Examples of Sufficient Statistics

Sufficient statistics in exponential families

If \( \mathcal{P} \) is an exponential family, then Theorem 2.2 can be applied with

\[
g_\theta(t) = \exp\{[\eta(\theta)]^\top t - \xi(\theta)\},
\]

i.e., \( T \) is a sufficient statistic for \( \theta \in \Theta \).

In Example 2.10 the joint distribution of \( X \) is in an exponential family with \( T(X) = \sum_{i=1}^n X_i \).

Hence, we can conclude that \( T \) is sufficient for \( \theta \in (0, 1) \) without computing the conditional distribution of \( X \) given \( T \).
Example 2.11 (Truncation families)

Let $\phi(x)$ be a positive Borel function on $(\mathcal{R}, \mathcal{B})$ such that $\int_a^b \phi(x)\,dx < \infty$ for any $a$ and $b$, $-\infty < a < b < \infty$.

Let $\theta = (a, b)$, $\Theta = \{(a, b) \in \mathcal{R}^2 : a < b\}$, and

$$f_\theta(x) = c(\theta)\phi(x)I_{(a,b)}(x), \quad c(\theta) = \left[\int_a^b \phi(x)\,dx\right]^{-1}$$

Then $\{f_\theta : \theta \in \Theta\}$, called a truncation family, is a parametric family dominated by the Lebesgue measure on $\mathcal{R}$.

Let $X_1, \ldots, X_n$ be i.i.d. random variables having the p.d.f. $f_\theta$.

Then the joint p.d.f. of $X = (X_1, \ldots, X_n)$ is

$$\prod_{i=1}^n f_\theta(x_i) = [c(\theta)]^n I_{(a,\infty)}(x_{(1)})I_{(-\infty,b)}(x_{(n)}) \prod_{i=1}^n \phi(x_i), \quad (1)$$

where $x_{(i)}$ is the $i$th ordered value of $x_1, \ldots, x_n$.

Let $T(X) = (X_{(1)}, X_{(n)})$, $g_\theta(t_1, t_2) = [c(\theta)]^n I_{(a,\infty)}(t_1)I_{(-\infty,b)}(t_2)$, and $h(x) = \prod_{i=1}^n \phi(x_i)$.

By (1) and Theorem 2.2, $T(X)$ is sufficient for $\theta \in \Theta$. 
Example 2.12 (Order statistics)

Let \( X = (X_1, \ldots, X_n) \) and \( X_1, \ldots, X_n \) be i.i.d. random variables having a distribution \( P \in \mathcal{P} \), where \( \mathcal{P} \) is the family of distributions on \( \mathcal{R} \) having Lebesgue p.d.f.'s.

Let \( X_{(1)}, \ldots, X_{(n)} \) be the order statistics given in Example 2.9.

Note that the joint p.d.f. of \( X \) is

\[
f(x_1) \cdots f(x_n) = f(x_{(1)}) \cdots f(x_{(n)}).
\]

Hence, \( T(X) = (X_{(1)}, \ldots, X_{(n)}) \) is sufficient for \( P \in \mathcal{P} \).

The order statistics can be shown to be sufficient even when \( \mathcal{P} \) is not dominated by any \( \sigma \)-finite measure, but Theorem 2.2 is not applicable.
Minimal Sufficiency

Convention: If a statement holds except for outcomes in an event A satisfying $P(A) = 0$ for all $P \in \mathcal{P}$, then we say that the statement holds a.s. $\mathcal{P}$.

Definition 2.5 Minimal sufficiency
Let $T$ be a sufficient statistic for $P \in \mathcal{P}$. $T$ is called a minimal sufficient Statistic iff, for any other statistic $S$ sufficient for $P \in \mathcal{P}$, there is a measurable function $\psi$ such that $T = \psi(S)$ a.s. $\mathcal{P}$

Existence
Minimal sufficient statistics exist under weak assumptions, e.g. $\mathcal{P}$ contains distributions on $\mathcal{R}^k$ dominated by a $\sigma$-finite measure (Bahadur, 1957).
Uniqueness of minimal sufficient statistics

If both $T$ and $S$ are minimal sufficient statistics, then by definition there is one-to-one measurable function $\psi$ such that $T = \psi(S)$ a.s. $P$.

Hence, the minimal sufficient statistic is unique in the sense that two statistics that are one-to-one measurable functions of each other can be treated as one statistic.

Example 2.13

Let $X_1, \ldots, X_n$ be i.i.d. random variables form $P_\theta$, the uniform distribution $U(\theta, \theta + 1)$, $\theta \in \mathbb{R}$, $n > 1$.

The joint Lebesgue p.d.f. of $(X_1, \ldots, X_n)$ is

$$f_\theta(x) = \prod_{i=1}^{n} l_{(\theta, \theta+1)}(x_i) = l_{(x_{(n)}-1, x_{(1)})}(\theta), \quad x = (x_1, \ldots, x_n) \in \mathcal{R}^n,$$

where $x_{(i)}$ denotes the $i$th smallest value of $x_1, \ldots, x_n$.

By Theorem 2.2, $T = (X_{(1)}, X_{(n)})$ is sufficient for $\theta$. 
Example 2.13 (continued)

We now show that \( T = (X_1, X_n) \) is minimal sufficient. Note that

\[ x_1 = \sup \{ \theta : f_\theta(x) > 0 \} \quad \text{and} \quad x_n = 1 + \inf \{ \theta : f_\theta(x) > 0 \}. \]

If \( S(X) \) is a statistic sufficient for \( \theta \), then by Theorem 2.2, there are Borel functions \( h \) and \( g_\theta \) such that \( f_\theta(x) = g_\theta(S(x)) h(x) \). For \( x \) with \( h(x) > 0 \),

\[ x_1 = \sup \{ \theta : g_\theta(S(x)) > 0 \} \quad \text{and} \quad x_n = 1 + \inf \{ \theta : g_\theta(S(x)) > 0 \}. \]

Hence, there is a measurable function \( \psi \) such that \( T(x) = \psi(S(x)) \) when \( h(x) > 0 \).

Since \( h > 0 \), a.s. \( P \), we conclude that \( T \) is minimal sufficient.
Theorem 2.3 (usefull tools for checking minimal sufficiency)

Let $\mathcal{P}$ be a family of distributions on $\mathbb{R}^k$.

(i) Suppose that $\mathcal{P}_0 \subset \mathcal{P}$ and a.s. $\mathcal{P}_0$ implies a.s. $\mathcal{P}$. If $T$ is sufficient for $P \in \mathcal{P}$ and minimal sufficient for $P \in \mathcal{P}_0$, then $T$ is minimal sufficient for $P \in \mathcal{P}$.

(ii) Suppose that $\mathcal{P}$ contains p.d.f.'s $f_0, f_1, f_2, \ldots$, w.r.t. a $\sigma$-finite measure. Let $f_\infty(x) = \sum_{i=0}^{\infty} c_i f_i(x)$, where $c_i > 0$ for all $i$ and $\sum_{i=0}^{\infty} c_i = 1$, and let $T_i(x) = f_i(x)/f_\infty(x)$ when $f_\infty(x) > 0$, $i = 0, 1, 2, \ldots$. Then $T(X) = (T_0, T_1, T_2, \ldots)$ is minimal sufficient for $P \in \mathcal{P}$. Furthermore, if 
\begin{align*}
\{x : f_i(x) > 0\} \subset \{x : f_0(x) > 0\}
\end{align*}
for all $i$, then we may replace $f_\infty(x)$ by $f_0(x)$, in which case $T(X) = (T_1, T_2, \ldots)$ is minimal sufficient for $P \in \mathcal{P}$. 

Suppose that $\mathcal{P}$ contains p.d.f.'s $f_p$ w.r.t. a $\sigma$-finite measure and that there exists a sufficient statistic $T(X)$ such that, for any possible values $x$ and $y$ of $X$, $f_p(x) = f_p(y)\phi(x, y)$ for all $P$ implies $T(x) = T(y)$, where $\phi$ is a measurable function. Then $T(X)$ is minimal sufficient for $P \in \mathcal{P}$.

Proof

(i) If $S$ is sufficient for $P \in \mathcal{P}$, then it is also sufficient for $P \in \mathcal{P}_0$ and, therefore, $T = \psi(S)$ a.s. $\mathcal{P}_0$. The result follows from that a.s. $\mathcal{P}_0$ implies a.s. $\mathcal{P}$.

(ii) Note that $f_\infty > 0$ a.s. $\mathcal{P}$. Let $g_i(T) = T_i$, $i = 0, 1, 2, \ldots$. Then $f_i(x) = g_i(T(x))f_\infty(x)$ a.s. $\mathcal{P}$. By Theorem 2.2, $T$ is sufficient for $P \in \mathcal{P}$. Suppose $S(X)$ is another sufficient statistic, and $f_i(x) = \tilde{g}_i(S(x))h(x)$, $i = 0, 1, 2, \ldots$. Hence

$$T_i(x) = \frac{\tilde{g}_i(S(x))}{\sum_{j=1}^{\infty} c_j\tilde{g}_j(S(x))}$$

for $x$'s satisfying $f_\infty(x) > 0$. By Definition 2.5, $T$ is minimal sufficient for $P \in \mathcal{P}$. The proof is the same when $f_\infty$ is replaced by $f_0$. 
From Bahadur (1957), there is a minimal sufficient statistic $S(X)$. The result follows if we can show that $T(X) = \psi(S(X))$ a.s. $\mathcal{P}$ for a measurable function $\psi$.

By Theorem 2.2, there are Borel functions $h$ and $g_P$ such that $f_P(x) = g_P(S(x))h(x)$ for all $P$. Let $A = \{x : h(x) = 0\}$. Then $P(A) = 0$ for all $P$. For $x$ and $y$ such that $S(x) = S(y)$, $x \not\in A$ and $y \not\in A$,

$$f_P(x) = g_P(S(x))h(x) = g_P(S(y))h(x) = f_P(y)h(x)/h(y)$$

for all $P$. Hence $T(x) = T(y)$. This shows that there is a function $\psi$ such that $T(x) = \psi(S(x))$ except for $x \in A$.

It remains to show that $\psi$ is measurable. Since $S$ is minimal sufficient, $g(T(X)) = S(X)$ a.s. $\mathcal{P}$ for a measurable function $g$. Hence $g$ is one-to-one and $\psi = g^{-1}$. By Theorem 3.9 in Parthasarathy (1967), $\psi$ is measurable.