## 1 Complete Statistics

Suppose $X \sim P_{\theta}, \theta \in \Theta$. Let $\left(X_{(1)}, \ldots, X_{(n)}\right)$ denote the order statistics.
Definition 1. A statistic $T=T(X)$ is complete if

$$
E_{\theta} g(T)=0 \quad \text { for all } \quad \theta
$$

implies

$$
P_{\theta}(g(T)=0)=1 \quad \text { for all } \quad \theta \text {. }
$$

(Note: $E_{\theta}$ denotes expectation computed with respect to $P_{\theta}$ ).
Example: $X=\left(X_{1}, \ldots, X_{n}\right)$ iid $\mathrm{N}(\theta, 1) . \quad T(X)=\left(X_{1}, X_{2}\right)$ is a statistic which is not complete because

$$
E\left(X_{1}-X_{2}\right)=0 \quad \text { for all } \theta .
$$

and $X_{1}-X_{2}$ here is a function of $T$ where $g: \mathbb{R}^{2} \rightarrow \mathbb{R}$ given by $\left(x_{1}, x_{2}\right) \mapsto x_{1}-x_{2}$ but $P\left(X_{1}-X_{2}=0\right) \neq 1$ for all $\theta$.
More formally: $T$ is not complete because the function $g$ satisfies

$$
\begin{aligned}
E_{\theta} g(T) & =0 \text { for all } \theta \text { but } \\
P_{\theta}(g(T)=0) & \neq 1 \text { for all } \theta .
\end{aligned}
$$

Example: $X=\left(X_{1}, \ldots, X_{n}\right)$ iid $\operatorname{Unif}(\theta, \theta+1) . T=T(X)=\left(X_{(1)}, X_{(n)}\right)$ is MSS. But $T$ is not complete. We know that $S(X)=X_{(n)}-X_{(1)}$ is ancillary. Thus

$$
E\left(X_{(n)}-X_{(1)}\right)=c
$$

where $c$ does not depend on $\theta$ and therefore,

$$
E\left(X_{(n)}-X_{(1)}-c\right)=0 \quad \text { for all } \quad \theta
$$

but clearly,

$$
P\left(X_{(n)}-X_{(1)}-c=0\right) \neq 1 \quad \text { for all } \theta .
$$

Hence taking $g(T)=X_{(n)}-X_{(1)}-c$, we can see that $T$ is not complete.
Example: $X=X_{1}, \ldots, X_{n}$ iid $\operatorname{Unif}(0, \theta) T=T(X)=X_{(n)}$ is MSS.
Fact: $T$ is also complete.

Proof. Assume there exists $g$ such that

$$
E_{\theta} g(T)=0 \quad \text { for all } \theta
$$

$T$ has cdf $H(t)=(t / \theta)^{n}, 0 \leq t \leq \theta$ and pdf $n t^{n-1} / \theta^{n}, 0 \leq t \leq \theta$. Then

$$
E g(T)=\int_{0}^{\theta} g(t) \frac{n t^{n-1}}{\theta^{n}} d t=0 \quad \text { for all } \quad \theta>0
$$

implies

$$
\int_{0}^{\theta} g(t) n t^{n-1} d t=0 \quad \text { for all } \quad \theta>0
$$

implies (by differentiating both sides and using the Fundamental Theorem of Calculus)

$$
g(\theta) n \theta^{n-1} d t=0 \quad \text { for all } \quad \theta>0
$$

implies

$$
g(t)=0 \quad \text { for all } \quad t>0
$$

implies

$$
P_{\theta}(g(T)=0)=1 \quad \text { for all } \quad \theta>0
$$

Theorem 1. Suppose $X_{1}, \ldots, X_{n}$ iid with pdf (pmf)

$$
f(x \mid \theta)=c(\theta) h(x) \exp \left\{\sum_{j=1}^{k} w_{j}(\theta) t_{j}(x)\right\}
$$

for $\theta=\left(\theta_{1}, \ldots, \theta_{k}\right) \in \Theta$. Let $X=\left(X_{1}, \ldots, X_{n}\right)$. Define

$$
T(X)=\left(\sum_{i=1}^{n} t_{1}\left(X_{i}\right), \sum_{i=1}^{n} t_{2}\left(X_{i}\right), \ldots, \sum_{i=1}^{n} t_{k}\left(X_{i}\right)\right)
$$

Then

1. $T(X)$ is sufficient statistic for $\theta$.
2. If $\Theta$ contains an open set in $\mathbb{R}^{k}$, then $T(X)$ is complete. (More precisely, if

$$
\left\{\left(w_{1}(\theta), w_{2}(\theta), \ldots, w_{k}(\theta)\right): \theta \in \Theta\right\}
$$

contains an open set in $\mathbb{R}^{k}$, then $T(X)$ is complete.)

## Remarks:

1. The statistic $T(X)$ in the Theorem is called the natural sufficient statistic.
2. $\eta=\left(\eta_{1}, \ldots, \eta_{k}\right) \equiv\left(w_{1}(\theta), \ldots, w_{k}(\theta)\right)$ is called the natural parameter of the exponential family.
3. Condition (2) is the "open set condition" (OSC). The OSC is easily verified by inspection. Let $A \subset \mathbb{R}^{k}$. A contains an open set in $\mathbb{R}^{k}$ iff A contains a $k$-dimensional ball. That is, there exists $x \in \mathbb{R}^{k}$ and $r>0$ such that $B(x, r) \subset A$. Here $B(x, r)$ denotes the ball of radius $r$ about $x$. Let $A \subset \mathbb{R}$ (take $k=1$ ). A contains an open set in $\mathbb{R}$ if and only if $A$ contains an interval. That is, there exists $c<d$ such that $(c, d) \subset A$.

## Facts:

1. Under weak conditions (which are almost always true, a complete sufficient statistic is also minimal. Abbreviation: CSS $\Rightarrow$ MSS. (but MSS does not imply CSS as we saw earlier).
2. A one-to-one function of a CSS is also a CSS (See later remarks). Reminder: A 1-1 function of an MSS is also an MSS.

Example: The Bernoulli pmf is an exponential family (1pef):

$$
p(x \mid \theta)=\theta^{x}(1-\theta)^{1-x}, x \in\{0,1\}=(1-\theta) I(x \in\{0,1\}) \exp \left[x \log \left(\frac{\theta}{1-\theta}\right)\right] .
$$

If $X_{1}, \ldots, X_{n}$ are iid $p(x \mid \theta)$, then $T=\sum_{i} X_{i}$ is a SS. It is also complete if $\Theta$ contains an interval. Here is a direct proof of completeness (not relying on our general theorem on exponential families).

Proof. We know $T \sim \operatorname{Binomial}(n, \theta)$. If $E g(T)=0$ for all $\theta \in \Theta \subset(0,1)$, then

$$
\sum_{k=0}^{n} g(k)\binom{n}{k} \theta^{k}(1-\theta)^{n-k}=0
$$

for all $\theta \in \Theta$. Then

$$
(1-\theta)^{n} \sum_{k=0}^{n} g(k)\binom{n}{k}\left(\frac{\theta}{1-\theta}\right)^{k}=0
$$

for all $\theta \in \Theta$. Then

$$
\sum_{k=0}^{n} a_{k} u^{k}=0
$$

for all $\theta \in \Theta$ where $a_{k}=g(k)\binom{n}{k}$ and $u=\theta /(1-\theta)$. This if $\Theta$ contains an interval, then the above implies that the polynomial $\psi(u)=\sum_{k=1}^{n} a_{k} u^{k}$ is identically zero in some interval. This implies all the coefficients $a_{k}$ must be zero, which further implies $g(k)=0$ for $k=0,1, \ldots, n$ so that $P_{\theta}(g(T)=0)=1$ for all $\theta \in \Theta$.

In a homework exercise, you show that $T=\sum_{i=1}^{n} X_{i}$ is complete when $X_{1}, \ldots, X_{n}$ are iid Poisson $(\lambda)$ in a similar way, using the fact that an infinite power series (an analytic function) is identically zero in some interval if and only if all the coefficients are zero.
Example: The $\mathrm{N}(\theta, 1)$ family is a 1 pef with $w(\theta)=\theta, t(x)=x$. Let $X=\left(X_{1}, \ldots, X_{n}\right)$ $\overline{\mathrm{iid}} \mathrm{N}(\theta, 1) . T(X)=\sum_{i=1}^{n} X_{i}$ is the natural SS (It is sufficient for any $\Theta$ ). Is $T$ complete? This depends on $\Theta$.

1. $\Theta=\mathbb{R}$. Yes. (OSC holds)
2. $\Theta=[0.01,0.02]$. Yes. (OSC holds)
3. $\Theta=(1,2) \cup\{4,7\}$. Yes. (OSC holds)
4. $\Theta=\mathbb{Z}$ (the integers). OSC fails so Theorem says nothing. But can show that it is not complete.
5. $\Theta=\mathbb{Q}$ (the rationals). OSC fails so Theorem says nothing. Yes or no? Don't know.
6. $\Theta=$ Cantor set. OSC fails so Theorem says nothing. Yes or no? Don't know.
7. $\Theta=$ finite set. OSC fails so Theorem says nothing. But can show that it is not complete.

Remark: In general, it is typically true that if $\Theta$ is finite and the support of $T=T(X)$ is infinite, then $T$ is not complete.
Example: The $\mathrm{N}\left(\mu, \sigma^{2}\right)$ family with $\theta=\left(\mu, \sigma^{2}\right)$ is a 2 pef with

$$
w(\theta)=\left(\frac{\mu}{\sigma^{2}}, \frac{-1}{2 \sigma^{2}}\right), \quad t(x)=\left(x, x^{2}\right) .
$$

Let $X=\left(X_{1}, \ldots, X_{n}\right)$ iid $\mathrm{N}\left(\mu, \sigma^{2}\right)$. Then $T(X)=\left(\sum_{i=1}^{n} X_{i}, \sum_{i=1}^{n} X_{i}^{2}\right)$ is the natural $S S$. (It is a SS for any $\Theta$ ). $T(X)$ is a one-to-one function of $U(X)=\left(\bar{x}, s^{2}\right)$. So $T$ is CSS iff $U$ is CSS. Is $T$ (or $U$ ) complete? That depends on $\Theta$.

1. $\Theta_{1}=\left\{\left(\mu, \sigma^{2}\right): \sigma^{2}>0\right\}$. OSC holds. Yes, complete.
2. $\Theta_{2}=\left\{\left(\mu, \sigma^{2}\right): \sigma^{2}=\sigma_{0}^{2}\right\}$. OSC fails. Theorem says nothing. But we can prove that $U$ is not complete.

Proof. Let $g\left(x_{1}, x_{2}\right)=x_{2}-\sigma_{0}^{2}$. Then $E g(U)=E\left(s^{2}-\sigma_{0}^{2}\right)=\sigma^{2}-\sigma_{0}^{2}=0$ for all $\theta \in \Theta_{2}$.
3. $\Theta_{3}=\left\{\left(\mu, \sigma^{2}\right): \mu=\mu_{0}, \sigma^{2}>0\right\}$. OSC fails. Theorem says nothing. But we can prove that $U$ is not complete.

Proof. Let $g\left(x_{1}, x_{2}\right)=x_{1}-\mu_{0}$. Then $E g(U)=E\left(\bar{x}-\mu_{0}\right)=\mu-\mu_{0}=0$ for all $\theta \in \Theta_{3}$.
4. $\Theta_{4}=\left\{\left(\mu, \sigma^{2}\right): \mu=\sigma^{2}, \sigma^{2}>0\right\}$. OSC fails. Theorem says nothing. But we can prove that $U$ is not complete.

Proof. Let $g\left(x_{1}, x_{2}\right)=x_{1}-x_{2}$. Then $E g(U)=E\left(\bar{x}-s^{2}\right)=\mu-\sigma^{2}=0$ for all $\theta \in \Theta_{4}$.
(Note: It is more natural to describe the families $\Theta_{2}, \Theta_{3}, \Theta_{4}$ as 1pef's. If you do this, you get different natural sufficient statistics, which turn out to be complete.)
5. $\Theta_{5}=\left\{\left(\mu, \sigma^{2}\right): \mu^{2}=\sigma^{2}, \sigma^{2}>0\right\}$. OSC fails. Theorem says nothing. But we can prove that $U$ is not complete.

Proof. Homework
6. $\Theta_{6}=[1,3] \times[4,6]$. OSC holds. Yes, complete.
7. $\Theta_{7}=\Theta_{6} \cup\{(5,1),(4,2)\}$. OSC holds. Yes, complete.
8. $\Theta_{8}=$ complicated wavy curve. OSC fails. Theorem says nothing. But hard to conclude anything.

Corollary 1. Suppose $X \in \mathbb{R}^{m}$ has joint pdf (pmf)

$$
f(x \mid \theta)=c(\theta) h(x) \exp \left\{\sum_{j=1}^{k} w_{j}(\theta) t_{j}(x)\right\}
$$

for all $x \in \mathbb{R}^{m}$ where $\theta=\left(\theta_{1}, \ldots, \theta_{k}\right) \in \Theta$. Define $T(X)=\left(t_{1}(X), t_{2}(X), \ldots, t_{k}(X)\right)$. Then

1. $T(X)$ is sufficient statistic for $\theta$.
2. If $\Theta$ contains an open set in $\mathbb{R}^{k}$, then $T(X)$ is complete.
(More precisely, if

$$
\left\{\left(w_{1}(\theta), w_{2}(\theta), \ldots, w_{k}(\theta)\right): \theta \in \Theta\right\}
$$

contains an open set in $\mathbb{R}^{k}$, then $T(X)$ is complete.)
Example: Return to Simple Linear Regression: $X_{1}, \ldots, X_{n}$ independent with $X_{i} \sim \mathrm{~N}\left(\beta_{0}+\right.$ $\overline{\left.\beta_{1} z_{i}, \sigma^{2}\right)}$. $\theta=\left(\beta_{0}, \beta_{1}, \sigma^{2}\right)$ and $\Theta=\mathbb{R}^{2} \times(0, \infty)$. Recall that the joint density of $X=$ $\left(X_{1}, \ldots, X_{n}\right)$ is

$$
\begin{aligned}
f(x \mid \theta)= & \left(2 \pi \sigma^{2}\right)^{-n / 2} \exp \left\{-\frac{1}{2 \sigma^{2}}\left(\sum_{i} x_{i}^{2}-2 \beta_{0} \sum_{i=1}^{n} x_{i}-2 \beta_{1} \sum_{i=1}^{n} z_{i} x_{i}+\sum_{i=1}^{n}\left(\beta_{0}+\beta_{1} z_{i}\right)^{2}\right)\right\} \\
= & {\left[\left(2 \pi \sigma^{2}\right)^{-n / 2} \exp \left\{-\frac{1}{2 \sigma^{2}} \sum_{i}\left(\beta_{0}+\beta_{1} z_{i}\right)^{2}\right\}\right] \times 1 } \\
& \times \exp \left\{\frac{-1}{2 \sigma^{2}} \sum_{i} x_{i}^{2}+\frac{\beta_{0}}{\sigma^{2}} \sum_{i} x_{i}+\frac{\beta_{1}}{\sigma^{2}} \sum_{i} z_{i} x_{i}\right\} \\
= & c(\theta) h(x) \exp \left\{\sum_{j=1}^{3} w_{j}(\theta) t_{j}(x)\right\}
\end{aligned}
$$

where

$$
\begin{aligned}
w(\theta) & =\left(w_{1}(\theta), w_{2}(\theta), w_{3}(\theta)\right)=\left(\frac{-1}{2 \sigma^{2}}, \frac{\beta_{0}}{\sigma^{2}}, \frac{\beta_{1}}{\sigma^{2}}\right) \\
t(x) & =\left(t_{1}(x), t_{2}(x), t_{3}(x)\right)=\left(\sum_{i=1}^{n} x_{i}^{2}, \sum_{i=1}^{n} x_{i}, \sum_{i=1}^{n} z_{i} x_{i}\right) .
\end{aligned}
$$

The data vector $X$ may be regarded as a single observation from an $n$-dimensional 3pef. Since $\Theta \subset \mathbb{R}^{3}$ satisfied the OSC, the statistic $T(X)=t(X)=\left(\sum_{i=1}^{n} X_{i}^{2}, \sum_{i=1}^{n} X_{i}, \sum_{i=1}^{n} z_{i} X_{i}\right)$ is a CSS.
Notation: $X \sim P_{\theta}, \theta \in \Theta . S(X)=\psi(T(X))$ for some $\psi . \Theta_{1} \subset \Theta_{2} \subset \Theta$.
Sufficiency:

1. If $S(X)$ is sufficient, then $T(X)$ is sufficient.
2. If $T(X)$ is sufficient for $\Theta_{2}$, then $T(X)$ is sufficient for $\Theta_{1}$.

Completeness:

1. If $T(X)$ is complete, then $S(X)$ is complete.
2. If $T(X)$ is complete for $\Theta_{1}$, then $T(X)$ is complete for $\Theta_{2}$ (under mild regularity conditions).

Proof. Proof of 1. $E_{\theta} g(S(X))=0$ for all $\theta \in \Theta \Longrightarrow E_{\theta} g(\psi(T(X)))=0$ for all $\theta \in$ $\Theta \Longrightarrow P_{\theta}\{g(\psi(T(X)))=0\}=1$ for all $\theta \in \Theta$ (by completeness of $T(X)$ ) which implies $P_{\theta}\{g(S(X))=0\}=1$ for all $\theta$.
Proof of 2. $E_{\theta} g(T(X))=0$ for all $\theta \in \Theta_{2} \Longrightarrow E_{\theta} g(T(X))=0$ for all $\theta \in \Theta_{1} \Longrightarrow$ $P_{\theta}\{g(T(X))=0\}=1$ for all $\theta \in \Theta_{1}$ (by completeness of $T(X)$ for $\Theta_{1}$ ) which implies $P_{\theta}\{g(T(X))=0\}=1$ for all $\theta \in \Theta_{2}$ (under mild assumptions).

Ancillarity:

1. If $T(X)$ is ancillary, then $S(X)$ is ancillary.
2. If $T(X)$ is ancillary for $\Theta_{2}$, then $T(X)$ is ancillary for $\Theta_{1}$.

Proof. Proof of 1. Uses $Y \stackrel{d}{=} Z \Longrightarrow \psi(Y) \stackrel{d}{=} \psi(Z)$.
Proof of 2. Trivial

## 2 Basu's results

Suppose $X \sim P_{\theta}, \theta \in \Theta$.
Lemma 1. (Basu's Lemma) If $T(X)$ is complete and sufficient (for $\theta \in \Theta$ ), and $S(X)$ is ancillary, then $S(X)$ and $T(X)$ are independent for all $\theta \in \Theta$.

In other words, a complete sufficient statistic is independent of any ancillary statistic.

### 2.1 Remarks:

Let $S=S(X), T=T(X)$. Let $E_{\theta}$ denote expectation w.r.t. $P_{\theta}$.

1. The joint distribution of $(S, T)$ depends on $\theta$, so in general it is possible for $S$ and $T$ to be independent for some values of $\theta$, but not for others. (Basu's Lemma says this does not happen in this case.)
2. For any rv's $Y$ and $Z$, we know that $E(Y \mid Z)=g(Z)$, i.e., the conditional expectation is a function of $Z$. If the joint distribution of $(Y, Z)$ depends on a parameter $\theta$, then $E_{\theta}(Y \mid Z)=g(Z, \theta)$, i.e., the conditional expectation is a function of both $Z$ and $\theta$. (However, this function may turn out to be constant in one or both variables.)
3. In general, $E(Y)=E\{E(Y \mid Z)\}$ and $E_{\theta}(Y)=E_{\theta} E_{\theta}(Y \mid Z)$.
4. To show that $Y$ and $Z$ are independent, it suffices to show that $\mathcal{L}(Y \mid Z)=\mathcal{L}(Y)$ which means that $P(Y \in A \mid Z)=P(Y \in A)$ for all (Borel) sets $A$. Let $w(Y)=$ $I(Y \in A)$. Then $P(Y \in A)=E w(Y)$ and $P(Y \in A \mid Z)=E(w(Y) \mid Z)$. The indicator function $w(Y)$ is a bounded (Borel measurable) function. Therefore we have:
To show that $Y$ and $Z$ are independent, it suffices to show $E(w(Y) \mid Z)=E w(Y)$ for all bounded (Borel measurable) functions.
5. Thus, to show that $S$ and $T$ are independent for all $\theta$, it suffices to show that $E_{\theta}(w(S) \mid T)=E_{\theta} w(S)$ for all $\theta$ and all bounded measurable functions $w(S)$.

### 2.2 Proof of Basu's Lemma

Proof. Let $w(S)$ be a given bounded function of $S$. Consider both sides of the identity:

$$
E_{\theta} w(S)=E_{\theta}\left[E_{\theta}(w(S) \mid T)\right]
$$

for all $\theta$. Consider the LHS. Since $S$ is ancillary, the distribution of $w(S)$ is the same for all $\theta$ so that the LHS is constant in $\theta$. Call this constant $c$.
Now consider the RHS. We know that $E_{\theta}(w(S) \mid T)$ will be some function of $\theta$ and $T$. However, since $T$ is a sufficient statistic, $\mathcal{L}(X \mid T)$ does not depend on $\theta$. Since $S=S(X)$, this implies $\mathcal{L}(S \mid T)$ does not depend on $\theta$ so that in turn $\mathcal{L}(w(S) \mid T)$ does not depend on $\theta$. Thus, by sufficiency, $E_{\theta}(w(S) \mid T)$ is constant in $\theta$ and must be a function of $T$ only. Call this function $\psi(T)$. That is,

$$
\psi(T)=E_{\theta}(w(S) \mid T)
$$

The original identity can now be written as

$$
\begin{aligned}
& c=E_{\theta} \psi(T) \text { for all } \theta \text { or equivalently } \\
& 0=E_{\theta}(\psi(T)-c) \text { for all } \theta .
\end{aligned}
$$

Since T is complete, this implies

$$
\begin{array}{rll}
P(\psi(T)-c=0) & =1 \quad \text { for all } \theta \quad \text { or equivalently } \\
\psi(T) & =c \quad \text { with probability one for all } \theta .
\end{array}
$$

This means

$$
E_{\theta}(w(S) \mid T)=E_{\theta} w(S) \quad \text { with probability one for all } \quad \theta
$$

Since $w(S)$ is an arbitrary bounded function, by the earlier discussion this implies $S$ and $T$ are independent for all $\theta$.

### 2.3 Applications of Basu's theorem

1. Example: Let $X=\left(X_{1}, \ldots, X_{n}\right)$ iid $\operatorname{Unif}(0, \theta)$. Recall: $T(X)=X_{(n)}=\max \left\{X_{i}\right\}$ is a CSS. Unif $(0, \theta)$ is a scale family, so any scale invariant statistic $\mathrm{S}(\mathrm{X})$ is ancillary. Thus, by Basus lemma, all of the following are independent of $X_{(n)}$ for all $\theta$ :

$$
\frac{\bar{x}}{s}, \frac{X_{(1)}}{X_{(n)}}, \quad\left(\frac{X_{(1)}}{X_{(n)}}, \ldots, \frac{X_{(n-1)}}{X_{(n)}}\right)
$$

2. Example: (Using Basu's Lemma to obtain an indirect proof of completeness.) Let $\left.\overline{X=\left(X_{1}\right.}, \ldots, X_{n}\right)$ iid $\operatorname{Unif}(\theta, \theta+1)$. Recall: $T(X)=\left(X_{(1)}, X_{(n)}\right)$ is a MSS. $S(X)=$ $X_{(n)}-X_{(1)}$ is ancillary. Since $S$ is a function of $T$, the $\operatorname{rvs} S$ and $T$ cannot be independent. Thus $T$ cannot be complete (for then we would get a contradiction with Basus Lemma).
3. Example: Let $X=\left(X_{1}, \ldots, X_{n}\right)$ be iid $\mathrm{N}\left(\mu, \sigma^{2}\right)$. Let $\bar{x}, s^{2}, Z$ be the sample mean, sample variance, and standardized residuals (z-scores) of the data $X$ (Recall: $Z=$ $\left(Z_{1}, \ldots, Z_{n}\right)$ with $Z_{i}=\left(X_{i}-\bar{x}\right) / s$.
Fact: $\bar{x}, s^{2}, Z$ are mutually independent.
Proof. We first show that the pair $\left(\bar{x}, s^{2}\right)$ is independent of $Z$, and then show that $\bar{x}$ and $s^{2}$ are independent. Each stage uses Basu's Lemma.
Stage 1: $\left(\bar{x}, s^{2}\right)$ is independent of $Z$ : Consider the family of all $N\left(\mu, \sigma^{2}\right)$ distributions (with both parameters allowed to vary). Recall: $\left(\bar{x}, s^{2}\right)$ is a CSS. This is a location-scale family so that any location-scale invariant statistic is ancillary. $Z$ is location-scale invariant. Thus, $Z$ is ancillary so that (by Basu's Lemma) it must be independent of $\left(\bar{x}, s^{2}\right)$ for all $\left(\mu, \sigma^{2}\right)$.
Stage 2: $\bar{x}$ and $s^{2}$ are independent: Fix $\sigma^{2}$ at an arbitrary value $\sigma_{0}^{2}$ and consider the family of $\mathrm{N}\left(\mu, \sigma_{0}^{2}\right)$ distributions, $\mu$ unknown. Recall: This is a 1 pef and the natural SS $\sum_{i} X_{i}$ is a CSS. $\bar{x}$ is a 1-1 function of this and so also a CSS. This is a location family so that any location invariant statistic is ancillary. $s^{2}$ is location invariant. Thus, $s^{2}$ is ancillary so (by Basu's Lemma) it must be independent of $\bar{x}$ for all $\mu$ (and also for all $\sigma^{2}$ since $\sigma_{0}^{2}$ is arbitrary).
4. Example: (an amusing calculation via Basu) Let $X=\left(X_{1}, \ldots, X_{n}\right) \operatorname{iid} \mathrm{N}\left(0, \sigma^{2}\right), \sigma^{2}>$ 0.

Goal: Compute $E S$ where

$$
S=\frac{\left(\sum_{i} X_{i}\right)^{2}}{\sum_{i} X_{i}^{2}}
$$

This is a 1pef with $\theta=\sigma^{2}, t(x)=x^{2}$ and $w(\theta)=-1 /\left(2 \sigma^{2}\right)$. Therefore $T(X)=\sum_{i} X_{i}^{2}$ is CSS. This is also a scale family so that scale invariant statistics are ancillary. $S$ is scale invariant implies $S$ ancillary which implies (by Basu) $S$ independent of $T$. Thus $E(S T)=(E S)(E T)$ which implies $E S=E(S T) / E T$ which becomes

$$
E S=\frac{E\left(\sum_{i} X_{i}\right)^{2}}{E \sum_{i} X_{i}^{2}}=\frac{n \sigma^{2}}{n \sigma^{2}}=1 .
$$

