Multivariate Analysis and Its Applications
IMS Lecture Notes – Monograph Series (1994) Volume 24

THE VIRTUAL SYSTEM METHOD FOR ESTIMATION OF PARAMETER IN SYSTEM TREE*

BY ZHENG ZHONGGUO

Peking University

A machine, or other type of "system", can often be divided into several subsystems and these subsystems again can be divided into several subsystems (second generation), $\cdots$. This process forms a "system tree". Suppose that the distributions of the life spans of subsystems in the system tree are exponential distributions. To estimate the parameter of the distribution of life span of the equipment (the entire "system tree") based on data collected from subsystems, the virtual system method, an alternative to the obvious ML method, is presented in this paper. It is proved that the series of estimators constructed by the virtual system method is asymptotically efficient and that the calculation of the estimator is quite simple while the likelihood equation of the system tree is complicated.

1. Introduction and Main Result.

In practice, an equipment (System) is usually divided into several subsystems and these subsystems again can be divided into several subsystems, $\cdots$. Finally a system tree is formed. In this paper a system tree is denoted by a finite set of indices $M = \{m = (i_1, \cdots, i_k)\}$ satisfying

(i) m is a finite series of natural numbers,

(ii) $m = (i_1, \cdots, i_k) \in M \Longrightarrow i_1 = 1$,

(iii) $(i_1, \cdots, i_k) \in M \Longrightarrow (i_1, \cdots, i_{k-1}) \in M, (i_1, \cdots, i_{k-1}, j) \in M, \ j = 1, \cdots, i_{k-1}$.

An example of the system tree is given in Figure 1. Later on, we also call m ($\in M$) a subsystem.

---

* This work was supported by NSFC and Doctoral Program Foundation of Institute of Higher Education.

AMS 1980 Subject Classifications: Primary 62F12, Secondary 90B25.

Key words and phrases: Asymptotical efficiencies, system tree and Virtual System Method.

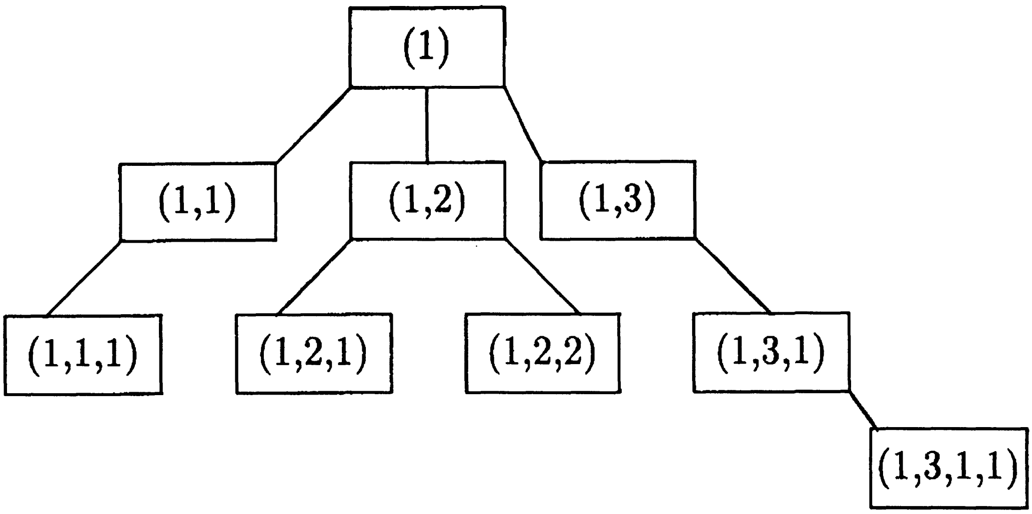

Figure 1. Example of system tree

DEFINITION 1.1. *Let $\widetilde{m}$, $m \in M$. $\widetilde{m}$ is said to be a subsystem of m, if $m = (i_1, \cdots, i_k)$ and $\widetilde{m} = (i_1, \cdots, i_k, \cdots, i_l)$. $\widetilde{m}$ is said to be the first generation subsystem of m if $m = (i_1, \cdots, i_k)$ and $\widetilde{m} = (i_1, \cdots, i_k, i_{k+1})$. m is said to be the last generation subsystem of the system tree M if no subsystem in M is a first generation subsystem of m.*

In Figure 1, (1,1,1), (1,2,1), (1,2,2), (1,3,1,1) are the last generation subsystems of the system tree. Denote

$$M_0 = \{m : m \text{ is the last generation subsystem of } M\}, \tag{1.1}$$

$$M(m) = \{\widetilde{m} : \widetilde{m} \text{ is the first generation subsystem of } m\}. \tag{1.2}$$

In [1], the lower confidence limit of the reliability of the system tree is constructed based on binomial trials. In this paper we suppose that every subsystem m in M has its own life span X_m.

DEFINITION 1.2. *$\{X_m : m \in M\}$ is said to be an exponential system if*

(i) The marginal distribution of X_m is exponential, i.e.,

$$P\{X_m \geq x\} = \exp\left\{-\frac{x}{\theta_m}\right\}, \qquad m \in M, \quad x > 0, \tag{1.3}$$

(ii) For every $m \in M \backslash M_0$, the parameter θ_m depends on $\{\theta_{\widetilde{m}}, \widetilde{m} \in M(m)\}$, i.e.,

$$\theta_m = \theta_m\big(\theta_{\widetilde{m}}, \widetilde{m} \in M(m)\big), \tag{1.4}$$

where the symbol $\theta_m(\theta_{\widetilde{m}}, \widetilde{m} \in M(m))$ means that θ_m is a function of the arguments $\{\theta_{\widetilde{m}}, \widetilde{m} \in M(m)\}$.

EXAMPLE 1.1. Let the subsystems be connected in series, i.e.,

$$X_m = \min_{\tilde{m} \in M(m)} \{X_{\tilde{m}}\} \tag{1.5}$$

and the marginal distributions of X_m, $m \in M$, are exponential distributions with parameters θ_m, $m \in M$. The parameters satisfy the following relation

$$\theta_m^{-1} = \sum_{\tilde{m} \in M(m)} \theta_{\tilde{m}}^{-1}. \tag{1.6}$$

According to Definition 1.2, $\{X_m,\ m \in M\}$ is an exponential system.

It is easy to show that the set of independent parameters of an exponential system is $\{\theta_m,\ m \in M_0\}$. In reliability engineering, people are interested in estimating the parameter of the machine (the first generation $m = (1)$ in the system tree) from the lifespans of machines and the lifespans of the components as well. All these lifespans are independently observed. Suppose that $\{x_{mi},\ m \in M,\ i = 1, \cdots, n_m\}$ are independently observed data, i.e., the joint density of $\{x_{mi},\ m \in M,\ i = 1, \cdots, n_m\}$ is in the form of

$$\prod_{m \in M} \theta_m^{-n_m} \exp\left\{ -\sum_{i=1}^{n_m} x_{m_i} \Big/ \theta_m \right\} = \prod_{m \in M} \theta_m^{-n_m} \exp\{ -T_m/\theta_m\}, \tag{1.7}$$

where

$$T_m \triangleq \sum_{i=1}^{n_m} x_{m_i}. \tag{1.8}$$

Let $\eta^\tau \hat{=} (\theta_{m_1}, \cdots, \theta_{m_k})$ be the independent parameters of the distribution family (1.7), where $\{m_1, \cdots, m_k\} = M_0$. The information matrix of $\{x_{m_i},\ m \in M,\ i = 1, \cdots, n_m\}$ is

$$I_\eta = \sum_{m \in M} \frac{n_m}{\theta_m^2} \frac{\partial \theta_m}{\partial \eta} \frac{\partial \theta_m}{\partial \eta^\tau}. \tag{1.9}$$

For the estimator of the parameter $\theta_{(1)}$, the Cramer lower bound of the variance of estimators is

$$\frac{\partial \theta_{(1)}}{\partial \eta^\tau} I_\eta^{-1} \frac{\partial \theta_{(1)}}{\partial \eta} \tag{1.10}$$

where $\theta_{(1)}$ is a composite function of η and $\frac{\partial \theta_{(1)}}{\partial \eta^\tau}$ is a row vector of the partial derivatives. In this paper we create a virtual system so that we can calculate the Cramer lower bound of variance for estimators of $\theta_{(1)}$ in the virtual system instead of the original one.

Let m_0 be a subsystem in M such that $M(m_0) \subset M_0$. Denote

$$M^{(1)} = M \backslash M(m_0). \tag{1.11}$$

Without loss of generality, suppose that $M_0=\{m_1,\cdots,m_k\}$ and $M(m_0) = \{m_{l+1},\cdots,m_k\}$. Let

$$M_0^{(1)} = \{m_1, \cdots, m_l, m_0\}, \tag{1.12}$$

$$\eta^{(1)^\tau} = (\theta_{m_1}, \cdots, \theta_{m_l}, \theta_{m_0}). \tag{1.13}$$

It is easy to show that $M^{(1)}$ is a new system tree. Define

$$n_m^{(1)} = \begin{cases} n_m, & m \in M^{(1)} \backslash m_0, \\ n_m + \dfrac{\theta_m^2}{\sum\limits_{\tilde{m}\in M(m)} \left(\dfrac{\partial\theta_m}{\partial\theta_{\tilde{m}}}\right)^2 \dfrac{\theta_{\tilde{m}}^2}{n_{\tilde{m}}}}, & m = m_0. \end{cases} \tag{1.14}$$

For the virtual system tree $M^{(1)}, \eta^{(1)}$ is the vector of independent parameters of the system. Define

$$I_{\eta^{(1)}} = \sum_{m\in M^{(1)}} \frac{n_m^{(1)}}{\theta_m^2} \frac{\partial\theta_m}{\partial\eta^{(1)}} \frac{\partial\theta_m}{\partial\eta^{(1)^\tau}} \tag{1.15}$$

as the information matrix of the virtual system $M^{(1)}$. The Cramer lower bound of variance of the estimator of $\theta_{(1)}$ in the new system is

$$\frac{\partial\theta_{(1)}}{\partial\eta^{(1)^\tau}} I_{\eta^{(1)}}^{-1} \frac{\partial\theta_{(1)}}{\partial\eta^{(1)}}. \tag{1.16}$$

For the two bounds (1.16) and (1.10), we have

THEOREM 1.1. *For the system M and its virtual system $M^{(1)}$, the following equality*

$$\frac{\partial\theta_{(1)}}{\partial\eta^\tau} I_\eta^{-1} \frac{\partial\theta_{(1)}}{\partial\eta} = \frac{\partial\theta_{(1)}}{\partial\eta^{(1)^\tau}} I_{\eta^{(1)}}^{-1} \frac{\partial\theta_{(1)}}{\partial\eta^{(1)}}. \tag{1.17}$$

holds in the sense that when we substitute $\eta^{(1)} = \eta^{(1)}(\eta)$ into the right hand side of (1.17), the two sides become identically equal.

REMARK 1.1. For the virtual system $M^{(1)}$, again we can construct a new system $M^{(2)}$ and by using Theorem 1.1, we obtain

$$\frac{\partial\theta_{(1)}}{\partial\eta^\tau} I_\eta^{-1} \frac{\partial\theta_{(1)}}{\partial\eta} = \frac{\partial\theta_{(1)}}{\partial\eta^{(2)^\tau}} I_{\eta^{(2)}}^{-1} \frac{\partial\theta_{(1)}}{\partial\eta^{(2)}}.$$

Continuing this procedure, finally we obtain $M^{(k_M)} = \{(1)\}$ with only one system as its member, where k_M is an integer. For the virtual system tree

$M^{(k_M)} = \{(1)\}, \eta^{(k_M)} = (\theta_{(1)}), I_{(\eta)}(k_M) = \frac{n_{(1)}^{(k_M)}}{\theta_{(1)}^2}$. By repeatedly using Theorem 1.1, we obtain

$$\frac{\partial\theta_{(1)}}{\partial\eta^\tau} I_\eta^{-1} \frac{\partial\theta_{(1)}}{\partial\eta} = \frac{\theta_{(1)}^2}{n_{(1)}^{(k_M)}}. \tag{1.18}$$

Before giving the definition of $\hat{\theta}_{(1)}$, we first introduce the definition of the virtual sample size N_m for every $m \in M$. When $m \in M_0$, N_m is defined by

$$N_m = n_m,$$

where n_m is the real sample size of the subsystem m. Suppose that $m \notin M_0$ and that for every $\widetilde{m} \in M(m)$, $N_{\widetilde{m}}$ is already defined, then N_m is defined by

$$N_m = n_m + \widetilde{n}_m,$$

$$\widetilde{n}_m = \frac{\theta_m^2}{\sum_{\widetilde{m}\in M(m)} \left(\frac{\partial\theta_m}{\partial\theta_{\widetilde{m}}}\right)^2 \frac{\theta_{\widetilde{m}}^2}{N_{\widetilde{m}}}}, \tag{1.19}$$

where θ_m is a known function of its arguments $\{\theta_{\widetilde{m}}, \widetilde{m} \in M(m)\}$. It is easy to know that for $m \notin M_0$, N_m and $\widetilde{n}_m$ are defined recursively and that N_m and $\widetilde{n}_m$ both are function of $(\theta_{\widetilde{m}}, N_{\widetilde{m}}, \widetilde{m} \in M(m))$. We denote them by $N_m(\theta_{\widetilde{m}}, N_{\widetilde{m}}, \widetilde{m} \in M(m))$ and $\widetilde{n}_m(\theta_{\widetilde{m}}, N_{\widetilde{m}}, \widetilde{m} \in M(m))$. The main interest of this paper is to estimate the parameter $\theta_{(1)}$ of the top subsystem $m = (1)$ (the parameter of the machine itself). The estimator $\hat{\theta}_{(1)}$ of $\theta_{(1)}$, and the related quantities $\hat{T}_m$ and $\hat{N}_m$ are defined recursively through the following steps.

(i) For $m \in M_0$, they are defined by

$$\hat{\theta}_m = \frac{\hat{T}_m}{\hat{N}_m}, \tag{1.20}$$

$$\hat{T}_m = T_m, \tag{1.21}$$

$$\hat{N}_m = n_m, \tag{1.22}$$

where T_m is the total time of test for the system m, n_m is the sample size of the subsystem m.

(ii) $m \notin M_0$. Suppose that for every $\widetilde{m} \in M(m)$, $\hat{T}_{\widetilde{m}}$, $\hat{N}_{\widetilde{m}}$ and $\hat{\theta}_{\widetilde{m}}$ have been defined. Then for the subsystem m, the corresponding estimators are defined by

$$\hat{\theta}_m = \frac{\hat{T}_m}{\hat{N}_m}, \tag{1.23}$$

$$\hat{T}_m = T_m + \hat{\widetilde{n}}_m \widetilde{\theta}_m, \tag{1.24}$$

$$\hat{N}_m = n_m + \hat{\tilde{n}}_m, \tag{1.25}$$

where

$$\tilde{\theta} \hat{=} \theta_m(\hat{\theta}_{\tilde{m}}, \tilde{m} \in M(m)), \tag{1.26}$$

$$\hat{\tilde{n}}_m \hat{=} \tilde{n}_m(\hat{\theta}_{\tilde{m}}, \hat{N}_m, \tilde{m} \in M(m)). \tag{2.27}$$

For the estimator $\widehat{\theta}_{(1)}$ of $\theta_{(1)}$,we obtain

THEOREM 1.2. *Suppose that for every θ_m, as a function of $(\theta_{\tilde{m}}, \tilde{m} \in M(m))$, $(m \notin M_0)$, the following holds*

$$\left(\frac{\partial \theta_m}{\partial \theta_{\tilde{m}}}, \ \tilde{m} \in M(m)\right) \neq 0.$$

Then, for the estimator $\widehat{\theta}_{(1)}$ of $\theta_{(1)}$ given by (1.23), we have

$$\left(\frac{\partial \theta_{(1)}}{\partial \eta^\tau} I_\eta^{-1} \frac{\partial \theta_{(1)}}{\partial \eta}\right)^{-1/2} \left(\widehat{\theta} - \theta_{(1)}\right) \xrightarrow{d} N(0,1). \tag{1.28}$$

REMARK 1.2. It is well known that it is very difficult to solve the likelihood equation to obtain the ML estimator of $\theta_{(1)}$. $\widehat{\theta}_{(1)}$ is an alternative estimator for $\theta_{(1)}$ which is also asymptotically efficient.

Suppose that $x_1, x_2, \cdots, x_n \sim$ iid $\frac{1}{\theta} e^{-\frac{x}{\theta}}, \theta > 0$. It is well known that the lower confidence limit with level 1 - α is given by

$$\underline{\theta} = \frac{T}{\Gamma_n^{-1}(1-\alpha)} \tag{1.29}$$

where $T = \sum_{i=1}^n x_i$ and $\Gamma_n^{-1}(1-\alpha)$ is the $(1-\alpha)$ quantile of the Γ distribution with parameter n, i.e. $\Gamma_n^{-1}(1-\alpha)$ is the solution of the equation

$$\frac{1}{\Gamma(n)} \int_0^x u^{n-1} e^{-u} \, du = 1 - \alpha.$$

For the exponential system $\{x_m, m \in M\}$ with data $\{ x_{m_i}, m \in M, i = 1, 2, \cdots, n_m \}$, we would like to construct a lower confidence limit of the parameter $\theta_{(1)}$. We consider $(\widehat{N}_{(1)}, \widehat{T}_{(1)})$ as a virtual exponential system where $\widehat{N}_{(1)}$ is the sample size of the virtual system and $\widehat{T}_{(1)}$ is the total experiment time of the system. As in (1.29), we use

$$\underline{\theta}_{(1)} = \frac{\widehat{T}_{(1)}}{\Gamma^{-1}_{\widehat{N}_{(1)}}(1-\alpha)} \tag{1.30}$$

as an approximate lower confidence limit of $\theta_{(1)}$. For $\underline{\theta}_{(1)}$, we have

THEOREM 1.3. *Under the condition of Theorem 1.2,*

$$P\left\{\theta_{(1)} \geq \underline{\theta}_{(1)}\right\} \longrightarrow 1-\alpha, \tag{1.31}$$

which shows that $\underline{\theta}_{(1)}$ is level consistent.

For the efficiency of $\underline{\theta}_{(1)}$, we have

THEOREM 1.4. *Under the condition of Theorem 1.2,*

$$\left(\frac{\partial\theta_{(1)}}{\partial\eta^{\tau}} I_{\eta}^{-1} \frac{\partial\theta_{(1)}}{\partial\eta}\right)^{-1/2} \left(\underline{\theta}_{(1)} - \theta_{(1)}\right) \xrightarrow{d} N(-u_{1-\alpha}, 1). \tag{1.32}$$

REMARK 1.3. In practice, we want to get the lower confidence limit of the reliability $R_{(1)} = \exp\{-\frac{t_0}{\theta_{(1)}}\}$. According to Theorem 1.4, the lower confidence limit of $R_{(1)}$ is

$$\underline{R_{(1)}} = \exp\left\{-\frac{t_0 \Gamma^{-1}_{\widehat{N}_{(1)}}(1-\alpha)}{\widehat{T}_{(1)}}\right\}.$$

In section 2, we will give the proofs.

2. The Proofs.

Proof of Theorem 1.1. Write

$$\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \tag{2.1}$$

where

$$\eta_1 = \begin{pmatrix} \theta_{m_1} \\ \vdots \\ \theta_{m_l} \end{pmatrix}, \qquad \eta_2 = \begin{pmatrix} \theta_{m_{l+1}} \\ \vdots \\ \theta_{m_k} \end{pmatrix},$$

$$M_0 = \{m_1, \cdots, m_k\}, \qquad M(m_0) = \{m_{l+1}, \cdots, m_k\}.$$

By using partitioned matrix calculations, we rewrite the matrix I_η in the form

$$I_\eta = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}, \tag{2.2}$$

where

$$\left.\begin{aligned}
D_{11} &= \sum_{m\in M^{(1)}\backslash m_0} \frac{n_m}{\theta_m^2}\frac{\partial\theta_m}{\partial\eta_1}\frac{\partial\theta_m}{\partial\eta_1^{\tau}},\\
D_{12} &= \sum_{m\in M^{(1)}\backslash m_0} \frac{n_m}{\theta_m^2}\frac{\partial\theta_m}{\partial\eta_1}\frac{\partial\theta_m}{\partial\theta_{m_0}}\frac{\partial\theta_{m_0}}{\partial\eta_1^{\tau}},\\
D_{21} &= D_{12}^{\tau},\\
D_{22} &= \sum_{m\in M^{(1)}\backslash m_0} \frac{n_m}{\theta_m^2}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2\frac{\partial\theta_{m_0}}{\partial\eta_2}\frac{\partial\theta_{m_0}}{\partial\eta_2^{\tau}}\\
&\quad+\frac{n_{m_0}}{\theta_{m_0}^2}\frac{\partial\theta_{m_0}}{\partial\eta_2}\frac{\partial\theta_{m_0}}{\partial\eta_2^{\tau}}+\Lambda^{-1},
\end{aligned}\right\} \tag{2.3}$$

$$\Lambda = \operatorname{diag}\left(\frac{\theta_{m_{l+1}}^2}{n_{m_{l+1}}},\cdots,\frac{\theta_{m_k}^2}{n_{m_k}}\right). \tag{2.4}$$

Similarly, we write the matrix $I_{\eta^{(1)}}$ in the form

$$I_{\eta^{(1)}} = \begin{pmatrix} D_{11}^{(1)} & D_{12}^{(1)} \\ D_{21}^{(1)} & D_{22}^{(1)} \end{pmatrix}, \tag{2.5}$$

where

$$\left.\begin{aligned}
D_{11}^{(1)} &= \sum_{m\in M^{(1)}\backslash m_0} \frac{n_m^{(1)}}{\theta_m^2}\frac{\partial\theta_m}{\partial\eta_1}\frac{\partial\theta_m}{\partial\eta_1^{\tau}} = D_{11},\\
D_{12}^{(1)} &= \sum_{m\in M^{(1)}\backslash m_0} \frac{n_m^{(1)}}{\theta_m^2}\frac{\partial\theta_m}{\partial\eta_1}\frac{\partial\theta_m}{\partial\theta_{m_0}},\\
D_{21}^{(1)} &= D_{12}^{(1)^{\tau}},\\
D_{22}^{(1)} &= \sum_{m\in M^{(1)}} \frac{n_m^{(1)}}{\theta_m^2}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2.
\end{aligned}\right\} \tag{2.6}$$

By the inverse formula for partitioned matrices

$$I_\eta^{-1} = \begin{pmatrix} D_{11}^{-1} + D_{11}^{-1}D_{12}\Delta^{-1}D_{21}D_{11}^{-1} & -D_{11}^{-1}D_{12}\Delta^{-1} \\ -\Delta^{-1}D_{21}D_{11}^{-1} & \Delta^{-1} \end{pmatrix}$$

where

$$\Delta = D_{22} - D_{21}D_{11}^{-1}D_{12}.$$

Hence

$$\frac{\partial\theta_{(1)}}{\partial\eta^\tau} I_\eta^{-1} \frac{\partial\theta_{(1)}}{\partial\eta} = \left(\frac{\partial\theta_{(1)}}{\partial\eta_1^\tau}, \frac{\partial\theta_{(1)}}{\partial\theta_{m_0}}\frac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\right)$$

$$\cdot\begin{pmatrix} D_{11}^{-1} + D_{11}^{-1}D_{12}\Delta^{-1}D_{21}D_{11}^{-1} & -D_{11}^{-1}D_{12}\Delta^{-1} \\ -\Delta^{-1}D_{21}D_{11}^{-1} & \Delta^{-1}\end{pmatrix}\begin{pmatrix}\dfrac{\partial\theta_{(1)}}{\partial\eta_1} \\ \dfrac{\partial\theta_{(1)}}{\partial\theta_{m_0}}\dfrac{\partial\theta_{m_0}}{\partial\eta_2}\end{pmatrix}$$

$$= \left(\frac{\partial\theta_{(1)}}{\partial\eta_1^\tau}, \frac{\partial\theta_{(1)}}{\partial\theta_{m_0}}\right)$$

$$\cdot\begin{pmatrix} D_{11}^{-1} + D_{11}^{-1}D_{12}\Delta^{-1}D_{21}D_{11}^{-1} & -D_{11}^{-1}D_{12}\Delta^{-1}\dfrac{\partial\theta_{m_0}}{\partial\eta_2} \\ -\dfrac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\Delta^{-1}D_{21}D_{11}^{-1} & \dfrac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\Delta^{-1}\dfrac{\partial\theta_{m_0}}{\partial\eta_2}\end{pmatrix}\begin{pmatrix}\dfrac{\partial\theta_{(1)}}{\partial\eta_1} \\ \dfrac{\partial\theta_{(1)}}{\partial\theta_{m_0}}\end{pmatrix}. \tag{2.7}$$

From the definition of D_{ij} and $D_{ij}^{(1)}$ we know that

$$D_{11}^{(1)} = D_{11},$$
$$D_{12}^{(1)}\frac{\partial\theta_{m_0}}{\partial\eta_2^\tau} = D_{12}.$$

Substituting these formulas into (2.7), we obtain

$$\frac{\partial\theta_{(1)}}{\partial\eta^\tau} I_\eta^{-1} \frac{\partial\theta_{(1)}}{\partial\eta} = \frac{\partial\theta_{(1)}}{\partial\eta^{(1)^\tau}}$$

$$\begin{pmatrix} D_{11}^{(1)^{-1}} + D_{11}^{(1)^{-1}}D_{12}^{(1)}D_{21}^{(1)}KD_{21}^{(1)}D_{11}^{(1)^{-1}} & -D_{11}^{(1)^{-1}}D_{12}^{(1)}K \\ -KD_{21}^{(1)}D_{11}^{(1)^{-1}} & K\end{pmatrix}\frac{\partial\theta_{(1)}}{\partial\eta^{(1)}},$$

where

$$K \triangleq \frac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\Delta^{-1}\frac{\partial\theta_{m_0}}{\partial\eta_2}.$$

From the above equality we know that to prove (1.17), it suffices to prove

$$\frac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\Delta^{-1}\frac{\partial\theta_{m_0}}{\partial\eta_2} = \left(D_{22}^{(1)} - D_{21}^{(1)}D_{11}^{(1)^{-1}}D_{12}^{(1)}\right)^{-1} \triangleq \Delta^{(1)^{-1}}. \tag{2.8}$$

By using the following identity for matrices

$$(A + UV^\tau)^{-1} = A^{-1} - \frac{A^{-1}UV^\tau A^{-1}}{1 + V^\tau A^{-1}U}.$$

where U and V are vectors, we obtain

$$K = \frac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\left(\Lambda^{-1} + \sum_{m\in M^{(1)}\setminus m_0}\frac{n_m}{\theta_m^2}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2\frac{\partial\theta_{m_0}}{\partial\eta_2}\frac{\partial\theta_{m_0}}{\partial\theta_2^\tau}\right.$$

$$+\frac{n_{m_0}}{\theta^2_{m_0}}\frac{\partial\theta_{m_0}}{\partial\eta_2}\frac{\partial\theta_{m_0}}{\partial\eta_2^\tau}-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\frac{\partial\theta_{m_0}}{\partial\eta_2}\frac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\Bigg)^{-1}\frac{\partial\theta_{m_0}}{\partial\eta_2}$$

$$=\frac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\Bigg[\Lambda-\Lambda\frac{\partial\theta_{m_0}}{\partial\eta_2}\frac{\partial\theta_{m_0}}{\partial\eta_2^\tau}\Lambda\Bigg(\frac{n_{m_0}}{\theta^2_{m_0}}$$

$$+\sum_{m\in M^{(1)}\backslash m_0}\frac{n_m}{\theta^2_m}\left(y\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\Bigg)\Bigg(1+\Bigg(\frac{n_{m_0}}{\theta^2_{m_0}}$$

$$+\sum_{m\in M^{(1)}\backslash m_0}\frac{n_m}{\theta^2_m}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\Bigg)\frac{\theta^2_{m_0}}{\widetilde{n}_{m_0}}\Bigg)^{-1}\Bigg]\frac{\partial\theta_{m_0}}{\partial\eta_2}$$

$$=\frac{\theta^2_{m_0}}{\widetilde{n}_{m_0}}-\left(\frac{\theta^2_{m_0}}{\widetilde{n}_{m_0}}\right)^2\Bigg(\frac{n_{m_0}}{\theta^2_{m_0}}+\sum_{m\in M^{(1)}\backslash m_0}\frac{n_{m_0}}{\theta^2_m}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2$$

$$-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\Bigg)\Bigg(1+\Bigg(\frac{n_{m_0}}{\theta^2_{m_0}}+\sum_{m\in M^{(1)}\backslash m_0}\frac{n_m}{\theta^2_m}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2$$

$$-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\Bigg)\frac{\theta^2_{m_0}}{\widetilde{n}_{m_0}}\Bigg)^{-1}$$

$$=\frac{\theta^2_{m_0}}{\widetilde{n}_{m_0}}\Bigg[1-\Bigg(\frac{n_{m_0}}{\widetilde{n}_{m_0}}+\frac{\theta^2_{m_0}}{\widetilde{n}_{m_0}}\Bigg(\sum_{m\in M^{(1)}\backslash m_0}\frac{n_m}{\theta^2_m}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2$$

$$-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\Bigg)\Bigg)\Bigg(1+\frac{n_{m_0}}{\widetilde{n}_{m_0}}+\frac{\theta^2_{m_0}}{\widetilde{n}_{m_0}}\Bigg(\sum_{m\in M^{(1)}\backslash m_0}\frac{n_m}{\theta^2_m}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2$$

$$-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\Bigg)\Bigg)^{-1}\Bigg]$$

$$=\Bigg(\frac{n_{m_0}+\widetilde{n}_{m_0}}{\theta^2_{m_0}}+\sum_{m\in M^{(1)}\backslash m_0}\frac{n_m}{\theta^2_m}\left(\frac{\partial\theta_m}{\partial\theta_{m_0}}\right)^2-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\Bigg)^{-1}$$

$$=\left(D^{(1)}_{22}-D^{(1)}_{21}D^{(1)^{-1}}_{11}D^{(1)}_{12}\right)^{-1},$$

where $\widetilde{n}_{m_0}$ and Λ are given by (1.19) and (2.4) respectively.

For the proof of the Theorem 1.2, some preliminary lemmas are first developed

LEMMA 2.1. *Let $\{X_m, m\in M\}$ be an exponential system and $\{x_{mi},\ m\in M,\ i=1,2,\cdots,n_m\}$ be independently observed data. Suppose that for every $m\notin M_0$, θ_m, as a function of $\{\theta_{\widetilde{m}}, \widetilde{m}\in M(m)\}$, satisfies the condition*

$$\left(\frac{\partial\theta_m}{\partial\theta_{\widetilde{m}}},\ \widetilde{m}\in M(m)\right)\neq 0.$$

Then, as $\min\{n_m, m \in M\} \longrightarrow \infty$,

$$\frac{\widehat{\theta}_m}{\theta_m} \longrightarrow 1, wp1, \tag{2.9}$$

$$\frac{\widehat{T}_m}{\widehat{N}_m} \longrightarrow \theta_m, wp1, \tag{2.10}$$

$$\frac{\widehat{N}_m}{N_m} \longrightarrow 1, wp1, \tag{2.11}$$

$$\frac{\widehat{T}_m - \widehat{N}_m\theta_m}{\sqrt{\widehat{N}_m\theta_m^2}} \xrightarrow{d} N(0,1), \tag{2.12}$$

$$\frac{\widetilde{\theta}_m - \theta_m}{\sqrt{\theta_m^2/\widetilde{n}_m}} \xrightarrow{d} N(0,1), \qquad m \notin M_0. \tag{2.13}$$

Proof of Lemma 2.1. We prove the lemma by induction. When $m \in M_0$, according to the definition of $\widehat{\theta}_m, \widehat{T}_m, \widehat{N}_m$ (see (1.20)–(1.22)), it is straightforward that (2.9)–(2.12) hold. Now suppose that (2.9)–(2.13) hold for all $\widetilde{m} \in M(m)$ (when $\widetilde{m} \in M_0$, we only require that (2.9)–(2.12) hold for $\widetilde{m}_m$). According to the definition of $\widehat{\theta}_m$, we have

$$\widehat{\theta}_m = \frac{T_m + \hat{\widetilde{n}}_m\widetilde{\theta}_m}{n_m + \hat{\widetilde{n}}_m}. \tag{2.14}$$

By the definition of $\widetilde{\theta}_m$ we know that

$$\widetilde{\theta}_m = \theta_m(\theta_{\widetilde{m}}, \widetilde{m} \in M(m))\Big|_{\theta_{\widetilde{m}} = \frac{\widehat{T}_{\widetilde{m}}}{\widehat{N}_{\widetilde{m}}}} \longrightarrow \theta_m(\theta_{\widetilde{m}}, \widetilde{m} \in M(m)), wp1, \tag{2.15}$$

from which, combining the consistency of T_m/n_m, we know that (2.15) implies that $\widehat{\theta}_m \longrightarrow \theta_m$, wp1 and $\widehat{T}_m/\widehat{N}_m \longrightarrow \theta_m$, wp1, i.e. (2.9) and (2.10) hold for m. By the definition of $\widehat{N}_m$ and N_m, we have

$$\frac{\widehat{N}_m}{N_m} = \left(n_m + \frac{\widetilde{\theta}_m^2}{\sum\left(\widehat{\frac{\partial\theta_m}{\partial\theta_{\widetilde{m}}}}\right)^2 \frac{\widehat{\theta}_{\widetilde{m}}^2}{\widehat{N}_{\widetilde{m}}}}\right) \Bigg/ \left(n_m + \frac{\theta_m^2}{\sum\left(\frac{\partial\theta_m}{\partial\theta_{\widetilde{m}}}\right)^2 \frac{\theta_{\widetilde{m}}^2}{N_{\widetilde{m}}}}\right).$$

From this expression, together with the facts that

$$\sum\left(\frac{\partial\theta_m}{\partial\theta_{\widetilde{m}}}\right)^2 \neq 0, \qquad \widehat{N}_{\widetilde{m}} = N_{\widetilde{m}}(1 + o(1)), wp1,$$

and $\widehat{\theta}_{\widetilde{m}} = \theta_{\widetilde{m}}(1 + o(1))$, wp1, we obtain

$$\frac{\widehat{N}_m}{N_m} = 1 + o(1), \quad wp1,$$

which shows that (2.11) holds for m. Using a Taylor expansion, we obtain

$$\begin{aligned}
\frac{\widetilde{\theta}_m - \theta_m}{\sqrt{\theta_m^2/\widetilde{n}_m}} &= \sqrt{\frac{\widetilde{n}_m}{\theta_m^2}} \left(\theta_m \left(\frac{\widehat{T}_{\widetilde{m}}}{\widehat{N}_{\widetilde{m}}}, \widetilde{m} \in M(m) \right) - \theta_m(\theta_{\widetilde{m}}, \widetilde{m} \in M(m)) \right) \\
&= \sqrt{\frac{\widetilde{n}_m}{\theta_m^2}} \sum \left(\frac{\partial \theta_m}{\partial \theta_{\widetilde{m}}} \right) \left(\frac{\widehat{T}_{\widetilde{m}}}{\widehat{N}_{\widetilde{m}}} - \theta_{\widetilde{m}} \right) \cdot (1 + o_p(1)) \\
&\xrightarrow{d} N(0,1),
\end{aligned}$$

which shows that (2.13) holds. From (2.13), we know that

$$\begin{aligned}
\frac{\widehat{T}_m - \widehat{N}_m \theta_m}{\sqrt{\widehat{N}_m \theta_m^2}} &= \frac{\sqrt{n_m}}{\sqrt{\widehat{N}_m \theta_m^2}} \left(\frac{1}{\sqrt{n_m}} (T_m - n_m \theta_m) \right) \\
&\quad + \frac{\sqrt{\widehat{\widetilde{n}}_m}}{\sqrt{\widehat{N}_m}} \left(\frac{\sqrt{\widehat{\widetilde{n}}_m}}{\theta_m} (\widetilde{\theta}_m - \theta_m) \right) \xrightarrow{d} N(0,1),
\end{aligned}$$

i.e., (2.12) holds. By induction, we know that the conclusion of the lemma follows.

Proof of Theorem 1.2. Examining the definition of the numbers $N_{(1)}$ and $n_{(1)}^{k_M}$ given in (1.18) and (1.25) respectively, we know that $n_{(1)}^{(k_M)} = N_{(1)}$. Therefore (1.28) follows from (2.12), i.e.

$$\left(\frac{\partial \theta_{(1)}}{\partial \eta^\tau} I_\eta^{-1} \frac{\partial \theta_{(1)}}{\partial \eta} \right)^{-\frac{1}{2}} (\widehat{\theta}_{(1)} - \theta_{(1)}) = \frac{\widehat{\theta}_{(1)} - \theta_{(1)}}{\sqrt{\theta_{(1)}^2 / N_{(1)}}} \xrightarrow{d} N(0,1).$$

Proof of Theorem 1.3. Let U be a random variable with conditional distribution

$$P\{U < x | \widehat{N}_{(1)}\} = \frac{1}{\Gamma(\widehat{N}_{(1)})} \int_0^x u^{\widehat{N}_{(1)} - 1} e^{-u} \, du.$$

It is easy to show that

$$\frac{U - \widehat{N}_{(1)}}{\sqrt{\widehat{N}_{(1)}}} | \widehat{N}_{(1)} \xrightarrow{d} N(0,1), wp1$$

from which it follows that

$$\frac{\Gamma^{-1}_{\widehat{N}_{(1)}}(1-\alpha)-\widehat{N}_{(1)}}{\sqrt{\widehat{N}_{(1)}}} \longrightarrow u_{1-\alpha}, wp1, \tag{2.16}$$

where $u_{1-\alpha}$ is the $1-\alpha$ quantile of the standard normal distribution. From (2.12), we obtain

$$\frac{\frac{\widehat{T}_{(1)}}{\theta_{(1)}}-\widehat{N}_{(1)}}{\sqrt{\widehat{N}_{(1)}}} \xrightarrow{d} N(0,1). \tag{2.17}$$

Combining (2.16), (2.17), we obtain

$$\begin{aligned}
&P\left\{\frac{\widehat{T}_{(1)}}{\Gamma^{-1}_{\widehat{N}_{(1)}}(1-\alpha)} \le \theta_{(1)}\right\} \\
=&P\left\{\left(\frac{\widehat{T}_{(1)}}{\theta_{(1)}}-\widehat{N}_{(1)}\right)\Big/\sqrt{\widehat{N}_{(1)}} < \left(\Gamma^{-1}_{\widehat{N}_{(1)}}(1-\alpha)-\widehat{N}_{(1)}\right)\Big/\sqrt{N_{(1)}}\right\} \\
\longrightarrow&1-\alpha.
\end{aligned}$$

Proof of Theorem 1.4. First we have

$$\begin{aligned}
\frac{\widehat{T}_{(1)}/\Gamma^{-1}_{\widehat{N}_{(1)}}(1-\alpha)-\theta_{(1)}}{\sqrt{\theta^2_{(1)}/\widehat{N}_{(1)}}} =& \frac{\widehat{T}_{(1)}/\widehat{N}_{(1)}-\theta_{(1)}}{\sqrt{\theta^2_{(1)}/\widehat{N}_{(1)}}} \\
&+\frac{\widehat{T}_{(1)}/\Gamma^{-1}_{\widehat{N}_{(1)}}(1-\alpha)-\widehat{T}_{(1)}/\widehat{N}_{(1)}}{\sqrt{\theta^2_{(1)}/\widehat{N}_{(1)}}}.
\end{aligned} \tag{2.18}$$

From (2.16), we obtain

$$\begin{aligned}
&\frac{\widehat{T}_{(1)}\Big/\Gamma^{-1}_{\widehat{N}_{(1)}}(1-\alpha)-\widehat{T}_{(1)}/\widehat{N}_{(1)}}{\sqrt{\theta^2_{(1)}/\widehat{N}_{(1)}}} \\
=&\frac{\widehat{T}_{(1)}\Big/(\widehat{N}_{(1)}+\sqrt{\widehat{N}_{(1)}}u_{1-\alpha}(1+o(1)))-\widehat{T}_{(1)}/\widehat{N}_{(1)}}{\sqrt{\theta^2_{(1)}/\widehat{N}_{(1)}}} \\
=&-\frac{\widehat{T}_{(1)}u_{1-\alpha}(1+o(1))}{\widehat{N}_{(1)}+\sqrt{\widehat{N}_{(1)}}u_{1-\alpha}(1+o(1))}\frac{1}{\theta_{(1)}} \\
\longrightarrow&-u_{1-\alpha}, wp1.
\end{aligned} \tag{2.19}$$

According to Theorem 1.3, we have

$$\left(\frac{\widehat{T}_{(1)}}{\widehat{N}_{(1)}} - \theta_{(1)}\right) \Bigg/ \sqrt{\frac{\theta^2_{(1)}}{\widehat{N}_{(1)}}} \xrightarrow{d} N(0,1),$$

which, together with (2.19), implies that (1.32) holds.

Acknowledgment. The author would like to thank Dr. Shi Peide for typing of this paper.

REFERENCE

ZHENG ZHONGGUO (1991). Virtual system for system tree. Tech. report, Peking University.

DEPARTMENT OF PROBABILITY
AND STATISTICS
PEKING UNIVERSITY
PEKING, CHINA