A statistic $V(X)$ is \textit{ancillary} if its distribution does not depend on the population $P$

$V(X)$ is \textit{first-order ancillary} if $E[V(X)]$ is independent of $P$.

A trivial ancillary statistic is the constant statistic $V(X) \equiv c \in \mathbb{R}$.

If $V(X)$ is a nontrivial ancillary statistic, then $\sigma(V(X)) \subset \sigma(X)$ is a nontrivial $\sigma$-field that does not contain any information about $P$.

Hence, if $S(X)$ is a statistic and $V(S(X))$ is a nontrivial ancillary statistic, it indicates that $\sigma(S(X))$ contains a nontrivial $\sigma$-field that does not contain any information about $P$ and, hence, the “data" $S(X)$ may be further reduced.

A sufficient statistic $T$ appears to be most successful in reducing the data if no nonconstant function of $T$ is ancillary or even first-order ancillary.

This leads to the following definition.
## Finding a complete and sufficient statistic

### Definition 2.6 (Completeness)

A statistic $T(X)$ is said to be complete for $P \in \mathcal{P}$ iff, for any Borel $f$, $E[f(T)] = 0$ for all $P \in \mathcal{P}$ implies $f = 0$ a.s. $\mathcal{P}$.

$T$ is said to be boundedly complete iff the previous statement holds for any bounded Borel $f$.

### Remarks

- A complete statistic is boundedly complete.
- If $T$ is complete (or boundedly complete) and $S = \psi(T)$ for a measurable $\psi$, then $S$ is complete (or boundedly complete).
- Intuitively, a complete and sufficient statistic should be minimal sufficient (Exercise 48).
- A minimal sufficient statistic is not necessarily complete; for example, the minimal sufficient statistic $(X_{(1)}, X_{(n)})$ in Example 2.13 is not complete (Exercise 47).
### Definition 2.6 (Completeness)

A statistic $T(X)$ is said to be *complete* for $P \in \mathcal{P}$ iff, for any Borel $f$, $E[f(T)] = 0$ for all $P \in \mathcal{P}$ implies $f = 0$ a.s. $P$.

$T$ is said to be *boundedly complete* iff the previous statement holds for any bounded Borel $f$.

### Remarks

- A complete statistic is boundedly complete.
- If $T$ is complete (or boundedly complete) and $S = \psi(T)$ for a measurable $\psi$, then $S$ is complete (or boundedly complete).
- Intuitively, a complete and sufficient statistic should be minimal sufficient (Exercise 48).
- A minimal sufficient statistic is not necessarily complete; for example, the minimal sufficient statistic $(X_{(1)}, X_{(n)})$ in Example 2.13 is not complete (Exercise 47).
Proposition 2.1

If $P$ is in an exponential family of full rank with p.d.f.'s given by

$$f_\eta(x) = \exp\{\eta^T T(x) - \zeta(\eta)\} h(x),$$

then $T(X)$ is complete and sufficient for $\eta \in \Xi$.

Proof

We have shown that $T$ is sufficient.
We now show that $T$ is complete.
Suppose that there is a function $f$ such that $E[f(T)] = 0$ for all $\eta \in \Xi$.
By Theorem 2.1(i),

$$\int f(t) \exp\{\eta^T t - \zeta(\eta)\} d\lambda = 0 \quad \text{for all } \eta \in \Xi,$$

where $\lambda$ is a measure on $(\mathbb{R}^p, \mathcal{B}^p)$. 
Proposition 2.1

If $P$ is in an exponential family of full rank with p.d.f.’s given by

$$f_{\eta}(x) = \exp\{\eta \cdot T(x) - \zeta(\eta)\} h(x),$$

then $T(X)$ is complete and sufficient for $\eta \in \Xi$.

Proof

We have shown that $T$ is sufficient. We now show that $T$ is complete. Suppose that there is a function $f$ such that $E[f(T)] = 0$ for all $\eta \in \Xi$. By Theorem 2.1(i),

$$\int f(t) \exp\{\eta \cdot t - \zeta(\eta)\} \, d\lambda = 0 \quad \text{for all } \eta \in \Xi,$$

where $\lambda$ is a measure on $(\mathbb{R}^p, \mathcal{B}^p)$. 
Proof (continued)

Let \( \eta_0 \) be an interior point of \( \Xi \). Then

\[
\int f_+(t)e^{\eta_0 \tau t}d\lambda = \int f_-(t)e^{\eta_0 \tau t}d\lambda \quad \text{for all } \eta \in N(\eta_0),
\]

where \( N(\eta_0) = \{ \eta \in \mathbb{R}^p : \| \eta - \eta_0 \| < \varepsilon \} \) for some \( \varepsilon > 0 \).

In particular,

\[
\int f_+(t)e^{\eta_0 \tau t}d\lambda = \int f_-(t)e^{\eta_0 \tau t}d\lambda = c.
\]

If \( c = 0 \), then \( f = 0 \) a.e. \( \lambda \).

If \( c > 0 \), then \( c^{-1}f_+(t)e^{\eta_0 \tau t} \) and \( c^{-1}f_-(t)e^{\eta_0 \tau t} \) are p.d.f.'s w.r.t. \( \lambda \) and result (1) implies that their m.g.f.'s are the same in a neighborhood of 0.

By Theorem 1.6(ii), \( c^{-1}f_+(t)e^{\eta_0 \tau t} = c^{-1}f_-(t)e^{\eta_0 \tau t} \), i.e., \( f = f_+ - f_- = 0 \) a.e. \( \lambda \).

Hence \( T \) is complete.
Example 2.15

Suppose that $X_1, \ldots, X_n$ are i.i.d. random variables having the $N(\mu, \sigma^2)$ distribution, $\mu \in \mathbb{R}$, $\sigma > 0$.

From Example 2.6, the joint p.d.f. of $X_1, \ldots, X_n$ is

$$(2\pi)^{-n/2} \exp \left\{ \eta_1 T_1 + \eta_2 T_2 - n \zeta(\eta) \right\},$$

where $T_1 = \sum_{i=1}^n X_i$, $T_2 = -\sum_{i=1}^n X_i^2$, and $\eta = (\eta_1, \eta_2) = \left( \frac{\mu}{\sigma^2}, \frac{1}{2\sigma^2} \right)$.

Hence, the family of distributions for $X = (X_1, \ldots, X_n)$ is a natural exponential family of full rank ($\Xi = \mathbb{R} \times (0, \infty)$).

By Proposition 2.1, $T(X) = (T_1, T_2)$ is complete and sufficient for $\eta$.

Since there is a one-to-one correspondence between $\eta$ and $\theta = (\mu, \sigma^2)$, $T$ is also complete and sufficient for $\theta$.

It can be shown that any one-to-one measurable function of a complete and sufficient statistic is also complete and sufficient (exercise).

Thus, $(\bar{X}, S^2)$ is complete and sufficient for $\theta$, where $\bar{X}$ and $S^2$ are the sample mean and sample variance, respectively.
Example 2.16

Let $X_1, \ldots, X_n$ be i.i.d. random variables from $P_\theta$, the uniform distribution $U(0, \theta)$, $\theta > 0$.

The largest order statistic, $X_{(n)}$, is complete and sufficient for $\theta \in (0, \infty)$. The sufficiency of $X_{(n)}$ follows from the fact that the joint Lebesgue p.d.f. of $X_1, \ldots, X_n$ is $\theta^{-n}I_{(0, \theta)}(x_{(n)})$.

From Example 2.9, $X_{(n)}$ has the Lebesgue p.d.f. $(nx^{n-1}/\theta^n)I_{(0, \theta)}(x)$.

Let $f$ be a Borel function on $[0, \infty)$ such that $E[f(X_{(n)})] = 0$ for all $\theta > 0$. Then

$$\int_0^\theta f(x)x^{n-1}dx = 0 \quad \text{for all } \theta > 0.$$ 

Let $G(\theta)$ be the left-hand side of the previous equation. Applying the result of differentiation of an integral (see, e.g., Royden (1968, §5.3)), we obtain that $G'(\theta) = f(\theta)\theta^{n-1}$ a.e. $m_+$, where $m_+$ is the Lebesgue measure on $([0, \infty), B_{[0, \infty)}$).

Since $G(\theta) = 0$ for all $\theta > 0$, $f(\theta)\theta^{n-1} = 0$ a.e. $m_+$ and, hence, $f(x) = 0$ a.e. $m_+$.

Therefore, $X_{(n)}$ is complete and sufficient for $\theta \in (0, \infty)$. 
Example 2.17

In Example 2.12, we showed that the order statistics

\[ T(X) = (X_{(1)}, \ldots, X_{(n)}) \]

of i.i.d. random variables \( X_1, \ldots, X_n \) is sufficient for \( P \in \mathcal{P} \), where \( \mathcal{P} \) is the family of distributions on \( \mathbb{R} \) having Lebesgue p.d.f.'s.

We now show that \( T(X) \) is also complete for \( P \in \mathcal{P} \).

Let \( \mathcal{P}_0 \) be the family of Lebesgue p.d.f.'s of the form

\[ f(x) = C(\theta_1, \ldots, \theta_n) \exp\{-x^{2n} + \theta_1 x + \theta_2 x^2 + \cdots + \theta_n x^n\}, \]

where \( \theta_j \in \mathbb{R} \) and \( C(\theta_1, \ldots, \theta_n) \) is a normalizing constant such that \( \int f(x) dx = 1 \).

Then \( \mathcal{P}_0 \subset \mathcal{P} \) and \( \mathcal{P}_0 \) is an exponential family of full rank.

Note that the joint distribution of \( X = (X_1, \ldots, X_n) \) is also in an exponential family of full rank.

Thus, by Proposition 2.1, \( U = (U_1, \ldots, U_n) \) is a complete statistic for \( P \in \mathcal{P}_0 \), where \( U_j = \sum_{i=1}^n X_i^j \).

Since a.s. \( \mathcal{P}_0 \) implies a.s. \( \mathcal{P} \), \( U(X) \) is also complete for \( P \in \mathcal{P} \).
Example 2.17 (continued)

The result follows if we can show that there is a one-to-one correspondence between $T(X)$ and $U(X)$.
Let $V_1 = \sum_{i=1}^{n} X_i$, $V_2 = \sum_{i<j} X_i X_j$, $V_3 = \sum_{i<j<k} X_i X_j X_k$, ..., $V_n = X_1 \cdots X_n$. From the identities

$$U_k - V_1 U_{k-1} + V_2 U_{k-2} - \cdots + (-1)^{k-1} V_{k-1} U_1 + (-1)^k k V_k = 0,$$

$k = 1, ..., n$, there is a one-to-one correspondence between $U(X)$ and $V(X) = (V_1, ..., V_n)$.
From the identity

$$(t - X_1) \cdots (t - X_n) = t^n - V_1 t^{n-1} + V_2 t^{n-2} - \cdots + (-1)^n V_n,$$

there is a one-to-one correspondence between $V(X)$ and $T(X)$. This completes the proof and, hence, $T(X)$ is sufficient and complete for $P \in \mathcal{P}$.
In fact, both $U(X)$ and $V(X)$ are sufficient and complete for $P \in \mathcal{P}$. 
The relationship between an ancillary statistic and a complete and sufficient statistic is characterized in the following result.

**Theorem 2.4 (Basu’s theorem)**

Let $V$ and $T$ be two statistics of $X$ from a population $P \in \mathcal{P}$. If $V$ is ancillary and $T$ is boundedly complete and sufficient for $P \in \mathcal{P}$, then $V$ and $T$ are independent w.r.t. any $P \in \mathcal{P}$.

**Proof**

Let $B$ be an event on the range of $V$.

Since $V$ is ancillary, $P(V^{-1}(B))$ is a constant.

As $T$ is sufficient, $E[I_B(V)|T]$ is a function of $T$ (not dependent on $P$). Because

$$E\{E[I_B(V)|T] - P(V^{-1}(B))\} = 0 \quad \text{for all } P \in \mathcal{P},$$

by the bounded completeness of $T$,

$$P(V^{-1}(B)|T) = E[I_B(V)|T] = P(V^{-1}(B)) \quad \text{a.s. } \mathcal{P}.$$
The relationship between an ancillary statistic and a complete and sufficient statistic is characterized in the following result.

**Theorem 2.4 (Basu’s theorem)**

Let $V$ and $T$ be two statistics of $X$ from a population $P \in \mathcal{P}$. If $V$ is ancillary and $T$ is boundedly complete and sufficient for $P \in \mathcal{P}$, then $V$ and $T$ are independent w.r.t. any $P \in \mathcal{P}$.

**Proof**

Let $B$ be an event on the range of $V$. Since $V$ is ancillary, $P(V^{-1}(B))$ is a constant. As $T$ is sufficient, $E[I_B(V)|T]$ is a function of $T$ (not dependent on $P$). Because

$$E\{E[I_B(V)|T] - P(V^{-1}(B))\} = 0 \quad \text{for all } P \in \mathcal{P},$$

by the bounded completeness of $T$,

$$P(V^{-1}(B)|T) = E[I_B(V)|T] = P(V^{-1}(B)) \quad \text{a.s. } \mathcal{P}.$$
Proof (continued)

Let $A$ be an event on the range of $T$. Then

$$P(T^{-1}(A) \cap V^{-1}(B)) = E\{E[I_A(T)I_B(V) | T]\} = E\{I_A(T)E[I_B(V) | T]\}$$

$$= E\{I_A(T)P(V^{-1}(B))\} = P(T^{-1}(A))P(V^{-1}(B)).$$

Hence $T$ and $V$ are independent w.r.t. any $P \in \mathcal{P}$.

Remark

Basu’s theorem is useful in proving the independence of two statistics.

Example 2.18

Suppose that $X_1, \ldots, X_n$ are i.i.d. random variables having the $N(\mu, \sigma^2)$ distribution, with $\mu \in \mathbb{R}$ and a known $\sigma > 0$.

It can be easily shown that the family $\{N(\mu, \sigma^2) : \mu \in \mathbb{R}\}$ is an exponential family of full rank with natural parameter $\eta = \mu / \sigma^2$.

By Proposition 2.1, the sample mean $\bar{X}$ is complete and sufficient for $\eta$ (and $\mu$).
Proof (continued)

Let $A$ be an event on the range of $T$. Then

$$P(T^{-1}(A) \cap V^{-1}(B)) = E\{E[I_A(T)I_B(V)|T]\} = E\{I_A(T)E[I_B(V)|T]\}$$

$$= E\{I_A(T)P(V^{-1}(B))\} = P(T^{-1}(A))P(V^{-1}(B)).$$

Hence $T$ and $V$ are independent w.r.t. any $P \in \mathcal{P}$.

Remark

Basu’s theorem is useful in proving the independence of two statistics.

Example 2.18

Suppose that $X_1, \ldots, X_n$ are i.i.d. random variables having the $N(\mu, \sigma^2)$ distribution, with $\mu \in \mathbb{R}$ and a known $\sigma > 0$.

It can be easily shown that the family $\{N(\mu, \sigma^2) : \mu \in \mathbb{R}\}$ is an exponential family of full rank with natural parameter $\eta = \mu / \sigma^2$.

By Proposition 2.1, the sample mean $\bar{X}$ is complete and sufficient for $\eta$ (and $\mu$).
Proof (continued)

Let $A$ be an event on the range of $T$. Then

$$P\left( T^{-1}(A) \cap V^{-1}(B) \right) = E\left\{ E[I_A(T)I_B(V) \mid T] \right\} = E\left\{ I_A(T)E[I_B(V) \mid T] \right\}$$

$$= E\left\{ I_A(T)P(V^{-1}(B)) \right\} = P(T^{-1}(A))P(V^{-1}(B)).$$

Hence $T$ and $V$ are independent w.r.t. any $P \in \mathcal{P}$.

Remark

Basu’s theorem is useful in proving the independence of two statistics.

Example 2.18

Suppose that $X_1, \ldots, X_n$ are i.i.d. random variables having the $N(\mu, \sigma^2)$ distribution, with $\mu \in \mathbb{R}$ and a known $\sigma > 0$.

It can be easily shown that the family $\{N(\mu, \sigma^2) : \mu \in \mathbb{R}\}$ is an exponential family of full rank with natural parameter $\eta = \mu / \sigma^2$.

By Proposition 2.1, the sample mean $\bar{X}$ is complete and sufficient for $\eta$ (and $\mu$).
Example 2.18 (continued)

Let $S^2$ be the sample variance. Since $S^2 = (n - 1)^{-1} \sum_{i=1}^{n} (Z_i - \bar{Z})^2$, where $Z_i = X_i - \mu$ is $N(0, \sigma^2)$ and $\bar{Z} = n^{-1} \sum_{i=1}^{n} Z_i$, $S^2$ is an ancillary statistic ($\sigma^2$ is known).

By Basu’s theorem, $\bar{X}$ and $S^2$ are independent w.r.t. $N(\mu, \sigma^2)$ with $\mu \in \mathbb{R}$.

Using the independence of $\bar{X}$ and $S^2$, we now show that $(n - 1)S^2/\sigma^2$ has the chi-square distribution $\chi^2_{n-1}$.

Note that

$$n \left( \frac{\bar{X} - \mu}{\sigma} \right)^2 + (n - 1)S^2 = \sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2.$$ 

From the properties of the normal distributions, $n(\bar{X} - \mu)^2/\sigma^2$ has the chi-square distribution $\chi^2_1$ with the m.g.f. $(1 - 2t)^{-1/2}$ and $\sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2$ has the chi-square distribution $\chi^2_n$ with the m.g.f. $(1 - 2t)^{-n/2}$, $t < 1/2$. 

Example 2.18 (continued)

Let $S^2$ be the sample variance.
Since $S^2 = (n-1)^{-1} \sum_{i=1}^{n} (Z_i - \bar{Z})^2$, where $Z_i = X_i - \mu$ is $N(0, \sigma^2)$ and $ar{Z} = n^{-1} \sum_{i=1}^{n} Z_i$, $S^2$ is an ancillary statistic ($\sigma^2$ is known).

By Basu’s theorem, $\bar{X}$ and $S^2$ are independent w.r.t. $N(\mu, \sigma^2)$ with $\mu \in \mathbb{R}$.
Since $\sigma^2$ is arbitrary, $\bar{X}$ and $S^2$ are independent w.r.t. $N(\mu, \sigma^2)$ for any $\mu \in \mathbb{R}$ and $\sigma^2 > 0$.

Using the independence of $\bar{X}$ and $S^2$, we now show that $(n-1)S^2/\sigma^2$ has the chi-square distribution $\chi^2_{n-1}$.

Note that
\[
n \left( \frac{\bar{X} - \mu}{\sigma} \right)^2 + \frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2.
\]

From the properties of the normal distributions, $n(\bar{X} - \mu)^2/\sigma^2$ has the chi-square distribution $\chi^2_1$ with the m.g.f. $(1 - 2t)^{-1/2}$ and $\sum_{i=1}^{n}(X_i - \mu)^2/\sigma^2$ has the chi-square distribution $\chi^2_n$ with the m.g.f. $(1 - 2t)^{-n/2}$, $t < 1/2$. 


Example 2.18 (continued)

By the independence of $\bar{X}$ and $S^2$, the m.g.f. of $(n - 1)S^2/\sigma^2$ is

$$(1 - 2t)^{-n/2}/(1 - 2t)^{-1/2} = (1 - 2t)^{-(n-1)/2}$$

for $t < 1/2$.

This is the m.g.f. of the chi-square distribution $\chi^2_{n-1}$ and, therefore, the result follows.