**Minimal sufficiency.** $T = T(X)$ is minimal sufficient if it provides the maximal amount of data reduction. i.e., for any sufficient statistics $U = U(X)$, there exists a function $H$ such that $T = H(U)$.

Usually, one can find minimal sufficient statistics applying the factorization theorem and inspection. However, there are some techniques to find them also.

**Theorem.** $\mathcal{P}$ is a family of probability models with common support and $\mathcal{P}_0 \subset \mathcal{P}$. If $T$ is minimal sufficient for $\mathcal{P}_0$ and sufficient for $\mathcal{P}$, then it is minimal sufficient for $\mathcal{P}$ also.

**Proof.** Let $U$ be any sufficient statistic for $\mathcal{P}$. Then it is sufficient for $\mathcal{P}_0$. But $T$ is minimal sufficient for $\mathcal{P}_0$. Therefore $T = H(U)$. Now consider $\mathcal{P}$. $T$ is minimal sufficient for $\mathcal{P}$ and for any other sufficient statistic $U$, $T = H(U)$. Therefore, $T$ is minimal sufficient.

**Theorem.** Let $f(x|\theta)$ be pmf or pdf of $X$. Suppose there exists a function $T(x)$ such that, for two sample points $x$ and $y$ the ratio $f(x|\theta)/f(y|\theta)$ is a constant function of $\theta$ iff $T(x) = T(y)$. Then $T(X)$ is minimal sufficient for $\theta$.

**Example.** $X_1, \ldots, X_n$ i.i.d Poisson($\lambda$). Then

$$f(x_1, \ldots, x_n|\lambda) = \exp(-n\lambda) \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}.$$

Therefore,

$$\frac{f(x_1, \ldots, x_n|\lambda)}{f(y_1, \ldots, y_n|\lambda)} = \lambda^{\left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)} \frac{\prod_{i=1}^n y_i!}{\prod_{i=1}^n x_i!}$$

is a constant function of $\lambda > 0$ iff $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. Therefore, $T(X_1, \ldots, X_n) = \sum_{i=1}^n X_i$ is actually minimal sufficient for $\lambda$.

**Proof (of Theorem).** Assume $f(x|\theta) > 0$ for all $x \in \mathcal{X}$ and $\theta \in \Theta$. Suppose there exists $T$ such that $f(x|\theta)/f(y|\theta)$ is a constant function of $\theta$ iff $T(x) = T(y)$. We show then that $T(X)$ is minimal sufficient for $\theta$. First, let us show that it is sufficient. The map $T$ is
$T : \mathcal{X} \to \mathcal{T} = \{t : t = T(x) \text{ for some } x \in \mathcal{X}\}$. Let $A_t = \{x \in \mathcal{X} : T(x) = t\}$. Then $\{A_t\}_{t \in \mathcal{T}}$ is a partition of $\mathcal{X}$. For each $A_t$, fix one element $x_t \in A_t$. For any $x \in \mathcal{X}$, we have that $x \in A_{T(x)}$, and hence $x_{T(x)}$ is the fixed element which belongs to the same partitioning set as $x$ does. $T(x) = T(x_{T(x)})$ since $x$ and $x_{T(x)}$ belong to $A_{T(x)}$. Hence $f(x|\theta)/f(x_{T(x)}|\theta)$ is a constant function of $\theta$. Then $h(x) = f(x|\theta)/f(x_{T(x)}|\theta)$ is independent of $\theta$ and $h : \mathcal{X} \to \mathcal{R}^+$.

Define $g$ by $g(t,\theta) = f(x_t|\theta)$ and $g : \mathcal{T} \times \Theta \to \mathcal{R}^+$. Then

$$f(x|\theta) = \frac{f(x|\theta)}{f(x_{T(x)}|\theta)} f(x_{T(x)}|\theta) = h(x)g(T(x),\theta).$$

Therefore, $T(X)$ is sufficient for $\theta$. Let $T'(X)$ be any other sufficient statistics. Then there exist $g'$ and $h'$ such that $f(x|\theta) = g'(T'(x),\theta)h'(x)$. Let $x$ and $y$ be any two sample points such that $T'(x) = T'(y)$. Then

$$\frac{f(x|\theta)}{f(y|\theta)} = \frac{g'(T'(x),\theta)h'(x)}{g'(T'(y),\theta)h'(y)} = \frac{h'(x)}{h'(y)},$$

which is independent of $\theta$. We already have that $T(x) = T(y)$ whenever $f(x|\theta)/f(y|\theta)$ is a constant function of $\theta$. Therefore, $T'(x) = T'(y)$ implies $T(x) = T(y)$. This means that $T$ is coarser than $T'$ or $T(x) = q(T'(x))$ for some function $q$. Therefore $T$ is minimal sufficient.

**Example.** Let $X_1, \ldots, X_n$ be i.i.d $\mathrm{Exp}(\theta)$, $\theta > 0$ with density $f(x|\theta) = \theta\exp(-\theta x)$ for $x > 0$. Then

$$f_{X_1,\ldots,X_n}(x_1,\ldots,x_n|\theta) = \prod_{i=1}^{n}(\theta\exp(-\theta x_i)) = \theta^n \exp(-\theta\sum_{i=1}^{n} x_i), x_i > 0, \theta > 0.$$

$$\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} = \frac{\theta^n \exp(-\theta\sum_{i=1}^n x_i)}{\theta^n \exp(-\theta\sum_{i=1}^n y_i)} = \exp(-\theta\{\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i\})$$

is a constant function of $\theta$ in the interval $\theta > 0$ iff $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. Therefore, $\sum_{i=1}^n X_i$ is minimal sufficient for $\theta$.

**Example.** Suppose $X - \theta \sim \mathrm{Exp}(\theta)$, $-\infty < \theta < \infty$. Then

$$f(x|\theta) = \begin{cases} \exp(-(x-\theta)) & x > \theta; \\ 0 & \text{otherwise.} \end{cases}$$

Then $\Theta = \mathcal{R}$ and the common support is $\mathcal{X} = \mathcal{R}$ also. Consider a random sample, $X_1, \ldots, X_n$ from this distribution. Then

$$\begin{aligned} f(x_1,\ldots,x_n|\theta) &= \begin{cases} \exp(-(\sum_{i=1}^n x_i - n\theta)) & x_i > \theta \text{ for all } i; \\ 0 & \text{otherwise.} \end{cases} \\ &= \begin{cases} \exp(-(\sum_{i=1}^n x_i - n\theta)) & x_{(1)} > \theta\ ; \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Since $f(x_1, \ldots, x_n|\theta) = \exp(-\sum_{i=1}^n x_i) \exp(n\theta) I(x_{(1)} > \theta)$, $X_{(1)}$ is sufficient. Further,

$$
\begin{aligned}
\frac{f(\mathbf{x}|\theta)}{f(\mathbf{y}|\theta)} &= \frac{\exp(-\sum_{i=1}^n x_i) \exp(n\theta) I(x_{(1)} > \theta)}{\exp(-\sum_{i=1}^n y_i) \exp(n\theta) I(y_{(1)} > \theta)} \\
&= \begin{cases} \exp(-(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i)) & \text{if } \theta < \min\{x_{(1)}, y_{(1)}\}; \\ 0 & \text{if } x_{(1)} < \theta < y_{(1)}; \\ \infty & \text{if } y_{(1)} < \theta < x_{(1)}; \\ \text{undefined elsewhere.} & \end{cases}
\end{aligned}
$$

This is a constant function of $\theta$ iff $x_{(1)} = y_{(1)}$. Therefore, $X_{(1)}$ is minimal also.