Now, let $g(\cdot)$ be any measurable function. We now analyze $\mathbb{E}_\mu(g(X_{(1)})$.

$$\mathbb{E}_\mu(g(X_{(1)})) = \int_\mu^\infty g(y)ne^{-n(y-\mu)}dy = 0$$

Now taking a derivate with respect to μ, we see

$$\begin{aligned}
0 &= \frac{\partial}{\partial\mu}\int_\mu^\infty g(y)ne^{-n(y-\mu)}dy \\
&= -g(\mu)ne^{-n(\mu-\mu)}\frac{d}{d\mu}\mu + \int_\mu^\infty \frac{\partial}{\partial\mu}g(y)ne^{-n(y-\mu)}dy \\
&= -g(\mu)ne^{-n(\mu-\mu)}\frac{d}{d\mu}\mu + \int_\mu^\infty \frac{\partial}{\partial\mu}g(y)ne^{-n(y-\mu)}dy \\
&= -g(\mu)n + \int_\mu^\infty g(y)n^2e^{-n(y-\mu)}dy \\
&= -ng(\mu) + n\mathbb{E}_\mu(g(X_{(1)}) \\
&= -ng(\mu)
\end{aligned}$$

Therefore, we see that $g(\mu) = 0$ for $\infty < \mu < y$. Recall that this calculation was for arbitrary x, so letting $x \to \infty$, we see that $g(x) \equiv 0$. Therfore, $X_{(1)}$ is a complete sufficent statistic for μ.

(b) To use Basu's theorem, we must first show that S^2 is an ancillary statistic. Let $\psi(t) = e^{-t}$. Then we see that $f(x|mu) = \psi(t-\mu)$. Therefore $f(x|\mu)$ is a location family. Thus, for each $X_1,\dots,X_n$ we have $X_i = Z_i + c$ for $Z_i \sim \psi$ which is *free from* μ. This will allow us to show S^2 is ancillary - but first consider the following calculation

$$\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{n}\sum_{i=1}^n Z_i + c = \overline{Z} + c$$

Hence we see that

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2 = \frac{1}{n-1}\sum_{i=1}^n (Z_i + c - \overline{Z} - c)^2 = \frac{1}{n-1}\sum_{i=1}^n (Z_i - \overline{Z})^2 = S_Z^2$$

We now see that S^2 is a combination of random variables that do not depend on μ. Therefore, we see that S^2 does not depend on μ. That is the distribution of S^2 is constant with respect to μ. Thus, S^2 is ancillary statistic for μ. Therefore, using Basu's theorem we see that $X_{(1)}$ and S^2 are independent.

9. (a) If $X_1, \dots, X_n \sim \frac{1}{a}\psi(\frac{x-b}{a})$ where $a > 0$ and $-\infty < b < \infty$ there there exists $Z_1, \dots, Z_n$ such that $X_i = aZ_i + b$ with $Z_i \sim \psi(z)$. With this and the property of the statistics, we see that

$$\begin{aligned}
\frac{T_1(X_1, \dots, X_n)}{T_2(X_1, \dots, X_n)} &= \frac{T_1(aZ_1 + b, \dots, aZ_n + b)}{T_2(aZ_1 + b, \dots, aZ_n + b)} \\
&= \frac{aT_1(Z_1, \dots, Z_n)}{aT_2(Z_1, \dots, Z_n)} \\
&= \frac{T_1(Z_1, \dots, Z_n)}{T_2(Z_1, \dots, Z_n)}
\end{aligned}$$

Now notice that since Z_i are independent of a, b, then so is $T_i(Z_1, \dots, Z_n)$. Therefore the distribution of T_1/T_2 is an ancillary statistic.

(b) If R is the sample range, than using the same notation as above, we see that

$$R_X = X_{(n)} - X_{(1)} = (aZ_{(n)} + b) - (aZ_{(1)} + b) = a(Z_{(n)} - Z_{(1)}) = aR_Z$$

Before we calculate the sample standard deviation, we have

$$\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{n}\sum_{i=1}^n (aZ_i + b) = a\overline{Z} + b$$

Then for the sample standard deviation, we have

$$\begin{aligned}
S &= \left(\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2\right)^{1/2} \\
&= \left(\frac{1}{n-1}\sum_{i=1}^n (aZ_i + b - a\overline{Z} - b)^2\right)^{1/2} \\
&= \left(\frac{1}{n-1}\sum_{i=1}^n a^2(Z_i - \overline{Z})^2\right)^{1/2} \\
&= a\left(\frac{1}{n-1}\sum_{i=1}^n (Z_i - \overline{Z})^2\right)^{1/2} \\
&= aS_Z
\end{aligned}$$

Therefore, using the result from above, we see that

$$R/S = R_Z/R_S$$

which is independent of a and b. Therefore R/S is an ancillary statistic.