Solutions to Assignment #2

(6.3) Joint pdf of $X_1, \ldots, X_n$ is

$$\prod_{i=1}^n f(x_i|\mu,\sigma) = \prod_{i=1}^n \left\{ \left[\frac{1}{\sigma} e^{-(x_i-\mu)/\sigma}\right] I_{\{x_i>\mu\}} \right\}$$

$$= \frac{1}{\sigma^n} e^{-(\sum_i^n x_i - n\mu)/\sigma} \prod_{i=1}^n I_{\{x_i>\mu\}}$$

$$= \frac{1}{\sigma^n} e^{-\frac{\sum_1^n x_i}{\sigma} + \frac{n\mu}{\sigma}} I_{\{x_{(1)}>\mu\}},$$

since the event $\{x_1>\mu, x_2>\mu, \ldots, x_n>\mu\}$ occurs $\Leftrightarrow$ $\{x_{(1)}>\mu\}$.

By the factorization theorem, $(\sum_{i=1}^n X_i, X_{(1)})$ is sufficient for (μ,σ).

(6.9) Apply Theorem 6.2.13 in each case.

(a) Let $(x_1,\ldots,x_n)$ and $(y_1,\ldots,y_n)$ be two sample points. Then

$$\frac{\prod_{i=1}^n f(x_i|\theta)}{\prod_{i=1}^n f(y_i|\theta)} = \frac{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\sum_i^n (x_i-\theta)^2/2}}{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\sum_i^n (y_i-\theta)^2/2}}$$

$$= \exp\left\{\left(-\sum_{i=1}^n x_i^2 - 2\theta\sum_{i=1}^n x_i + n\theta^2 + \sum_{i=1}^n y_i^2 + 2\theta\sum_{i=1}^n y_i - n\theta^2\right)/2\right\}$$

$$= \exp\left\{\left(-\sum x_i^2 + \sum y_i^2\right)/2 - \left(\sum_{i=1}^n x_i - \sum_{i=1}^n y_i\right)\theta\right\}.$$

This is constant as a function of θ if and only if (continued)

6.9 (a) (continued) $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$.

Thus by Thm 6.2.13, $T(X_1, \ldots, X_n) = \sum_{i=1}^{n} X_i$ is minimal sufficient for θ.

<u>Note</u>: there are other (equivalent) minimal sufficient statistics. E.g. write $(\sum x_i - \sum y_i)\theta$ as $n(\bar{x} - \bar{y})\theta$ to see that $\bar{X} = \frac{\sum X_i}{n}$ is also minimal sufficient.

(b) Here we have
$$\frac{\prod_{i=1}^{n} f(x_i|\theta)}{\prod_{i=1}^{n} f(y_i|\theta)} = \frac{e^{-\sum(x_i-\theta)}\, I_{\{x_{(1)}>\theta\}}}{e^{-\sum(y_i-\theta)}\, I_{\{y_{(1)}>\theta\}}}$$

$$= \exp\left[-\sum x_i + \sum y_i + n\theta - n\theta\right] \frac{I_{\{x_{(1)}>\theta\}}}{I_{\{y_{(1)}>\theta\}}}$$

$$= e^{-\sum x_i + \sum y_i} \frac{I_{\{x_{(1)}>\theta\}}}{I_{\{y_{(1)}>\theta\}}}$$

With $\frac{0}{0}$ defined to be equal to 1, the only way that the ratio of the two indicator functions can be a constant value (namely 1) for <u>all</u> $\theta \in \mathbb{R}$ is to have $x_{(1)} = y_{(1)}$. It follows from Thm 6.2.13 that $T(X_1, \ldots, X_n) = X_{(1)}$ is minimal sufficient for θ.

6.9(c) $$\frac{\prod_{i=1}^n f(x_i|\theta)}{\prod_{i=1}^n f(y_i|\theta)} = \frac{e^{-\sum_1^n x_i + n\theta}}{\prod_{i=1}^n \left[1+e^{-(x_i-\theta)}\right]^2} \cdot \frac{\prod_{i=1}^n \left[1+e^{-(y_i-\theta)}\right]^2}{e^{-\sum y_i + n\theta}}$$

$$= e^{-\sum x_i + \sum y_i} \prod_{i=1}^n \left[\frac{1+e^{-(y_i-\theta)}}{1+e^{-(x_i-\theta)}}\right]^2.$$

The only way that this can be constant as a function of θ is for $(y_1, y_2, \ldots, y_n)$ to be some permutation of $(x_1, \ldots, x_n)$, in which case all the order statistics must coincide, i.e. $x_{(1)} = y_{(1)}, x_{(2)} = y_{(2)}, \ldots, x_{(n)} = y_{(n)}$. Thus the minimal sufficient statistic is the vector of order statistics $T(X_1, X_2, \ldots, X_n) = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$

Note: In any case where $X_1, \ldots, X_n$ are iid with joint pdf $\prod_{i=1}^n f(x_i|\theta)$, the order statistics are always sufficient for θ. If no further "reduction" is possible, then the order statistics are minimal sufficient.

6.9 (d) and (e): $T(X_1, \ldots, X_n) = (X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ is minimal sufficient, for the same reason as in 6.9(c).

6.10 $X_1, \ldots, X_n \overset{iid}{\sim} \text{Unif}\ (\theta, \theta+1)$.

It was shown in the text example that $T(X_1, \ldots, X_n) = (X_{(1)}, X_{(n)})$ is minimal sufficient. To show that T is not complete, we need only exhibit some non-vanishing function $g(X_{(1)}, X_{(n)})$ such that $E_\theta[g(X_{(1)}, X_{(n)}) = 0]$ for all $\theta \in \mathbb{R}$.

A natural way to go about finding such a function g is first to compute $E_\theta(X_{(1)})$ and $E_\theta[X_{(n)}]$ and then to play around with these expectations.

To compute $E_\theta[X_{(n)}]$, first set $U_i = X_i - \theta$, $i = 1, \ldots, n$, and note that the U_i's are iid Unif $[0,1]$. For $v \in (0,1)$, $P[U_{(n)} \leq v] = \prod_{i=1}^{n} P[U_i \leq v] = v^n$, so the p.d.f. of $U_{(n)}$ is $f(v) = nv^{n-1}$, $0 < v < 1$. So $E[U_{(n)}] =$ $= \int_0^1 v\,n v^{n-1}\,dv = n\int_0^1 v^n\,dv = \frac{n}{n+1}$. A similar calculation

(continued)

6.10 (continued)

(or arguing by symmetry) yields that $E(U_{(1)}) = \frac{1}{n+1}$.

Hence $E_\theta(X_{(n)}) = \theta + E(U_{(n)}) = \theta + \frac{n}{n+1}$

and $E_\theta(X_{(1)}) = \theta + \frac{1}{n+1}$.

This suggests knocking out θ by considering the difference of the two statistics. We note that

$$E_\theta\left[X_{(n)} - X_{(1)}\right] = \theta + \frac{n}{n+1} - \left(\theta + \frac{1}{n+1}\right) = \frac{n-1}{n+1}.$$

Hence $E_\theta\left[X_{(n)} - X_{(1)} - \frac{(n-1)}{n+1}\right] = 0$ for all θ.

Hence $g(X_{(1)}, X_{(n)}) = X_{(n)} - X_{(1)} - \frac{(n-1)}{n+1}$

is a non-vanishing function of the minimal sufficient statistic $(X_{(1)}, X_{(n)})$ whose expectation is 0 under all $\theta \in \mathbb{R}$. Hence, by the definition, $(X_{(1)}, X_{(n)})$ is not complete.

(6.13) $X_1, X_2 \overset{iid}{\sim}$ pdf $f(x|\alpha) = \alpha x^{\alpha-1} e^{-x^\alpha}$, $x>0$, $\alpha>0$

To show that $(\log X_1)/(\log X_2)$ is ancillary, we need to verify that its distribution does not depend on α.

First we work out the pdf of $Y_i = \log X_i$, $i=1,2$.

$y = \log x$ is a 1-1 map of $(0,\infty)$ onto $(-\infty,\infty)$, with inverse $x = e^y$. So Y_i has density

$$g(y) = f(e^y)\,|e^y| = \alpha e^{y(\alpha-1)} e^{-e^{y\alpha}} \cdot e^y$$
$$= \alpha e^{y\alpha} e^{-e^{y\alpha}}, \quad -\infty<y<\infty, \ \alpha>0.$$

We could work out the jt pdf of Y_1/Y_2 and some other variable, then integrate out the other variable to get the pdf of Y_1/Y_2. However the form of the pdf of Y_i suggests a simpler method — namely, make the further change of variable $Z_i = \alpha Y_i$, $i=1,2$. (so $Z_i = \alpha \log X_i$; $i=1,2$). Then the pdf of Z_i is $h(z) = g\left(\frac{z}{\alpha}\right)\cdot\frac{1}{\alpha} = \alpha e^z e^{-e^z}\frac{1}{\alpha} = e^z e^{-e^z}$, $-\infty<z<\infty$, $\alpha>0$

which does not depend on α.

(continued)

6.13 (continued)

So the joint pdf. of Z_1 and Z_2 does not depend on α, so the distr of any function of Z_1 and Z_2 does not depend on α.

So $\dfrac{\log X_1}{\log X_2}$ has the same distribution as $\dfrac{\alpha \log X_1}{\alpha \log X_2}$, which has the same distribution as that of $\dfrac{Z_1}{Z_2}$, which does not depend on α.

So $(\log X_1)/(\log X_2)$ is ancillary.

6.15 (a) The parameter space for the class of all normal distributions $N(\mu, \sigma^2)$ is $-\infty < \mu < \infty$, $\sigma^2 > 0$, i.e. the two-dimensional set $\mathbb{R} \times \mathbb{R}^+$. For the subclass $\{N(\theta, a\theta^2) : \theta > 0\}$, the corresponding parameter space is the half-line $\sigma^2 = a\mu$. This is a one-dimensional set, and so cannot contain a two-dimensional open set.

(b) The factorization in Example 6.2.9 (top of page 279) still holds, with (μ, σ^2) replaced by $(\theta, a\theta)$:

(continued)

6.15(b) (continued)

We have

$$f(\underset{\sim}{x}|\theta) = g(T_1(\underset{\sim}{x}), T_2(\underset{\sim}{x})|\theta)\, h(\underset{\sim}{x}),$$

when $T_1(\underset{\sim}{x}) = \bar{x}$ and $T_2(\underset{\sim}{x}) = s^2 = \sum_1^n (x_i - \bar{x})^2/(n-1)$,

where $h(\underset{\sim}{x}) = 1$ and

$$g(t_1, t_2|\theta) = (2\pi a\theta^2)^{-n/2} \exp(-n(t_1-\theta)^2 + (n-1)t_2)/(2a\theta^2))$$

By the Factorization Theorem $T = (\bar{X}, S^2)$ is sufficient for θ.

To show that the family of distributions of T is not complete, one could use the result of part (a) and apply Theorem 6.2.25 provided that the phrase "as long as" in the statement of the theorem is taken to mean "if and only if." But if "as long as" is taken to mean "provided that" (the safer interpretation), then we need to show non-completeness by exhibiting a non-vanishing function of $\bar{X}$ and S^2 whose expectation vanishes for all $\theta \in \Omega$. Since $E_\theta(\bar{X}) = \theta$, and $E_\theta(S^2) = a\theta^2$, we have to stop to think about how to find a function of X and S^2 whose expectation is constant. ~~We can exploit the statistical independence~~

~~(continued)~~

6.15(b) (continued)

Obviously, a linear combination of $\bar{X}$ and S^2 won't work.

So try $\bar{X}^2$: $E_\theta(\bar{X}^2) = Var(\bar{X}) + [E(\bar{X})]^2$

$$= \frac{\sigma^2}{n} + \mu^2 = \frac{a\theta^2}{n} + \theta^2 = \theta^2\left(\frac{a}{n}+1\right).$$

So $E_\theta\left[\frac{\bar{X}^2}{\frac{a}{n}+1}\right] = E_\theta\left[\frac{S^2}{a}\right] = \theta^2$ for all θ.

So define $g(\bar{X}, S^2) = \frac{\bar{X}^2}{\frac{a}{n}+1} - \frac{S^2}{a}$.

This is a non-vanishing function of $T = (\bar{X}, S^2)$ with the property that $E_\theta\, g(\bar{X}, S^2) = \theta^2 - \theta^2 = 0$ for all θ. Hence T is not complete.

(Note: there are a lot of other ways to choose $g(\bar{X}, S^2)$ that will work).

6.18

If $X_1, \ldots, X_n \overset{iid}{\sim}$ Poisson(λ). Then the distribution of $\sum_{i=1}^n X_i$ (using, e.g., moment-generating functions) is seen to be Poisson ($n\lambda$). That is,

$$P_\lambda\left(\sum_{i=1}^n X_i = x\right) = e^{-n\lambda}\frac{(n\lambda)^x}{x!}, \quad x = 0, 1, 2, \ldots$$

(continued)

6.18 (continued)

Write $T = \sum_{i=1}^{n} X_i$. Let $g(T)$ be any function of T such that $E_\lambda[g(T)] = 0$ for all $\lambda > 0$. To verify completeness, we need to show that this implies that $g(t) = 0$ for $t = 0, 1, 2, \ldots$

Now $E_\lambda g(T) = 0$ means that

$$\sum_{t=0}^{\infty} g(t)\, e^{-n\lambda} \frac{(n\lambda)^t}{t!} = 0 \quad \text{for } \lambda > 0.$$

Divide by the positive constant $e^{-n\lambda}$ to obtain

$$\sum_{t=0}^{\infty} g(t) n^t \frac{\lambda^t}{t!} = 0 \quad \text{for all } \lambda > 0.$$

However, it is well known that a power series $\sum_{t=0}^{\infty} a(t) \lambda^t$ which vanishes for all λ must have $a(t) = 0$ for $t = 0, 1, 2, \ldots$

Hence $g(t) \frac{n^t}{t!} = 0$ for $t = 0, 1, 2, \ldots$

Hence $g(t) = 0$, $t = 0, 1, 2, \ldots$

Thus T is complete.

6.19 Consider first Distribution 1. Let g be any function defined on $\{0,1,2\}$ such that $E_p[g(X)]=0$ for $0<p<1/4$.

That is, (*) $p\,g(0)+3p\,g(1)+(1-4p)\,g(2)=0$ for $0<p<1/4$.

If this implies $g(0)=g(1)=g(2)=0$, then the family of distribution of X is complete. Otherwise, — not complete.

Write (*) as

$$p\left[g(0)+3g(1)-4g(2)\right]+g(2)=0.$$

The only way this polynomial in p can vanish for $p\in(0,1/4)$ is for $g(0)+3g(1)-4g(2)=0$ and $g(2)=0$.

We thus have $g(2)=0$ and $g(0)+3g(1)=0$.

There are lots of non-zero solutions, e.g. $g(0)=3$, $g(1)=-1$ and $g(2)=0$. So this family is not complete.

Now consider Distribution 2: $E_p(X)=0$ for $p\in(0,1/2)$ means $p\,g(0)+p^2 g(1)+(1-p-p^2)\,g(2)=0$ for $p\in(0,1/2)$.

OR $p^2[g(1)-g(2)]+p[g(0)-g(2)]+g(2)=0$ for $p\in(0,1/2)$.

This implies $g(1)-g(2)=0$, $g(0)-g(2)=0$ and $g(2)=0$.

So $g(0)=g(1)=g(2)=0$. Hence the family is complete.

6.20 (a) $\prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \left[\frac{2x_i}{\theta^2} I_{\{x_i<\theta\}} \right]$

$$= \frac{2^n \prod_{i=1}^{n} x_i}{\theta^{2n}} I_{\{\max x_i < \theta\}}$$

By the Factorization Theorem, $X_{(n)} = \max_{i=1..n} X_i$ is sufficient for θ.

Let $T = X_{(n)}$. We need to check whether $E_\theta\, g(T) = 0$ for all $\theta > 0$ implies that $g(t) = 0$ for all $t > 0$. First we work out the distribution of T. Note that $P_\theta[X_1 \le x] = \frac{x^2}{\theta^2}$ for $0 \le x \le \theta$,

so $P_\theta[T \le x] = \{P_\theta[X_1 \le x]\}^n = \frac{x^{2n}}{\theta^{2n}}$, $0 \le x \le \theta$,

so the pdf of T is $f_\theta(x) = \frac{2n x^{2n-1}}{\theta^{2n}}$, $0 < x < \theta$.

So $E_\theta\, g(T) = 0$ for all $\theta > 0 \Rightarrow \int_0^\theta \frac{g(t)\, 2n\, t^{n-1}}{\theta^{2n}}\, dt = 0$

for all $\theta > 0 \Rightarrow \int_0^\theta 2n\, g(t)\, t^{n-1}\, dt = 0$ for all $\theta > 0$

$\Rightarrow \frac{d}{d\theta} \int_0^\theta 2n\, g(t)\, t^{n-1}\, dt = 0 \Rightarrow 2n\, g(\theta)\, \theta^{n-1} = 0$ for all $\theta > 0$

$\Rightarrow g(\theta) = 0$ for all $\theta > 0 \Rightarrow g(t) = 0$ for all $t > 0$.

Thus $T = X_{(n)}$ is complete.

6.20(b) Write in exponential family form:

$$f(x|\theta) = \theta\,(1+x)^{-(1+\theta)} = \theta\, e^{-(1+\theta)\log(1+x)}$$

$$= h(x)\, c(\theta)\, \exp[w(\theta)\, t(x)] \text{ , where}$$

$k=1$, $h(x)=1$, $c(\theta)=\theta$, $w(\theta) = -(1+\theta)$, $t(x) = \log(1+x)$.

By Theorem 6.2.10, $\sum_{j=1}^{n} t(X_j) = \sum_{j=1}^{n} \log(1+X_j)$ is sufficient for θ [and so is the equivalent statistic $\prod_{j=1}^{n}(1+X_j)$].

By Thm 6.2.10, since θ is one-dimensional and the value of k in the exponential family representation is 1, the sufficient statistic is complete.

6.20(c) Another one-dimensional exponential family:

$$f(x|\theta) = \frac{\log\theta}{\theta-1}\, e^{x\log\theta} \quad , \; 0 < x < 1, \quad \theta > 1$$

Arguing as above, $\sum_{j=1}^{n} T_1(X_j) = \sum_{j=1}^{n} X_j$ is sufficient and complete

6.20(d) $f(x|\theta) = e^{-x}\, e^{\theta}\, \exp(e^{\theta} e^{-x})$

$$= h(x)\; c(\theta)\, \exp[w(\theta)\, t_1(x))$$

So arguing as in 6.20(b), $\sum_{j=1}^{n} t_1(X_j) = \sum_{j=1}^{n} e^{-X_j}$ is complete and sufficient.

6.20(c). This is just the binomial (n, θ) distr. with $n=2$ & already worked out in class that $\sum_{i=1}^{n} X_i$ is sufficient and complete by direct calculations. You can also apply the "exponential family" theorems: write

$$f(x|\theta) = \binom{2}{x}\theta^x(1-\theta)^{2-x} = \binom{2}{x}(1-\theta)^2\left(\frac{\theta}{1-\theta}\right)^x$$

$$= \binom{2}{x}(1-\theta)^2 e^{x\log\left(\frac{\theta}{1-\theta}\right)}, \quad 0 \le \theta \le 1, \quad x = 0, \ldots, n.$$

exponential family with $k=1$ and $t_1(x) = x$.

So, by the 2 theorems, $\sum_{j=1}^{n} t_1(X_j) = \sum_{j=1}^{n} X_j$ is sufficient and complete

(6.23) Use the Factorization Theorem to find a sufficient statistic.

$$\prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n}\left[\frac{1}{\theta} I_{\{x_i > \theta\}} \cdot I_{\{x_i < 2\theta\}}\right]$$

$$= \frac{1}{\theta^n}\prod_{i=1}^{n}\left(I_{\{x_i > \theta\}}\right) \cdot \prod_{i=1}^{n}\left[I_{\{x_i < 2\theta\}}\right]$$

$$= \frac{1}{\theta^n} I_{\{x_{(1)} > \theta\}} I_{\{x_{(n)} < 2\theta\}}$$

By Factorization Theorem, $(X_{(1)}, X_{(n)})$ is sufficient for θ.

(continued)

6.23 (continued)

To show that the sufficient statistic $(X_{(1)}, X_{(n)})$ is minimal sufficient, note that

$$\frac{\prod_{i=1}^{n} f(x_i|\theta)}{\prod_{i=1}^{n} f(y_i|\theta)} = \frac{\frac{1}{\theta^n} I_{\{x_{(1)}>\theta\}} I_{\{x_{(n)}<2\theta\}}}{\frac{1}{\theta^n} I_{\{y_{(1)}>\theta\}} I_{\{y_{(n)}<2\theta\}}} = \frac{I_{\{x_{(1)}>\theta\}} \cdot I_{\{x_{(n)}<2\theta\}}}{I_{\{y_{(1)}>\theta\}} I_{\{y_{(n)}<2\theta\}}}$$

is constant as a function of $\theta > 0$ only when $(x_{(1)}, x_{(n)}) = (y_{(1)}, y_{(n)})$. Thus, $(X_{(1)}, X_{(n)})$ is minimal sufficient.

Noting that $U_i = \frac{X_i - \theta}{\theta}$ are iid Unif$(0,1)$, we argue as in problem 6.10 to work out the expected values of $X_{(1)}$ and $X_{(n)}$:

$$E(X_{(1)}) = \theta + \frac{1}{n+1}\theta = \frac{n+2}{n+1}\theta$$

and

$$E(X_{(n)}) = 2\theta - \frac{1}{n+1}\theta = \frac{2n+1}{n+1}\theta.$$

So $E_\theta\left[\frac{n+1}{2n+1} X_{(n)} - \frac{n+1}{n+2} X_{(1)}\right] = \theta - \theta = 0$ for all $\theta > 0$.

Since $g(X_{(1)}, X_{(n)}) = \frac{n+1}{2n+1} X_{(n)} - \frac{n+1}{n+2} X_{(1)}$ is a non-vanishing function whose expected value is 0 for all θ it follows that $(X_{(1)}, X_{(n)})$ is not complete.

6.30 (a) $\prod_{i=1}^{n} f(x_i|\mu) = \prod_{i=1}^{n} \left[e^{-(x_i-\mu)} I_{\{x_i>\mu\}} \right]$

$$= e^{-\Sigma x_i} e^{n\mu} I_{\{x_{(1)}>\mu\}}$$

so by the Factorization, $X_{(1)}$ is sufficient.

To show completeness, we need to find the distribution of $X_{(1)}$.

Note first $P[X_i > x] = \int_x^\infty e^{-(t-\mu)} dt = e^{-(x-\mu)}$ for $x > \mu$.

So $P[X_{(1)} > x] = \left[P[X_1 > x]\right]^n = e^{-n(x-\mu)}$, $x > \mu$.

So the pdf of $X_{(1)}$ is $n e^{-n(x-\mu)}$, $x > \mu$.

Suppose $E_\mu g(X_{(1)}) = 0$ for all μ.

Then $\int_\mu^\infty g(x) n e^{-nx} e^{n\mu} dx = 0$ for all μ,

or $\int_\mu^\infty g(x) e^{-nx} dx = 0$ for all μ.

So $\frac{d}{d\mu} \int_\mu^\infty g(x) e^{-nx} dx = -g(\mu) e^{-n\mu} = 0$ for all μ.

Since $e^{-n\mu} > 0$ for all μ, we must have $g(\mu) = 0$

for all $\mu \in \mathbb{R}$. Thus $g(x) = 0$ for all $x \in \mathbb{R}$.

It follows that $X_{(1)}$ is complete.

(continued)

6.30(b)

Since $X_{(1)}$ is sufficient for μ, and complete, independence of $X_{(1)}$ and S^2 will follow once we show that S^2 is ancillary — that is, the distribution of S^2 does not depend on μ. The hard way to show this is to try to work out the distribution of $S^2 = \frac{\sum(X_i-\bar{X})^2}{n-1}$ when $X_1, \ldots, X_n \overset{iid}{\sim} f(x_i|\mu) = e^{-(x_i-\mu)}$. The easy way avoids the hard work by using the fact that the family of distr. of the X_i's is a location family $f_0(x_i-\mu)$ for some fixed f_0, $-\infty<\mu<\infty$. Write $X_i = \mu + Y_i$, $i=1,\ldots,n$, where the Y_i's are iid with pdf $f_0(y_i)$. Then the distr. of $S^2 = \frac{\sum(X_i-\bar{X})^2}{n-1} = \frac{\sum[(X_i-\mu)-(\bar{X}-\mu)]^2}{n-1} = \frac{\sum(Y_i-\bar{Y})^2}{n-1}$ clearly does not depend on μ. So S^2 is ancillary, hence indep of $X_{(1)}$ by Basu's Theorem.

Note: this argument also works if S^2 is replaced by any statistic that depends on $(X_1, \ldots, X_n)$ only through $(X_1-\bar{X}, \ldots, X_n-\bar{X})$.