The 2nd method of deriving a UMVUE when a sufficient and complete statistic is available

- Find an unbiased estimator of $\vartheta$, say $U(X)$.
- Conditioning on a sufficient and complete statistic $T(X)$: $E[U(X) | T]$ is the UMVUE of $\vartheta$.
- We need to derive an explicit form of $E[U(X) | T]$.
- We do not need the distribution of $T$.
  But we need to work out the conditional expectation $E[U(X) | T]$.
- From the uniqueness of the UMVUE, it does not matter which $U(X)$ is used.
  Thus, we should choose $U(X)$ so as to make the calculation of $E[U(X) | T]$ as easy as possible.
Example 3.3

Let $X_1, \ldots, X_n$ be i.i.d. from the exponential distribution $E(0, \theta)$. 
$F_\theta(x) = (1 - e^{-x/\theta})I_{(0, \infty)}(x)$.

Consider the estimation of $\vartheta = 1 - F_\theta(t)$.

$\bar{X}$ is sufficient and complete for $\theta > 0$.

$I_{(t, \infty)}(X_1)$ is unbiased for $\vartheta$,

$$E[I_{(t, \infty)}(X_1)] = P(X_1 > t) = \vartheta.$$

Hence

$$T(X) = E[I_{(t, \infty)}(X_1)|\bar{X}] = P(X_1 > t|\bar{X})$$

is the UMVUE of $\vartheta$.

If the conditional distribution of $X_1$ given $\bar{X}$ is available, then we can calculate $P(X_1 > t|\bar{X})$ directly.

By Basu’s theorem (Theorem 2.4), $X_1/\bar{X}$ and $\bar{X}$ are independent.

By Proposition 1.10(vii),

$$P(X_1 > t|\bar{X} = \bar{x}) = P(X_1/\bar{X} > t/\bar{x}|\bar{X} = \bar{x}) = P(X_1/\bar{X} > t/\bar{x}).$$
To compute this unconditional probability, we need the distribution of

$$X_1 / \sum_{i=1}^{n} X_i = X_1 / \left( X_1 + \sum_{i=2}^{n} X_i \right).$$

Using the transformation technique discussed in §1.3.1 and the fact that $\sum_{i=2}^{n} X_i$ is independent of $X_1$ and has a gamma distribution, we obtain that $X_1 / \sum_{i=1}^{n} X_i$ has the Lebesgue p.d.f.

$$(n - 1)(1 - x)^{n-2} I_{(0,1)}(x).$$

Hence

$$P(X_1 > t | \bar{X} = \bar{x}) = (n - 1) \int_{t/(n\bar{x})}^{1} (1 - x)^{n-2} dx = \left( 1 - \frac{t}{n\bar{x}} \right)^{n-1}$$

and the UMVUE of $\vartheta$ is

$$T(X) = \left( 1 - \frac{t}{n\bar{x}} \right)^{n-1}.$$
Example 3.4

Let $X_1, \ldots, X_n$ be i.i.d. from $N(\mu, \sigma^2)$ with unknown $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. From Example 2.18, $T = (\bar{X}, S^2)$ is sufficient and complete for $\theta = (\mu, \sigma^2)$.

$\bar{X}$ and $(n - 1)S^2/\sigma^2$ are independent.

$\bar{X}$ has the $N(\mu, \sigma^2/n)$ distribution.

$S^2$ has the chi-square distribution $\chi^2_{n-1}$.

Using the method of solving for $h$ directly, we find that

- the UMVUE for $\mu$ is $\bar{X}$;
- the UMVUE of $\mu^2$ is $\bar{X}^2 - S^2/n$;
- the UMVUE for $\sigma^r$ with $r > 1 - n$ is $k_{n-1,r}S^r$, where

$$k_{n,r} = \frac{n^{r/2}\Gamma\left(\frac{n}{2}\right)}{2^{r/2}\Gamma\left(\frac{n+r}{2}\right)}$$

- the UMVUE of $\mu/\sigma$ is $k_{n-1,-1}\bar{X}/S$, if $n > 2$. 
Example 3.4 (continued)

Suppose that $\vartheta$ satisfies $P(X_1 \leq \vartheta) = p$ with a fixed $p \in (0, 1)$. Let $\Phi$ be the c.d.f. of the standard normal distribution. Then

$$\vartheta = \mu + \sigma \Phi^{-1}(p)$$

and its UMVUE is

$$\bar{X} + k_{n-1.1} S \Phi^{-1}(p).$$

Let $c$ be a fixed constant and

$$\vartheta = P(X_1 \leq c) = \Phi \left( \frac{c - \mu}{\sigma} \right).$$

We can find the UMVUE of $\vartheta$ using the method of conditioning. Since $I_{(-\infty, c)}(X_1)$ is an unbiased estimator of $\vartheta$, the UMVUE of $\vartheta$ is

$$E[I_{(-\infty, c)}(X_1) \mid T] = P(X_1 \leq c \mid T).$$

By Basu’s theorem, the ancillary statistic $Z(X) = (X_1 - \bar{X})/S$ is independent of $T = (\bar{X}, S^2)$. 

Example 3.4 (continued)

Then, by Proposition 1.10(vii),

\[ P \left( X_1 \leq c \mid T = (\bar{x}, s^2) \right) = P \left( Z \leq \frac{c - \bar{X}}{S} \mid T = (\bar{x}, s^2) \right) \]

\[ = P \left( Z \leq \frac{c - \bar{x}}{S} \right). \]

It can be shown that \( Z \) has the Lebesgue p.d.f.

\[ f(z) = \frac{\sqrt{n}\Gamma \left( \frac{n-1}{2} \right)}{\sqrt{\pi(n-1)}\Gamma \left( \frac{n-2}{2} \right)} \left[ 1 - \frac{n z^2}{(n-1)^2} \right]^{(n/2)-2} I_{(0,(n-1)/\sqrt{n})}(|z|) \]

Hence the UMVUE of \( \vartheta \) is

\[ P(X_1 \leq c \mid T) = \int_{-(n-1)/\sqrt{n}}^{(c-\bar{x})/S} f(z) \, dz \]
Example 3.4 (continued)

Suppose that we would like to estimate

\[ \vartheta = \frac{1}{\sigma} \Phi' \left( \frac{c - \mu}{\sigma} \right), \]

the Lebesgue p.d.f. of \( X_1 \) evaluated at a fixed \( c \), where \( \Phi' \) is the first-order derivative of \( \Phi \).

By the previous result, the conditional p.d.f. of \( X_1 \) given \( \bar{X} = \bar{x} \) and \( S^2 = s^2 \) is

\[ s^{-1} f \left( \frac{x - \bar{x}}{s} \right). \]

Let \( f_T \) be the joint p.d.f. of \( T = (\bar{X}, S^2) \).

Then

\[ \vartheta = \int \int \frac{1}{s} f \left( \frac{c - \bar{x}}{s} \right) f_T(t)dt = E \left[ \frac{1}{S} f \left( \frac{c - \bar{X}}{S} \right) \right]. \]

Hence the UMVUE of \( \vartheta \) is

\[ \frac{1}{S} f \left( \frac{c - \bar{X}}{S} \right). \]
Example

Let $X_1, \ldots, X_n$ be i.i.d. with Lebesgue p.d.f. $f_\theta(x) = \theta x^{-2}I_{(\theta, \infty)}(x)$, where $\theta > 0$ is unknown.

Suppose that $\vartheta = P(X_1 > t)$ for a constant $t > 0$.

The smallest order statistic $X_{(1)}$ is sufficient and complete for $\theta$.

Hence, the UMVUE of $\vartheta$ is

$$P(X_1 > t | X_{(1)}) = P(X_1 > t | X_{(1)} = x_{(1)})$$

$$= P\left(\frac{X_1}{X_{(1)}} > \frac{t}{x_{(1)}} \bigg| X_{(1)} = x_{(1)}\right)$$

$$= P\left(\frac{X_1}{X_{(1)}} > \frac{t}{x_{(1)}} \bigg| X_{(1)} = x_{(1)}\right)$$

$$= P\left(\frac{X_1}{X_{(1)}} > \frac{t}{x_{(1)}} \bigg| X_{(1)} = x_{(1)}\right)$$

$$= P\left(\frac{X_1}{X_{(1)}} > s\right)$$

(Basu’s theorem), where $s = t/x_{(1)}$.

If $s \leq 1$, this probability is 1.
Example (continued)

Consider $s > 1$ and assume $\theta = 1$ in the calculation:

\[
P\left(\frac{X_1}{X(1)} > s\right) = \sum_{i=1}^{n} P\left(\frac{X_1}{X(1)} > s, X(1) = X_i\right)
\]

\[
= \sum_{i=2}^{n} P\left(\frac{X_1}{X(1)} > s, X(1) = X_i\right)
\]

\[
= (n - 1) P\left(\frac{X_1}{X(1)} > s, X(1) = X_n\right)
\]

\[
= (n - 1) P(X_1 > sX_n, X_2 > X_n, \ldots, X_{n-1} > X_n)
\]

\[
= (n - 1) \int_{x_1 > sx_n, x_2 > x_n, \ldots, x_{n-1} > x_n} \prod_{i=1}^{n} \frac{1}{x_i^2} dx_1 \cdots dx_n
\]

\[
= (n - 1) \int_{1}^{\infty} \left[ \int_{s x_n}^{\infty} \prod_{i=2}^{n-1} \left( \int_{x_n}^{\infty} \frac{1}{x_i^2} dx_i \right) \frac{1}{x_n^2} \right] \frac{1}{x_n^2} dx_n
\]

\[
= (n - 1) \int_{1}^{\infty} \frac{1}{sx_n^{n+1}} dx_n = \frac{(n - 1)x(1)}{nt}
\]
Example (continued)

This shows that the UMVUE of $P(X_1 > t)$ is

$$ h(X_{(1)}) = \begin{cases} 
\frac{(n-1)X_{(1)}}{nt} & X_{(1)} < t \\
1 & X_{(1)} \geq t 
\end{cases} $$

Another solution

The UMVUE must be $h(X_{(1)})$

The Lebesgue p.d.f. of $X_{(1)}$ is

$$ n\theta^n \frac{1}{x^{n+1}} I_{(\theta, \infty)}(x). $$

Use the method of finding $h$

If $\theta \geq t$, then $P(X_1 > t) = 1$ and $P(t > X_{(1)}) = 0$.

Hence, if $X_{(1)} \geq t$, $h(X_{(1)})$ must be 1 a.s. $P_\theta$

The value of $h(X_{(1)})$ for $X_{(1)} < t$ is not specified.
If $\theta < t$,

$$
E[h(X(1))] = \int_{\theta}^{\infty} h(x) \frac{n\theta^n}{x^{n+1}} \, dx
$$

$$
= \int_{\theta}^{t} h(x) \frac{n\theta^n}{x^{n+1}} \, dx + \int_{t}^{\infty} \frac{n\theta^n}{x^{n+1}} \, dx = \int_{\theta}^{t} h(x) \frac{n\theta^n}{x^{n+1}} \, dx + \frac{\theta^n}{t^n}
$$

Since $P(X_1 > t) = \theta / t$, we have

$$
\frac{\theta}{t} = \int_{\theta}^{t} h(x) \frac{n\theta^n}{x^{n+1}} \, dx + \frac{\theta^n}{t^n}
$$

i.e.,

$$
\frac{1}{t^{\theta n-1}} = \int_{\theta}^{t} h(x) \frac{n}{x^{n+1}} \, dx + \frac{1}{t^n}
$$

Differentiating both sizes w.r.t. $\theta$ leads to

$$
- \frac{n-1}{t\theta^n} = -h(\theta) \frac{n}{\theta^{n+1}}
$$

Hence, for any $X(1) < t$,

$$
h(X(1)) = \frac{(n-1)X(1)}{nt}.
$$