(7) * We are to find the minimal sufficient statistic for the RS $X_1 \dots X_n$ for each of the distributions below:

a) $f(x|\theta) = \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2}$, $-\infty < x < \infty$, $-\infty < \theta < \infty$:

* The joint :

$$\rightarrow f(x_1 \dots x_n | \theta) = \left(\frac{1}{2\pi}\right)^{n/2} e^{-\frac{1}{2}(\sum x_i^2 - 2\theta \sum x_i + n\theta^2)}$$

* If we were then to consider dividing the above by the exact same distribution of Y_i's, it is apparent the minimal statistic must be met when :

$$\rightarrow 2\theta\left(\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i\right) \Longrightarrow 2\theta n(\bar{x} - \bar{y}) \Longrightarrow \bar{x} = \bar{y}$$, to keep θ constant

$\boxed{\therefore T(x) = \bar{x} \text{ is the minimal sufficient statistic}}$

b) $f(x|\theta) = e^{-(x-\theta)}$, $\theta < x < \infty$, $-\infty < \theta < \infty$

* The joint :

$$\rightarrow f(x_1 \dots x_n | \theta) = e^{-\sum x_i} e^{n\theta}$$

* Since our X_i's are factored from θ, this means our minimal stat comes from the constraint :

$$\Rightarrow \min(x) > \theta \Rightarrow X_{(1)}$$

$\boxed{\therefore T(x) = X_{(1)} \text{ is our minimal sufficient statistic}}$

c) $f(x|\theta) = \frac{1}{\pi[1 + (x-\theta)^2]}$; $-\infty < x < \infty$, $-\infty < \theta < \infty$

* The joint :

$$\rightarrow f(x_1, \dots x_n \mid \theta) = \frac{1}{\pi^n \prod_{i=1}^{n} [1 + (x - \theta)^2]}$$

* To find the minimal :

$$\rightarrow \frac{f(x_1, \dots x_n \mid \theta)}{f(y_1, \dots y_n \mid \theta)} = \frac{\prod_{i=1}^{n} [1 + (y_i - \theta)^2]}{\prod_{i=1}^{n} [1 + (x_i - \theta)^2]} = \prod_{i=1}^{n} \left[\frac{1 + (y_i - \theta)^2}{1 + (x_i - \theta)^2} \right]$$

* The above expression is only constant in θ if Y and X have the same order statistics :

$$\boxed{\therefore T(X) = \{X_{(1)}, X_{(2)}, \dots, X_{(n)}\} \text{ is the minimal sufficient statistic}}$$

(9) * We are told $X_1 \dots X_n$ is an RS from a location family. We are to prove that $M - \bar{X}$ is an ancillary statistic, where:

- $M = (X_{(1)} + X_{(n)})/2$, the median ; $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$, the sample mean

* We will call $Z = M - \bar{X}$ and $X_i = Y_i + \theta$:

* We then have :

$$\rightarrow Pr(Z \le z) = Pr(M - \bar{X} \le z) = Pr\left[\frac{X_{(n)} + X_{(1)}}{2} - \bar{X} \le z\right]$$

$$= Pr\left[\frac{1}{2}(\max(X_i) + \min(X_i)) - \frac{1}{n}\sum_{i=1}^{n} X_i \le z\right]$$

$$= Pr\left[\frac{1}{2}(\max(Y_i+\theta) + \min(Y_i+\theta)) - \frac{1}{n}\sum_{i=1}^{n}(Y_i+\theta) \le z\right]$$

$$= Pr\left[\frac{1}{2}(\max(Y_i)+\theta \min(Y_i)+\theta) - \frac{1}{n}\sum_{i=1}^{n} Y_i + \theta \le z\right]$$

$$= Pr\left[\frac{1}{2}(\max(Y_i) + \min(Y_i)) + \theta - \frac{1}{n}\sum_{i=1}^{n} Y_i - \theta \le z\right]$$

$$= Pr\left[\frac{1}{2}(\max(Y_i) + \min(Y_i)) - \frac{1}{n}\sum_{i=1}^{n} Y_i \le z\right]$$

* From the above, we can see that all Y_i are independent of θ under the transformation to the RV Z. Since the Y_i are just a repositioning of all X_i from the location family :

$$\boxed{\therefore Z = M - \bar{X} \text{ is an ancillary statistic of all } X_i}$$