Limit Theorems for Winsorized Mean and Wrong-Model Likelihood Estimation*

by

Heebok Park

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series No. 55

October, 1965

---

* This paper is based on the author's Ph.D. thesis at the University of Chicago.

Limit Theorems for Winsorized Mean and Wrong-Model Likelihood Estimation*

by

Heebok Park

Purdue University

1. Introduction and Summary. Suppose $X_1 < X_2 < \ldots < X_n$ are order statistics in a random sample of size n from $g(\cdot)$. Let p_1, p_2 be prescribed with $o < p_1 < p_2 < 1$; let $q = 1 - p_i$. For each n, ν_1 and ν_3 are defined by $\nu_1 = [np_1]$, and $\nu_3 = [nq_2]$. The notation $[x]$ denotes the largest integer not exceeding x. Let I_1, I_2 and I_3 be sets of integers

$$I_1 = \{1, 2, \ldots, \nu_1\}$$

$$I_2 = \{\nu_1 + 2, \nu_1 + 3, \ldots, n - \nu_3 - 1\}$$

$$I_3 = \{n - \nu_3 + 1, n - \nu_3 + 2, \ldots, n\}$$

and set $\nu_2 = n - \nu_1 - \nu_3 - 2$, i.e.; ν_j is the number of integers contained in I_j. Let $\psi(\cdot)$ be a function. Our object here is to develope an asymptotics of statistics in a form of $\sum_{i \in I_j} \psi(X_i)$. Such a statistics appears, for instance, in the expression of Winsorized mean, $\bar{X}^W_{p_1,p_2}$, obtained for prescribed $o < p_1 < p_2 < 1$:

$$\bar{X}^W_{p_1,p_2} \equiv \frac{1}{n}\left[(\nu_1 + 1)\, X_{\nu_1+1} + \sum_{i \in I_2} X_i + (\nu_3 + 1)\, X_{n-\nu_3}\right].$$

---

* This paper is based on the author's Ph.D. Thesis at the University of Chicago.

The log-likelihood based on the entire sample is

$$\sum_{i\in I_1} \log g(x_i) + \log g(x_{\nu_1+1}) + \sum_{i\in I_2} \log g(x_i)$$

$$+ \log g(x_{n-\nu_3}) + \sum_{i\in I_3} \log g(x_i)$$

where as

$$\text{constant} + \nu_1 \log G(x_{\nu_1+1}) + \log g(x_{\nu_1+1})$$

$$+ \sum_{i\in I_2} \log g(x_i) + \log g(x_{n-\nu_3})$$

$$+ \nu_3 \log [1 - G(x_{n-\nu_3})]$$

is the log-likelihood based on trimmed sample $\{X_{\nu_1+1}, X_{\nu_1+2}, \ldots, X_{n-\nu_3}\}$.

Section 2 contains asymptotic theory of such statistics. For the future use, it is supposed that we have two such functions $\psi_k(.)$, $k = 1,2$. This will provide us with a joint asymptotic distribution of two different statistics, for instance, the maximum likelihood estimate based on the entire sample and Winsorized mean. Extension to the case $k > 2$ will be immediate.

As an immediate application to this result, asymptotic theory is given for Winsorized mean, and corresponds to a central limit theorem in the complete sample case.

Section 3 contains a maximum likelihood theory that appears to be new, even for complete sample. The asymptotic distribution theory of maximum likelihood estimates based on trimmed sample is developed when data are sampled from $g(\cdot|\theta)$, but when the estimates were computed with the assumptions that they are $f(\cdot|\theta)$. Such a wrong-model maximum likelihood estimate is typically a consistent and asymptotically normal estimate not of θ but of some function $\theta'(\theta)$ of θ.

2. Asymptotic theory for partial sums of order statistics.

Let η_i be the p_i fractile of $g(\cdot)$. Particularly for the p_i sample fractile, we use a distinct notation:

$$U_{n1} \equiv X_{\nu_1+1}$$

$$U_{n2} \equiv X_{n-\nu_3}.$$

For any $u_1 < u_2$, define the density truncated to $[-x, u_1]$, $[u_1, u_2]$ and $[u_2\ x]$ as

$$g_1(x \mid u_1, u_2) = \begin{cases} \dfrac{g(x)}{G(u_1)} & ; \quad x < u_1 \\ 0 & ; \quad \text{otherwise,} \end{cases}$$

$$(1.1) \qquad g_2(x \mid u_1, u_2) = \begin{cases} \dfrac{g(x)}{G(u_2) - G(u_1)} & ; \quad u_1 < x < u_2 \\ 0 & ; \quad \text{otherwise,} \end{cases}$$

$$g_3(x \mid u_1, u_2) = \begin{cases} \dfrac{g(x)}{1 - G(u_2)} & ; \quad u_2 < x \\ 0 & ; \quad \text{otherwise.} \end{cases}$$

Let

$$(1.2) \qquad \mu^{kj}(u_1, u_2) = \int \psi_k(x)\, g_j(x \mid u_1, u_2)\, dx,$$

$$(1.3) \qquad \mu^{kj}_{to} = \left[\frac{\partial}{\partial u_t} \mu^{kj}(u_1, u_2) \right]_{u_1 = \eta_1,\ u_2 = \eta_2}, \qquad t = 1,2$$

(1.4) $$\sigma^j_{kk'}(u_1, u_2) = \int \left[\psi_k(x) - \mu^{kj}(u_1, u_2)\right]\left[\psi_{k'}(x) - \mu^{k'j}(u_1, u_2)\right] \cdot$$

$$\cdot\ g_j\ (x|u_1, u_2)\ dx,\ k,\ k' = 1,\ 2,$$

(1.5) $$\sum^j (u_1, u_2) = \begin{pmatrix} \sigma^j_{11}\ (u_1, u_2) & \sigma^j_{12}\ (u_1, u_2) \\ \sigma^j_{21}\ (u_1, u_2) & \sigma^j_{22}\ (u_1, u_2) \end{pmatrix},$$

Denote by $\left[\sum^j (u_1, u_2)\right]^{1/2}$ any square root of the matrix $\sum^j (u_1, u_2)$.

In the subsequent discussion, $X_n|Y_n = y \xrightarrow{\mathcal{L}} Z$ is understood to mean the sequence $\{D(X_n|Y_n = y)\}$ converges weakly to $D(Z)$ with the convergence in probability to a constant, in the notation $\xrightarrow{p}$, as a special case. If μ is a distribution then, $X_n|Y_n = y \xrightarrow{\mathcal{L}} \mu$ means that $D(X_n|Y_n = y)$ converges weakly to μ. To obtain the desired asymptotics, we assume that the functions $\psi_k(x)$, $k = 1, 2$ and $g(x)$ satisfy:

(A 1) $g(x)$ is continuous in the neighborhood of S_i of η_i, $i = 1, 2$

(A 2) Matrix $\sum^j(\eta_1, \eta_2)$ is positive definite

(A 3) $\int |\ \psi_k\ (x)|^3\ g(x)\ dx < \infty$.

Theorem 1. If $X_1 < X_2 < \dots < X_n$ are order statistics in a random sample of size n from a density $g(x)$; if the function $\psi_k(x)$, $k = 1, 2$ and the model $g(x)$ satisfy assumptions (A 1), (A 2) and (A 3), then for fixed $o < p_1 < p_2 < 1$, the mean of $\psi_k(X_i)$ over index set I_j,

$$T^{kj}_n = \frac{1}{\nu_j} \sum_{i \in I_j} \psi_k\ (X_i)$$

have the following properties:

for each j

$$(1.6) \quad \begin{pmatrix} T_n^{1j} \\ T_n^{2j} \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \mu^{1j} (\eta_1, \eta_2) \\ \mu^{2j} (\eta_1, \eta_2) \end{pmatrix} ;$$

for each j and sufficiently large n

$$(1.7) \quad \begin{pmatrix} T_n^{1j} \\ \\ T_n^{2j} \end{pmatrix} = \begin{pmatrix} \mu^{1j} (\eta_1, \eta_2) \\ \\ \mu^{2j} (\eta_1, \eta_2) \end{pmatrix} + \frac{1}{\sqrt{n}} \begin{pmatrix} \mu_{10}^{1j} X_{n1} + \mu_{20}^{1j} X_{n2} \\ \\ \mu_{10}^{2j} X_{n1} + \mu_{20}^{2j} X_{n2} \end{pmatrix}$$

$$+ \frac{1}{\sqrt{n}} \left[\Sigma^{j} (\eta_1, \eta_2) \right]^{1/2} \cdot \begin{pmatrix} Y_n^{1j} \\ Y_n^{2j} \end{pmatrix}$$

where

(1.8) $X_{ni} \doteq \sqrt{n} \; (U_{ni} - \eta_i)$, $i = 1, 2$ and Y_n^{kj}, $k = 1, 2$ for $j = 1, 2, 3$

are random variables such that

$$(1.9) \quad (Y_n^{11} \; Y_n^{21} \; Y_n^{12} \; Y_n^{22} \; Y_n^{13} \; Y_n^{23} \; X_{n1} \; X_{n2})'$$

$$\xrightarrow{L} N_8 \left[0, \Lambda \right]$$

where

$$(1.10) \qquad \Lambda = \begin{bmatrix} \frac{1}{p_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{p_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{1-p_1-q_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{1-p_1-q_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{q_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{q_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & c_{11} & c_{12} \\ 0 & 0 & 0 & 0 & 0 & 0 & c_{12} & c_{22} \end{bmatrix}$$

with

$$(1.11) \qquad c_{ij} = \frac{p_i \; q_j}{g\;(\eta_i)\; g\;(\eta_j)} \; ; \quad i, j = 1, 2.$$

<u>Proof</u>. Suppose that Z is a random variable distributed with density $g_j\;(\cdot|u_1, u_2)$. For finite (u_1, u_2), $u_1 < u_2$ and that

$$\beta^{kj}(u_1, u_2) = \int |\psi_k(z) - \mu^{kj}(u_1, u_2)|^3 \; g_j(z|u_1, u_2)\; dz$$

is finite. Also $\int_{u_1}^{u_2} \left[\psi_k\;(z)\right]^r g(z|u_1, u_2)\; dz$ is continuous in u_1, u_2, and all moments of $\psi_k(Z)$ under $g_j(Z|u_1, u_2)$ are continuous and differentiable in u_1, u_2 for $U_i \epsilon S_1$, $i = 1, 2$. Then, there exists constants a_1 and a_2 such that

$$|\beta^{kj}(u_1, u_2)| < a_2, \quad \sigma^j_{kk'}(u_1, u_2) < a_2,$$

and

$$\text{Det.} \textstyle\sum^{j}(u_1, u_2) > a_1 > 0$$

for all $u_1 \in S_i'$, $i = 1, 2$ where

$$\eta_i \in S_i' \subset \overline{S_i'} \subset S_i, \; i = 1, 2.$$

Moreover, if we denote the cofactor of the (k, k')th entry of $\sum^{j}(u_1, u_2)$ by $\bar{\sigma}^{j}_{kk'}(u_1, u_2)$, we have

$$|\bar{\sigma}^{j}_{kk'}(u_1, u_2)| < \infty.$$

In particular, $\bar{\sigma}^{j}_{kk'}(u_1, u_2)$ is simply $(-1)^{k+k'} \sigma^{j}_{kk'}(u_1, u_2)$ in the present situation.

Hence,

$$(1.12) \qquad \underset{k,k'=1,2}{\text{Max}} \left| \frac{\bar{\sigma}^{j}_{kk'}(u, \nu)}{\text{Det.} \sum^{j}(u_1, u_2)} \right| < \frac{a_2}{a_1}$$

uniformely in $u_i \in S_i'$, $i = 1, 2$.

Denote by F the distribution function of $N_2\left[0, \sum^{j}(u_1, u_2)\right]$ and by F_{ν_j} the distribution function of

$$\sqrt{\nu_j} \begin{pmatrix} \frac{1}{\nu_j} \sum\limits_{i\in I_j} \psi_1(Z_i) - \mu^{1j}(u_1, u_2) \\ \frac{1}{\nu_j} \sum\limits_{i\in I_j} \psi_2(Z_i) - \mu^{2j}(u_1, u_2) \end{pmatrix}$$

where $(Z_1, Z_2, \ldots, Z_{\nu_j})$ is a random sample of size ν_j from $g_j(\cdot|u_1, u_2)$. It follows from the theorem of Bergstróm (1945) that there exists a numerical constant b such that

$$(1.13) \quad \sup_{-\infty < y_1, y_2 < \infty} |F_{\nu_j}(y_1, y_2) - F(y_1, y_2)|$$

$$\leq \frac{b}{\sqrt{\nu_j}} \left(\sum_{k=1}^{2} \beta^{kj} (u_1, u_2) \right) \left(\underset{k,k' = 1,2}{\text{Max}} \left| \frac{\bar{\sigma}^j_{kk'}(u_1, u_2)}{\text{Det.} \sum^j_{\sim}(u_1, u_2)} \right| \right)^{3/2}$$

$$\leq \frac{b}{\sqrt{\nu_j}} \cdot 2\, a_2 \cdot \left(\frac{a_2}{a_1} \right)^{3/2}$$

for all $u_i \in S_i'$, $i = 1, 2$. Thus, the bivariate central limit theorem holds uniformly for a random sample $(Z_1, Z_2, \ldots, Z_{\nu_j})$ of size ν_j from $g_j\,(\cdot|u_1, u_2)$.

Conditionally on $U_{n1} = u_1$, $U_{n2} = u_2$, the order statistics $\{X_i\ ;\ i\epsilon I_j\}$ are distributed like ν_j order statistics $\tilde{Z} = \{Z_{(1)}, Z_{(2)}, \ldots Z_{(\nu_j)}\}$ in a random sample $(Z_1, Z_2, \ldots, Z_{\nu_j})$;

$$D\,(\tilde{Z}) = D\,(X_1, X_2, \ldots, X_{\nu_j} | U_{n1} = u_1, U_{n2} = u_2).$$

Hence, we have

$$(1.14) \quad \begin{pmatrix} T_n^{1j} \\ T_n^{2j} \end{pmatrix} \Bigg| U_{n1} = u_1, U_{n2} = u_2 \xrightarrow{p} \begin{pmatrix} \mu^{1j}(u_1, u_2) \\ \mu^{2j}(u_1, u_2) \end{pmatrix}$$

and

$$(1.15) \quad \sqrt{n} \left(\frac{\sqrt{\nu_j}}{\sqrt{n}} \right) \left[\sum^j_{\sim}(U_{n1}, U_{n2}) \right]^{-\frac{1}{2}} \begin{pmatrix} T_n^{1j} - \mu^{1j}(U_{n1}, U_{n2}) \\ T_n^{2j} - \mu^{2j}(U_{n1}, U_{n2}) \end{pmatrix} \Bigg| \begin{matrix} U_{n1} = u_1 \\ U_{n2} = u_2 \end{matrix}$$

$$\xrightarrow[\substack{n \to x \\ \nu_j \to x}]{L} N_2\,[o, I].$$

Let

$$(1.16) \qquad \begin{pmatrix} \bar{Y}_n^{1j} \\ \\ \bar{Y}_n^{2j} \end{pmatrix} = \sqrt{n} \left[\textstyle\sum^j (U_{n1}, U_{n2}) \right]^{-\frac{1}{2}} \begin{pmatrix} T_n^{1j} - \mu^{1j}(U_{n1}, U_{n2}) \\ \\ T_n^{2j} - \mu^{2j}(U_{n1}, U_{n2}) \end{pmatrix} .$$

Then, conditionally on $U_{n1} = u_1$, $U_{n2} = u_2$, three random vectors $(\bar{Y}_n^{1j}, \bar{Y}_n^{2j})'$, $j = 1, 2, 3$, are independent of each other and have the property

$$(1.17) \quad \left[\begin{matrix} \bar{Y}_n^{11} \\ \bar{Y}_n^{21} \\ \bar{Y}_n^{12} \\ \bar{Y}_n^{22} \\ \bar{Y}_n^{13} \\ \bar{Y}_n^{23} \end{matrix} \;\middle|\; \begin{matrix} U_{n1} = u_1 \\ \\ U_{n2} = u_2 \end{matrix} \right] \xrightarrow{L} N_6 \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{p_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{p_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{1-p_1-q_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{1-p_1-q_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{q_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{q_2} \end{pmatrix} \right]$$

The entries of the covariance matrix are due to the relations

$$\lim \frac{\nu_1}{n} = p_1, \quad \lim \frac{\nu_2}{n} = 1 - p_1 - q_2, \quad \lim \frac{\nu_3}{n} = q_2.$$

It is well known (Cramer, 1958, (28.5)) that the sample fractiles U_{n1} and U_{n2} satisfy:

$$U_{ni} \xrightarrow{p} \eta_i, \qquad i = 1, 2$$

and for

$$X_{ni} = \sqrt{n}\,(U_{ni} - \eta_i), \qquad i = 1, 2,$$

joint weak convergence holds;

$$\begin{pmatrix} X_{n1} \\ X_{n2} \end{pmatrix} \xrightarrow{L} N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C_{11} & C_{12} \\ C_{12} & C_{22} \end{pmatrix} \right]$$

where

$$C_{ij} = \frac{p_i \; q_j}{g(\eta_i) \cdot g(\eta_j)} \quad .$$

Thus, applying these together with (1.17) to Theorem 1 of Appendix, we have a relation

(1.18) $$(\bar{Y}_n^{11} \; \bar{Y}_n^{21} \; \bar{Y}_n^{12} \; \bar{Y}_n^{22} \; \bar{Y}_n^{13} \; \bar{Y}_n^{23} \; X_{n1} \; X_{n2})' \xrightarrow{L} N_8 \left[0, \; \Lambda \right]$$

where Λ is a 8×8 matrix given by (1.10).

Now, let X_{01}, X_{02} and Y_0^{1j}, Y_0^{2j}, $j = 1, 2, 3$ be random variables such that

(1.19) $$(Y_0^{11} \; Y_0^{21} \; Y_0^{12} \; Y_0^{22} \; Y_0^{13} \; Y_0^{23} \; X_{01} \; X_{02})' \sim N_8 \left[0, \; \Lambda \right] .$$

Since $\sigma_{kk'}^j (u_1, u_2)$ is jointly continuous in u_1, u_2 for $u_i \in S_i$, $i = 1, 2$, we have

$$\sigma_{kk'}^j (U_{n1}, U_{n2}) \xrightarrow{p} \sigma_{kk'}^j (\eta_1, \eta_2),$$

and, it follows that

$$\begin{pmatrix} Y_n^{1j} \\ \\ Y_n^{2j} \end{pmatrix} \equiv \sqrt{n} \left[\sum^{j} (\eta_1, \eta_2) \right]^{-\frac{1}{2}} \begin{pmatrix} T_n^{1j} - \mu^{1j}(U_{n1}, U_{n2}) \\ \\ T_n^{2j} - \mu^{2j}(U_{n1}, U_{n2}) \end{pmatrix}$$

and $(\bar{Y}_n^{1j}, \bar{Y}_n^{2j})'$ have the same limiting distribution.

Hence, by applying Theorem 2 of Appendix, we now obtain the following relation:

$$(1.20) \qquad \begin{pmatrix} T_n^{1j} \\ \\ T_n^{2j} \end{pmatrix} = \begin{pmatrix} \mu^{1j} (U_{n1}, U_{n2}) \\ \\ \mu^{2j} (U_{n1}, U_{n2}) \end{pmatrix} + \frac{1}{\sqrt{n}} \left[\sum^{j} (\eta_1, \eta_2) \right]^{\frac{1}{2}} \begin{pmatrix} Y_n^{1j} \\ \\ Y_n^{2j} \end{pmatrix}$$

where

$$(1.21) \qquad (Y_n^{11}, Y_n^{21}, Y_n^{12}, Y_n^{22}, Y_n^{13}, Y_n^{23}, X_{n1}, X_{n2})'$$

$$\xrightarrow{L} (Y_0^{11}\ Y_0^{21}\ Y_0^{12}\ Y_0^{22}\ Y_0^{13}\ Y_0^{23}\ X_{01}\ X_{02})'.$$

Let

$$U_{ni} = \eta_i + \frac{1}{\sqrt{n}} X_{ni}, \qquad i = 1, 2.$$

Since $\mu^{kj}(u_1, u_2)$ is continuous and differentiable in u_1 and u_2 in the neighborhood of η_1 and η_2 respectively, we can expand $\mu^{kj}(\eta_1 + x_{n1}/\sqrt{n}, \eta_2 + x_{n2}/\sqrt{n})$ in x_{n1} and x_{n2} near zero:

$$\mu^{kj}(U_{n1}, U_{n2}) = \mu^{kj}(\eta_1, \eta_2) + \frac{1}{\sqrt{n}} \left(\mu_{10}^{kj} x_{n1} + \mu_{20}^{kj} x_{n2} \right)$$

$$+ o\left(\frac{1}{\sqrt{n}}\right)$$

where μ_{10}^{kj} and μ_{20}^{kj} are derivatives defined by (1.3). Thus as $n \longrightarrow \infty$

$$\mu^{kj}(U_{n1}, U_{n2}) \xrightarrow{p} \mu^{kj}(\eta_1, \eta_2),$$

$$\sqrt{n}\left[\mu^{kj}(U_{n1}, U_{n2}) - \mu^{kj}(\eta_1, \eta_2)\right]$$

$$\xrightarrow{L} \mu_{10}^{kj} X_{01} + \mu_{20}^{kj} X_{02}.$$

Furthermore, weak convergence holds jointly for X_{n1}, X_{n2} and Y_n^{kj}; $j = 1, 2, 3$, $k = 1, 2$. The result of the theorem now follows.

The asymptotics of Winsorized mean $X_{p1,p2}^{W}$ will now be obtained applying the theorem to $k = 1$, $\psi_1(y) \equiv y$. I recall the deffinition of Winsorized mean

(1.22) $$\bar{X}_{p_1,p_2}^{W} = \frac{1}{n}\left[(\nu_1 + 1)\, U_{n1} + \left\{\sum_{i\in I_2} X_i\right\} + (\nu_3 + 1)\, U_{n2}\right].$$

The assumption (A3) will now accordingly be changed

(A3)' $$\sigma_{11}^2(\eta_1, \eta_2) > 0.$$

Corollary 1. If $\bar{X}_{p_1,p_2}^{W}$ is, for each n and for prescribed $0 < p_1 < p_2 < 1$, a Winsorized mean obtained from a random sample of size n from a density $g(\cdot)$; if the model $g(\cdot)$ satisfies the assumptions (A1), (A2) and (A3)' then the following properties hold:

(1.23) $$\bar{X}_{p_1,p_2}^{W} \xrightarrow{p} p_1 \eta_1 + q_2 \eta_2 + (1 - p_1 - q_2)\, \mu^{12}(\eta_1, \eta_2) \equiv \mu^*;$$

(1.24) $$\sqrt{n}\,(\bar{X}_{p_1,p_2}^{W} - \mu^*) \xrightarrow{L} \left[p_1 + (1 - p_1 - q_2)\, \mu_{10}^{12}\right] X_{01}$$

$$+ \left[q_2 + (1 - p_1 - q_2)\, \mu_{20}^{12}\right] X_{02}$$

$$+ (1 - p_1 - q_2)\left[\sigma_{11}^2(\eta_1, \eta_2)\right]^{\frac{1}{2}}\, Y_0^{12}$$

where X_{01}, X_{02} and Y_0^{12} are random variables such that

(1.25)
$$\begin{pmatrix} Y_0^{12} \\ X_{01} \\ X_{02} \end{pmatrix} \sim N_3 \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{1-p_1-q_2} & 0 & 0 \\ 0 & C_{11} & C_{12} \\ 0 & C_{12} & C_{22} \end{pmatrix} \right].$$

<u>Proof</u>. From Theorem 1, we can write

(1.26)
$$\frac{1}{\nu_2} \sum_{i \in I_2} X_i = \mu^{12} (\eta_1, \eta_2) + \frac{1}{\sqrt{n}} (\mu_{10}^{12} X_{n1} + \mu_{01}^{12} X_{n2}) + \frac{1}{\sqrt{n}} \left[\sigma_{11}^2 (\eta_1, \eta_2) \right]^{\frac{1}{2}} Y_n^{12},$$

where

(1.27)
$$U_{n_i} = \eta_i + \frac{1}{\sqrt{n}} X_{n_i}, \qquad i = 1, 2,$$

while

$$\begin{pmatrix} Y_n^{12} \\ X_{n1} \\ X_{n2} \end{pmatrix} \xrightarrow{L} \begin{pmatrix} Y_0^{12} \\ X_{01} \\ X_{02} \end{pmatrix}$$

Now, substituting (1.26) and (1.27) to the right hand side of (1.22), we obtain

(1.28)
$$\bar{X}^W_{p_1, p_2} = \frac{\nu_1}{n} \eta_1 + \frac{\nu_3}{n} \eta_2 + \frac{\nu_2}{n} \mu^{12} (\eta_1, \eta_2)$$

$$+ \frac{1}{\sqrt{n}} \left(\frac{\nu_1}{n} + \frac{\nu_2}{n} \mu_{10}^{12} \right) X_{n1}$$

$$+ \frac{1}{\sqrt{n}} \left(\frac{\nu_3}{n} + \frac{\nu_2}{n} \mu_{20}^{12} \right) X_{n2}$$

$$+ \frac{\nu_3}{n} \left[\sigma_{11}^2 (\eta_1, \eta_2) \right]^{\frac{1}{2}} \frac{Y_n^{12}}{\sqrt{n}} .$$

The results follows as $n \to \infty$.

3. Maximum likelihood estimation under the incorrect family of densities.

Let $\{g(x|\theta) ; \theta = (\theta_1, \theta_2) \in \Theta\}$ and $\{f(x|\varphi); \varphi = (\varphi_1, \varphi_2) \in \Phi\}$ denote two parametric families of densities with parameter spaces which are open subsets of 2-dimensional Euclidian space. The two families might be alternative models for observed data. In this section, we shall develope an asymptotic theory of maximum likelihood estimation based on a trimmed sample

$$\tilde{X}_{p_1,p_2} \equiv (U_{n1}, W_n, U_{n2}) ; W_n \equiv \{X_i, i \in I_2\},$$

when the data are incorrectly assumed to be sampled from $f(x|\varphi)$, $\varphi \in \Phi$, but in truth are sampled from some $g(x|\theta)$, $\theta \in \Theta$.

In order to interprete the estimate, some correspondence between Θ and Φ is needed. For example, if Θ and Φ are both the real line, and, if θ and φ are the means of their respective distributions, one possible correspondence is $\varphi \longleftrightarrow \theta$. For another example, that θ and φ are respectively the scale paramers for the normal and the Laplace distributions. Equatite variance would yield to correspondence

$$\sqrt{2}\, \varphi \longleftrightarrow \theta$$

whereas equatile quartiles would yield

$$\varphi \longleftrightarrow 0.6745\, \theta.$$

Assume that a one-to-one correspondence has been established for interpretations. In the technical developement, a second correspondence arises. Denote it by $\theta \longrightarrow \varphi^*(\theta)$. Roughly, $\varphi^*(\theta)$ is the value of φ to which the estimates obtained under the model $f(\cdot)$ converges when θ is the actual parameter of the true model.

Under the (incorrect) assumption that $\tilde{X}_{p_1,p_2}$ is from $f(\cdot|\varphi)$, the sampling density would be

$$(2.1) \qquad h(u_{n1}, w_n, u_{n2} \mid \varphi) = \frac{n!}{\nu_1! \, \nu_3!} F^{\nu_1}(u_{n1} \mid \varphi) \, f(u_{n1} \mid \varphi) \cdot$$

$$\cdot \Big[\prod_{i \in I_2} f(x_i \mid \varphi)\Big] f(u_{n2} \mid \varphi) \Big[1 - F(u_{n2} \mid \varphi)\Big]^{\nu_3}.$$

We shall denote the (incorrect) log-likelihood by $L_n(\varphi)$;

$$(2.2) \qquad L_n(\varphi) \equiv \log h(u_{n1}, w_n, u_{n2} \mid \varphi).$$

Let

$$(2.3) \qquad L_{n1}(\varphi \,;\, W_n \mid U_{n1}, U_{n2}) = \sum_{i \in I_2} \log f(x_i|\varphi)$$

$$(2.4) \qquad L_{n2}(\varphi \,;\, U_{n1}, U_{n2})$$

$$= \log \Big[F^{\nu_1}(u_{n1}|\varphi) \, f(u_{n1}|\varphi) \, f(u_{n2}|\varphi) \left\{1 - F(u_{n2}|\varphi)\right\}^{\nu_3}\Big]$$

so that

$$(2.5) \qquad L_n(\varphi) = \log \frac{n!}{\nu_1! \, \nu_3!} + L_{n1}(\varphi \,;\, W_n|U_{n1}, U_{n2})$$

$$+ L_{n2}(\varphi \,;\, U_{n1}, U_{n2}).$$

Since the estimate will be taken to be the solution of the (incorrect) system

$$(2.6) \qquad \frac{\partial}{\partial \varphi_i} L_n(\varphi) = 0, \qquad i = 1, 2,$$

we need some conditions which ensure the existence of the partial derivatives involved, and some regularity conditions slightly different from those are necessary to obtain the correct-model likelihood theory based on $\tilde{X}_{p_1,p_2}$ (c. f. Halperin, 1952, for correct model likelihood estimation). These assumptions are collected together in the following.

(For a simplicity in notation, the symbol $E^{\theta}_{u_1, u_2}$ denote the operation taking expected value under the truncated density $g(\cdot|\theta, u_1, u_2)$:

$$(2.7) \qquad g(x|\theta, u_1, u_2) \equiv \begin{cases} \dfrac{g(x|\theta)}{G(u_2|\theta) - G(u_1|\theta)} & ; \quad u_1 < x < u_2 \\ 0 & ; \quad \text{otherwise} \end{cases} .$$

We use the subscript notation for partial derivatives of $\log f(x|\varphi)$:

$$\psi_i(x, \varphi) = \frac{\partial}{\partial\varphi_i} \log f(x|\varphi), \text{ etc.}$$

Assume that for any $\theta^o \in \Theta$, with η_i defined by $p_i = \int_{-x}^{\eta_i} g(x|\theta^o)\,dx$, for $i = 1, 2$, there exists neighborhoods N of θ^o and S_i of η_i, $i = 1, 2$, with S_1 and S_2 disjoint such that:

(B 1) for almost all x, $\psi_i(x, \varphi)$, $\psi_{ij}(x, \varphi)$ and $\psi_{ijk}(x, \varphi)$ exist and are continuous for all $\varphi \in \Phi$;

(B 2) there exists an open set N_1 in Φ such that

$$\int_{-x}^{x} \sup_{\varphi \in N_1} |\psi_i(x, \varphi)|\, g(x|\theta^o)\,dx < \infty;$$

(B 3) there exists a unique solution $\varphi^o = \varphi^*(\theta^o)$ in N_1 of

(2.8) $$\lambda_i(\varphi|\theta^0, \eta_1, \eta_2) \equiv p_1 \frac{\partial}{\partial\varphi_i} \log F(\eta_1|\varphi) + q_2 \frac{\partial}{\partial\varphi_i} \log\left[1 - F(\eta_2|\varphi)\right] + \int_{\eta_1}^{\eta_2} \psi_i(x, \varphi)\, g(x|\theta^0)\, dx = 0, \quad i = 1, 2;$$

(B 4) there exists a function R and a constant M for all $\theta \in N$, $\varphi \in \Phi'$ where $\varphi^0 \in \Phi' \subset N_1$ such that

$$|\psi_i(x, \varphi)| < R(x), \quad |\psi_{ij}(x, \varphi)| < R(x),$$

$$|\psi_{ijk}(x, \varphi)| < R(x), \quad |\psi_i(x, \varphi)|^3 < R(x)$$

where

$$\int_{-x}^{x} R(x)\, g(x|\theta) < M;$$

(B 5) $f(x|\varphi)$, $\psi_i(x, \varphi)$, $\psi_{ij}(x, \varphi)$ $\psi_{ijk}(x, \varphi)$ are jointly continuous in x and φ for $x \in S_1 \cup S_2$ and $\varphi \in \Phi'$ and, $g(x|\theta)$ is jointly continuous in x and θ for $x \in S_1 \cup S_2$ and $\theta \in N$.

Addition to these assumptions, we have to impose conditions on some 2 x 2 matrices. Let

(2.9) $$A_{ij}(\varphi|\theta, u_1, u_2) = E^{\theta}_{u_1,u_2}\left[\psi_i(x, \varphi) \cdot \psi_j(x, \varphi)\right] - \left[E^{\theta}_{u_1,u_2}\left\{\psi_i(x, \varphi)\right\}\right]\left[E^{\theta}_{u_1,u_2}\left\{\psi_j(x, \varphi)\right\}\right],$$

(2.10) $$B_{ij}(\varphi|\theta, u_1, u_2) = -\int_{u_1}^{u_2} \psi_{ij}(x, \varphi) \cdot g(x|\theta)\, dx - p_1 \frac{\partial^2}{\partial\varphi_i \partial\varphi_j} \log F(u_1|\varphi) - q_2 \frac{\partial^2}{\partial\varphi_i \partial\varphi_j} \log\left[1 - F(u_2|\varphi)\right],$$

(B 6) The 2 x 2 matrices

$$A(\varphi^{o}|\theta^{o}, \eta_1, \eta_2) = \{A_{ij}(\varphi^{o}|\theta^{o}, \eta_1, \eta_2)\}$$

and

$$B(\varphi^{o}|\theta^{o}, \eta_1, \eta_2) = \{B_{ij}(\varphi^{o}|\theta^{o}, \eta_1, \eta_2)\}$$

are positive definite and the 2 x 2 matrix

$$\left\{E^{\theta^{o}}_{\eta_1,\eta_2}\left[\psi_{ij}(x, \varphi)\right]_{\varphi = \varphi^{o}}\right\}$$

is nonsingular.

In the statement of the following theorem, there appears a 2 x 2 matrix $K(\varphi^{o}|\theta^{o}, \eta_1, \eta_2)$ with (i, j) element $K_{ij}(\varphi^{o}|\theta^{o}, \eta_1, \eta_2)$:

(2.11)
$$K_{ij}(\varphi^{o}|\theta^{o}, \eta_1, \eta_2) = (1 - p_1 - q_2)\, A_{ij}(\varphi^{o}|\theta^{o}, \eta_1, \eta_2) + \sum_{r,t=1,2} l_{ir}\,(\varphi^{o})\, c_{rj},$$

where

$$l_{ik}(\varphi) = \frac{\partial}{\partial u_k}\left[\lambda_i\ (\varphi|\theta^{o}, u_1, u_2)\right]_{u_1 = \eta_1,\ u_2 = \eta_2}, \quad k = 1, 2,$$

and where

$$c_{rt} = \frac{p_r\ q_t}{g(\eta_r|\theta^{o})\ g(\eta_t|\theta^{o})}, \quad r \leq t,$$

$$c_{21} = c_{12}.$$

<u>Theorem 2</u>. If $\tilde{X}_{p_1,p_2}$ is for each n and for prescribed $0 < p_1 < p_2 < 1$, a trimmed sample from density $g(\cdot|\theta^{o})$; if the (correct) model $g(x|\theta)$ and the (incorrect) model $f(x|\varphi)$ satisfy assumptions (B 1) - (B 6), then for each sufficiently large n, the (incorrect) likelihood equation

$$\frac{\partial}{\partial \varphi_i} L_n(\varphi) = 0, \quad i = 1, 2$$

has a solution $\hat{\varphi}_n$ such that sequence $\{\hat{\varphi}_n\}$ has the following properties:

(i) $\hat{\varphi}_n \xrightarrow{p} \varphi^o = \varphi^*(\theta^o)$.

(ii) $\hat{\varphi}_n$ is a local maximum of $L_n(\varphi)$ with probability going to one.

(iii) If $\bar{\varphi}_n$ is any other sequence of solution of $\frac{\partial}{\partial\varphi} L_n(\varphi) = 0$ such that $\bar{\varphi}_n \xrightarrow{p} \varphi^o$, then $\hat{\varphi}_n = \bar{\varphi}_n$ with probability going to one as $n \longrightarrow \infty$.

(iv) $\sqrt{n}\ (\hat{\varphi}_n - \varphi^o) \xrightarrow{L} N_2\left[0,\ Q(\varphi^o|\theta^o,\ \eta_1,\ \eta_2)\right]$

where

$$Q(\varphi^o|\theta^o,\ \eta_1,\ \eta_2) = B^{-1}(\varphi^o|\theta^o,\ \eta_1,\ \eta_2)\cdot$$
$$\cdot\ K(\varphi^o|\theta^o,\ \eta_1,\ \eta_2)\ B^{-1}\ (\varphi^o|\theta^o,\ \eta_1,\ \eta_2).$$

To prove the theorem, two lemmas are given in the following.

<u>Lemma 1</u>. If assumptions (B 1) - (B 6) holds between $g(x|\theta)$ and $f(x|\varphi)$ for fixed $0 < p_1 < p_2 < 1$, then the central portion of the (incorrect) log-likelihood $L_{n1}(\varphi;\ W_n|U_{n1},\ U_{n2})$ has the following properties:

$$(2.12)\quad \left\{\frac{1}{\nu_2} \frac{\partial}{\partial\varphi_i} L_{n1}(\varphi;\ W_n|U_{n1},\ U_{n2})\right\}_{\theta^o} \xrightarrow{p} E^{\theta^o}_{\eta_1,\eta_2}\left[\psi_i(x,\ \varphi)\right],\quad i = 1,\ 2.$$

for sufficiently large n

$$(2.13)\quad \begin{pmatrix} \frac{1}{\nu_2} \frac{\partial}{\partial\varphi_1} L_{n1}(\varphi;\ W_n|U_{n1},\ U_{n2}) \\ \frac{1}{\nu_2} \frac{\partial}{\partial\varphi_2} L_{n1}(\varphi;\ W_n|U_{n1},\ U_{n2}) \end{pmatrix}_{\theta^o} = \begin{pmatrix} E^{\theta^o}_{\eta_1,\eta_2}\left[\psi_1(x,\ \varphi)\right] \\ E^{\theta^o}_{\eta_1,\eta_2}\left[\psi_2(x,\ \varphi)\right] \end{pmatrix}$$
$$+ \frac{1}{\sqrt{n}} \begin{pmatrix} Q^1_{10}\ X_{n1} + Q^1_{20}\ X_{n2} \\ Q^2_{10}\ X_{n1} + Q^2_{20}\ X_{n2} \end{pmatrix}$$
$$+ \frac{1}{\sqrt{n}} \left[A(\varphi|\theta^o,\ \eta_1,\ \eta_2)\right]^{\frac{1}{2}} \begin{pmatrix} Y^1_n \\ Y^2_n \end{pmatrix}$$

where

$$Q_{ko}^{i} = \frac{\partial}{\partial u_k} E^{\theta^o}_{u_1,u_2}\left[\psi_i(x, \varphi)\right]_{u_1=\eta_1,\ u_2=\eta_2}, \ k = 1, 2,$$

and X_{n1}, X_{n2}, Y_n^1 and Y_n^2 are random variables such that

$$\begin{pmatrix} Y_n^1 \\ Y_n^2 \\ X_{n1} \\ X_{n2} \end{pmatrix} \xrightarrow{L} N_4 \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{1-p_1-q_2} & 0 & 0 & 0 \\ 0 & \frac{1}{1-p_1-q_2} & 0 & 0 \\ 0 & 0 & C_{11} & C_{12} \\ 0 & 0 & C_{12} & C_{22} \end{pmatrix} \right];$$

and for each i and j

(2.14)
$$\left[\frac{1}{\nu_2} \frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_{n1}(\varphi ; W_{n1}, U_{n2})\right]_{\theta^o}$$

$$\xrightarrow{p} E^{\theta^o}_{\eta_1,\eta_2}\left[\psi_{ij}(x, \varphi)\right], \quad i, j = 1, 2.$$

Proof of Lemma 1.

In addition to assumptions (B 1) - (B 6), if we assume that there exists a positive constant m for $\theta \in N$ and $u_i \in S_i$, $i = 1, 2$ such that $[G(u_2|\theta) - G(u_1|\theta)]^{-1} < m$, then we have by (B 4)

$$E^{\theta}_{u_1,u_2}|\psi_i(x, \varphi)| < m\ M, \qquad E^{\theta}_{u_1,u_2}|\psi_i(x, \varphi)\ \psi_j(x, \varphi)| < m\ M$$

and

$$\beta^{\theta}_{u_1,u_2}\left[\psi_i(x, \varphi)\right] \equiv E^{\theta}_{u_1,u_2}\left|\psi_i(x, \varphi) - E^{\theta}_{u_1,u_2}\left[\psi_i(x, \varphi)\right]\right|^3 < 8\ m\ M.$$

Moreover, by assumption (B 5), $\beta^{\theta}_{u_1,u_2}\left[\psi_i(z, \varphi)\right]$ is jointly continuous in θ, φ, u_1 and u_2 for $\theta \in N$, $\varphi \in \Phi'$ and $u_i \in S_i$, $i = 1, 2$. Also covariance

$$A_{ij}(\varphi|\theta, u_1, u_2) \equiv E^{\theta}_{u_1,u_2}\left[\psi_i(x, \varphi) \cdot \psi_j(x,\varphi)\right]$$

$$- E^{\theta}_{u_1,u_2}\left[\psi_i(x, \varphi)\right] E^{\theta}_{u_1,u_2}\left[\psi_j(x, \varphi)\right]$$

is a jointly continuous function in u_1, u_2, θ and φ. Thus, the determinant of the matrix $A(\varphi|\theta, u_1, u_2)$, being a continuous function of it's entries, is also jointly continuous in u_1, u_2, θ and φ. Hence, there exists $\gamma < \infty$, $\beta > 0$ such that

$$\text{Det. } A(\varphi|\theta, u_1, u_2) > \beta$$

(2.15)

$$\beta^{\theta}_{u_1,u_2}\left[\psi_i(z, \varphi)\right] < \gamma$$

for all $\theta \in N'$, $\varphi \in \Phi''$ and $u_1 \in S_i'$, $i = 1, 2$ where

$$\theta^{o} \in N' \subset \overline{N'} \subset N, \qquad \varphi^{o} \in \Phi'' \subset \overline{\Phi''} \subset \Phi'$$

and

$$\eta_i \in \overline{S_i'} < S_i, \quad i = 1, 2.$$

Since the determinant of the covariance matrix $A(\varphi|\theta, u_1, u_2)$ is bounded away from zero, the matrix is positive definite for all $\theta \in N'$, $\varphi \in \Phi''$ and $u_1 \in S_i'$, $i = 1, 2$.

Hence, Theorem 1 can be applied to

$$\frac{\partial}{\partial\varphi_i} L_{1n}(\varphi;\, W_n \mid U_{n1},\, U_{n2}) = \sum_{k\in I_2} \psi_i\,(x_k,\, \varphi)$$

and the first half of the lemma now follows.

By assumptions (B 4), we find

$$E^{\theta}_{u_1,u_2}\, \psi_{ij}(x,\, \varphi) < m\, M$$

for all $\theta \in N$, $\varphi \in \Phi'$ and $(u_1,\, u_2) \in S_1 \times S_2$.

Hence, if $(Z_1,\, Z_2,\, \ldots,\, Z_{\nu_2})$ is a random sample of size ν_2 from $g(\cdot|\theta^o,\, u_1,\, u_2)$, then the relation

$$\frac{1}{\nu_2}\Big[\sum_{k=1}^{\nu_2} \psi_{ij}\,(Z_k,\, \varphi)\Big]_{\theta^o} \xrightarrow{p} E^{\theta^o}_{u_1,u_2}\Big[\psi_{ij}\,(x,\, \varphi)\Big], \quad i,\, j = 1,\, 2$$

holds uniformly in $(u_1,\, u_2) \in S_1 \times S_2$. Thus, conditionally on $U_{n1} = u_2$ $U_{n2} = u_2$, we have a conditional convergence for the order statistics X_k, $k \in I_2$:

$$\frac{1}{\nu_2}\Big[\sum_{k\in I_2} \psi_{ij}\,(x_k,\, \varphi)\Big]_{\theta^o} \Big|\; U_{n1} = u_1,\;\; U_{n2} = u_2$$

$$\xrightarrow{p} E^{\theta^o}_{u_1,u_2}\Big[\psi_{ij}\,(x,\, \varphi)\Big], \quad i,\, j = 1,\, 2.$$

Denote a 2×2 matrix with (i, j) element t_{ij} by $\{t_{ij}\}$. Since $\{E^{\theta}_{u_1,u_2}[\psi_{ij}(x, \varphi)]\}$ is non-singular at $\theta = \theta^{\circ}$, $\varphi = \varphi^{\circ}$ and $u_1 = \eta_i$, $i = 1, 2$, and, since $E^{\theta}_{u_1,u_2}[\psi_{ij}(x, \varphi)]$ is continuous in θ, φ and u_1, $i = 1, 2$, in the respective neighborhoods of θ°, φ° and η_i, $i = 1, 2$, there exist neighborhoods N', Φ'' and S_i' of θ°, φ°, and η_i respectively such that $\{E^{\theta}_{u_1,u_2}[\psi_{ij}(x, \varphi)]\}$ is nonsingular for all $\theta \in N'$, $\varphi \in \Phi''$ and $(u_1, u_2) \in S_1' \times S_2'$. Thus, the inverse matrix $\{E^{\theta}_{u_1,u_2}[\psi_{ij}(x, \varphi)]\}^{-1}$ exists, and it follows from (2.16) that

$$(2.17) \qquad \left\{E^{\theta}_{u_{n1},u_{n2}}\left[\psi_{ij}(x, \varphi)\right]\right\}^{-1} \left\{\frac{1}{\nu_2} \sum_{k \in I_2} \psi_{ij}(x_k, \varphi)\right\}_{\theta^{\circ}} \Big| U_{n1} = u_1, U_{n2} = u_2$$

$$\xrightarrow{p} I.$$

Now, we obtain the unconditional convergence applying Theorem 2 of Appendix to (2.17):

(2.18)

$$\left\{E^{\theta^{\circ}}_{u_{n1},u_{n2}}\left[\psi_{ij}(x, \varphi)\right]\right\}^{-1} \left\{ \frac{1}{\nu_2} \frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_{n1}\ (\varphi;\ W_n | U_{n1}, U_{n2})\right\}$$

$$\xrightarrow{p} I.$$

We now have the second half of the lemma by applying the property,

$$U_{n1} \xrightarrow{p} \eta_1\ , \quad U_{n2} \xrightarrow{p} \eta_2$$

Lemma 2. If assumptions (B 1) - (B 6) hold between $g(\cdot|\theta)$ and $f(\cdot|\varphi)$ for fixed $0 < p_1 < p_2 < 1$, then the remaining portion $L_{n2}(\varphi;\ U_{n1},\ U_{n2})$ of the (incorrect) log-likelihood has the following properties:

$$\frac{1}{n}\left[\frac{\partial}{\partial\varphi_1} L_{n2}(\varphi;\ U_{n1},\ U_{n2})\right]_{\theta^\circ} \xrightarrow{p} \left[\frac{\partial}{\partial\varphi_i} T(\varphi|\theta^\circ,\ \eta_1,\ \eta_2)\right], \quad i = 1,\ 2,$$

where

$$T\ (\varphi|\theta,\ u_1,\ u_2) = p_1 \log F\ (u_1|\varphi) + q_2 \log\left[1 - F\ (u_2|\varphi)\right],$$

and, for sufficiently large n,

$$\begin{pmatrix} \frac{1}{n}\frac{\partial}{\partial\varphi_1} L_{n2}(\varphi;\ U_{n1},\ U_{n2}) \\ \\ \frac{1}{n}\frac{\partial}{\partial\varphi_2} L_{n2}(\varphi;\ U_{n1},\ U_{n2}) \end{pmatrix}_{\theta^\circ} = \begin{pmatrix} \frac{\partial}{\partial\varphi_1} T\ (\varphi|\theta^\circ,\ \eta_1,\ \eta_2) \\ \\ \frac{\partial}{\partial\varphi_2} T\ (\varphi|\theta^\circ,\ \eta_1,\ \eta_2) \end{pmatrix}$$

$$+ \frac{1}{\sqrt{n}} \begin{pmatrix} \frac{\partial}{\partial u_1}\left[\frac{\partial}{\partial\varphi_1} T(\varphi|\theta^\circ,u_1,u_2)\right] & \frac{\partial}{\partial u_2}\left[\frac{\partial}{\partial\varphi_1} T(\varphi|\theta^\circ,u_1,u_2)\right] \\ \\ \frac{\partial}{\partial u_1}\left[\frac{\partial}{\partial\varphi_2} T(\varphi|\theta^\circ,u_1,u_2)\right] & \frac{\partial}{\partial u_2}\left[\frac{\partial}{\partial\varphi_2} T(\varphi|\theta^\circ,u_1,u_2)\right] \end{pmatrix} \begin{pmatrix} X_{n1} \\ \\ X_{n2} \end{pmatrix}$$

where random variables X_{n_i}, $i = 1,\ 2$ are defined in Lemma 1;

and for each i and j

$$\frac{1}{n}\left[\frac{\partial}{\partial\varphi_i\partial\varphi_j} L_{n2}(\varphi; U_{n1}, U_{n2})\right]_{\theta^o} \xrightarrow{p} \left[\frac{\partial}{\partial\varphi_i\partial\varphi_j} T(\varphi|\theta^o, \eta_1, \eta_2)\right].$$

Proof of Lemma 2. By definition (2.4) $\frac{1}{n}\frac{\partial}{\partial\varphi_i} L_{n2}(\varphi; U_{n1}, U_{n2})$ can be expressed as

$$(2.19) \qquad \frac{1}{n}\frac{\partial}{\partial\varphi_i} L_{n2}(\varphi; U_{n1}, U_{n2})$$

$$= \frac{\nu_1}{n}\frac{\partial}{\partial\varphi_i} \log F(u_{n1}|\varphi) + \frac{\nu_3}{n}\frac{\partial}{\partial\varphi_i} \log\left[1 - F(u_{n2}|\varphi)\right]$$

$$+ \frac{1}{n}\frac{\partial}{\partial\varphi_i}\left[f(u_{n1}|\varphi)\ f(u_{n2}|\varphi)\right].$$

Since $f(x|\varphi)$ is continuous at $x = \eta_1$ and $x = \eta_2$, we have

$$\lim_{n\to\infty} \frac{1}{n}\frac{\partial}{\partial\varphi_i} \log\left[f(u_{n1}|\varphi)\ f(u_{n2}|\varphi)\right] = 0.$$

Let random variable X_{ni}, $i = 1, 2$, be such that

$$X_{ni} = \sqrt{n}\ (U_{ni} - \eta_i).$$

By expanding $\frac{\partial}{\partial\varphi_i} \log F(u_{n1}|\varphi)$ and $\frac{\partial}{\partial\varphi_i} \log\left[1 - F(u_{n2}|\varphi)\right]$ in x_{ni} near zero, we have

$$\frac{\partial}{\partial\varphi_i} \log F(\eta_i + \frac{x_{n1}}{\sqrt{n}} |\varphi) = \frac{\partial}{\partial\varphi_i} \log F(\eta_1|\varphi)$$

$$+ \frac{x_{n1}}{\sqrt{n}} \left[\frac{\partial}{\partial u_1} \frac{\partial}{\partial\varphi_i} \log F(u_1|\varphi\right]_{u_1=\eta_1} + o(\frac{1}{\sqrt{n}}),$$

$$\frac{\partial}{\partial\varphi_i} \log \left[1 - F(\eta_2 + \frac{x_{n2}}{\sqrt{n}} |\varphi\right] = \frac{\partial}{\partial\varphi_i} \log \left[1 - F(\eta_2|\varphi)\right]$$

$$+ \frac{x_{n2}}{\sqrt{n}} \left[\frac{\partial}{\partial u_2} \frac{\partial}{\partial\varphi_i} \log \left(1 - F(u_2|\varphi)\right) \right]_{u_2=\eta_2} + o(\frac{1}{\sqrt{n}}).$$

Hence

$$\frac{1}{n} \frac{\partial}{\partial\varphi_i} L_{n2} (\varphi ; U_{n1}, U_{n2})$$

$$= \frac{\nu_1}{n} \frac{\partial}{\partial\varphi_i} \log F(\eta_i|\varphi) + \frac{\nu_3}{n} \frac{\partial}{\partial\varphi_i} \log \left[1 - F(\eta_2|\varphi)\right]$$

$$+ \frac{\nu_1}{n} \frac{X_{n1}}{\sqrt{n}} \left[\frac{\partial}{\partial u_1} \frac{\partial}{\partial\varphi_i} \log F(u_1|\varphi)\right]_{u_1=\eta_1}$$

$$+ \frac{\nu_3}{n} \frac{X_{n2}}{\sqrt{n}} \left[\frac{\partial}{\partial u_2} \frac{\partial}{\partial\varphi_i} \log \left(1 - F(u_2|\varphi)\right) \right]_{u_2=\eta_2} + O(\frac{1}{\sqrt{n}}).$$

As n tends to ∞, $\frac{\nu_1}{n} \longrightarrow p_1$ and $\frac{\nu_3}{n} \longrightarrow q_2$, and, the first half now follows. The second half can be shown similarly by expanding $\frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_{n2} (\varphi ; U_{n1}, U_{n2})$ around η_1 and η_2.

We are now ready to prove Theorem 2.

Proof of Theorem 2. Applying the results of Lemmas 1 and 2 to

$$\frac{1}{n}\frac{\partial}{\partial\varphi_i} L_n (\varphi) = \frac{1}{n}\frac{\partial}{\partial\varphi_i} L_{n1} (\varphi ; W_n | U_{n1}, U_{n2})$$

$$+ \frac{1}{n}\frac{\partial}{\partial\varphi_i} L_{n2} (\varphi ; U_{n1}, U_{n2}),$$

we obtain

$$\frac{1}{n}\left[\frac{\partial}{\partial\varphi_i} L_n (\varphi)\right]_{\theta^o} \xrightarrow{p} \int_{\eta_1}^{\eta_2} \psi_i (x, \varphi)\, g(x|\theta^o)\, dx$$

$$+ p_1 \frac{\partial}{\partial\varphi_i} \log F(\eta_1|\varphi) + q_2 \frac{\partial}{\partial\varphi_i} \log \left[1 - F(\eta_2|\varphi)\right]$$

$$\equiv \lambda_i (\varphi|\theta^o, \eta_1, \eta_2), \quad i = 1, 2.$$

However, by assumption (B 3), φ^o is an unique solution such that $\lambda_i (\varphi^o|\theta^o, \eta_1, \eta_2) = 0$. Hence,

(2.20) $$\frac{1}{n}\left[\frac{\partial}{\partial\varphi_i} L_n (\varphi)\right]_{\varphi^o, \theta^o} \xrightarrow{p} 0, \quad i = 1, 2.$$

Moreover, by Lemmas 1 and 2, we find

$$(2.21) \qquad \frac{1}{\sqrt{n}} \begin{pmatrix} \frac{\partial}{\partial \varphi_1} L_n(\varphi) \\ \\ \frac{\partial}{\partial \varphi_2} L_n(\varphi) \end{pmatrix}_{\varphi^o, \theta^o} \xrightarrow{L} N_2 \left[0, K(\varphi^o | \theta^o, \eta_1, \eta_2) \right]$$

where the (ij) element of the covariance matrix $K(\varphi^o|\theta^o, \eta_1, \eta_2)$ is given by (2.11). Since the matrix $A(\varphi^o|\theta^o, \eta_1, \eta_2)$ is positive definite by (B 6), $K(\varphi^o|\theta^o, \eta_1, \eta_2)$ is also a positive definite matrix. Also it follows from Lemmas 1 and 2 that

$$(2.22) \qquad \left[\frac{1}{n} \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} L_n(\varphi)\right]_{\varphi^o, \theta^o} \xrightarrow{p} - B_{ij}(\varphi^o|\theta^o, \eta_1, \eta_2), \quad ij = 1, 2$$

where $B_{ij}(\varphi^o|\theta^o, \eta_1, \eta_2)$ is given by (2.10).

Now, if $\varphi \in \Phi'$, then by the mean value theorem

$$(2.23) \qquad \frac{1}{n} \frac{\partial}{\partial \varphi_i} L_n(\varphi) = \frac{1}{n} \left[\frac{\partial}{\partial \varphi_i} L_n(\varphi)\right]_{\varphi^o} + \sum_j (\varphi - \varphi^o) \cdot \frac{1}{n} \cdot \left[\frac{\partial^2}{\partial \varphi_i \partial \varphi_j} L_n(\varphi)\right]_{\varphi^o}$$

$$+ \alpha |\varphi - \varphi^o| \; H(u_{n1}, w_n, u_{n2})$$

where $|\alpha| \leq 2$, $|\varphi - \varphi^o|$ denotes length in E^2. The function $H(u_{n1}, w_{n1}, u_{n2})$ is such that

$$(2.24) \qquad \left| \frac{1}{n} \frac{\partial^3}{\partial\varphi_i \partial\varphi_j \partial\varphi_k} L_n(\varphi) \right|_{\varphi'} < H\,(u_{n1}, w_n, u_{n2}),\ \varphi' \in \Phi'$$

and

$$(2.25) \qquad H\,(U_{n1}, W_n, U_{n2}) \xrightarrow{p} M$$

where M is positive constant. Such a function H exists by the assumption (B 4). Since the matrix $B\,(\varphi^o|\theta^o, \eta_1, \eta_2)$ is positive definite by (B 6), there exists $\alpha' > 0$ such that

$$t'\,B\,(\varphi^o|\theta^o, \eta_1, \eta_2)\,t \geq \alpha'$$

for any vector t in E^2 such that $|\,t\,| = 1$. For given $\epsilon > 0$ choose $\delta > 0$ in such a way that

$$(2.26) \qquad \delta < \epsilon,\ \{\varphi : |\varphi - \varphi^o| \leq \delta\} \subset \Phi',\ \delta < \frac{\alpha}{10(M+1)}\,.$$

Then, by (2.20), (2.22), and (2.25), for large $n_o(\epsilon)$ and if $n > n_o(\epsilon)$,

$$\left| \frac{1}{n} \frac{\partial}{\partial\varphi_i} L_n\,(\varphi) \right|_{\varphi^o} < \delta^2,$$

$$\left| \frac{1}{n} \left[\frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_n(\varphi)\right]_{\varphi^o} - B_{ij}\,(\varphi^o|\theta^o, \eta_1, \eta_2) \right| < \delta,$$

$$o < H\,(u_{n1}, w_n, u_{n2}) < M + 1$$

with probability exceeding $(1 - \epsilon)$. If $|\varphi - \varphi^{o}| < \delta$, then from (2.23),

$$\left| \frac{1}{n} \frac{\partial}{\partial \varphi_i} L_n(\varphi) + \sum_j B_{ij} (\varphi^{o} | \theta^{o}, \eta_1, \eta_2) \cdot (\varphi_j - \varphi_j^{o}) \right|$$

$$\leq \delta^2 + 2\delta |\varphi - \varphi^{o}| + 2 |\varphi - \varphi^{o}|^2 (M + 1) \leq 5\delta^2 (M + 1).$$

Hence, if $|\varphi - \varphi^{o}| = \delta$, then,

$$\sum_i \left[\frac{1}{n} \frac{\partial}{\partial \varphi_i} L_n (\varphi)\right] \cdot (\varphi_i - \varphi_i^{o})$$

$$\leq \sum_{i,j} B_{ij} (\varphi^{o} | \theta^{o}, \eta_1, \eta_2) \cdot (\varphi_i - \varphi_i^{o})(\varphi_j - \varphi_j^{o})$$

$$+ 10 \delta^3 (M + 1) \leq - \beta |\varphi - \varphi^{o}| + 10\delta^3 (M + 1)$$

$$= - \beta\delta^2 + 10\delta^3 (M + 1) < 0.$$

It follows from Lemma 2 of Aitchson and Silvey (1958), that there is a value $\hat{\varphi}_n$ of φ which satisfies

$$\frac{\partial}{\partial \varphi} L_n (\varphi) = 0$$

and

$$| \hat{\varphi}_n - \varphi^{o} | \leq \delta.$$

Since ϵ and δ were chosen arbitrarily, this proves the existence of the solution $\hat{\varphi}_n$ of the incorrect likelihood equation $\frac{\partial}{\partial \varphi} L_n(\varphi) = 0$ such that $\hat{\varphi}_n \xrightarrow{p} \varphi^{o}$.

Using the mean value theorem again, if $\varphi \in \Phi'$

$$(2.27) \quad \frac{1}{n} \frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_n(\varphi) = \frac{1}{n} \left[\frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_n(\varphi)\right]_{\varphi^o} + \alpha \, |\varphi - \varphi'| \, H \, (u_{n1}, w_n, u_{n2}).$$

Choose δ' so small that $\{\varphi : |\varphi - \varphi^o| < \delta'\} \subset \Phi'$, and so small that for any 2×2 symmetric matrix $\{b_{ij}\}$ with

$$| \, b_{ij} - B_{ij} \, (\varphi^o | \theta^o, \eta_1, \eta_2) \, | < \delta' + 2\delta'(M + 1)$$

is positive definite. For thus chosen δ', we have by (2.20), (2.22) and (2.25),

$$\left| \, \frac{1}{n} \left[\frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_n \, (\varphi)\right]_{\varphi^o} + B_{ij} \, (\varphi' | \theta^o, \eta_1, \eta_2) \, \right| < \delta'$$

and

$$0 < H \, (u_{n1}, w_n, u_{n2}) < M + 1$$

with probability going to one. Hence by (2.27), for all $\varphi \in \{\varphi : |\varphi - \varphi'| < \delta'\}$ we find

$$\left| \, \frac{1}{n} \frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_n(\varphi) + B_{ij}(\varphi^o | \theta^o, \eta_1, \eta_2) \, \right| < \delta' + 2\delta' \, (M + 1).$$

Therefore, the matrix $\left\{\frac{\partial^2}{\partial\varphi_i \partial\varphi_j} L_n(\varphi)\right\}$ is negative definite for every φ such that $|\varphi - \varphi'| < \delta'$ and it follows that $\frac{\partial}{\partial\varphi} L_n(\varphi) = 0$ has at most one solution with $|\varphi - \varphi^o| < \delta'$ and $L_n(\varphi)$ achieves maximum.

Now there remains the proof of the limiting distribution of $\sqrt{n} \, (\hat{\varphi}_n - \varphi^o)$. Since $\hat{\varphi}_n$ lies in Φ' with probability going to one, we have by (2.23),

$$\frac{1}{\sqrt{n}}\left[\frac{\partial}{\partial\varphi_i} L_n(\varphi)\right]_{\varphi^o} + \sum_j \sqrt{n}\ (\hat{\varphi}_j - \varphi_j^o)\left[\frac{1}{n}\frac{\partial^2}{\partial\varphi_i\partial\varphi_j} L_n(\varphi)\right]_{\varphi^o}$$

$$+ \alpha|\hat{\varphi}_n - \varphi^o| \cdot |\sqrt{n}\ (\hat{\varphi}_n - \varphi^o)|\ H\ (u_{n1}, w_n, u_{n2}) = 0, \quad i = 1, 2.$$

From (2.20), (2.22), and the fact that $\hat{\varphi}_n \xrightarrow{p} \varphi^o$,

$$\left|\ \frac{1}{n}\left(\frac{1}{\partial\varphi} L_n(\varphi)\right)_{\varphi^o} - B\ (\varphi^o|\theta^o, \eta_1, \eta_2) \cdot \sqrt{n}\ (\hat{\varphi}_n - \varphi^o)\ \right|$$

$$\leq \epsilon_n |\ \sqrt{n}\ (\hat{\varphi}_n - \varphi^o)|,$$

where $\epsilon_n \xrightarrow{p} 0$. Moreover, we have the relation (2.21), and it follows from Theorem 10.1 of Billingsley (1961) that

$$B(\varphi^o|\theta^o, \eta_1, \eta_2) \cdot \sqrt{n}\ (\hat{\varphi}_n - \varphi^o)$$

$$\xrightarrow{L} N_2\left[0, K(\varphi^o|\theta^o, \eta_1, \eta_2)\right].$$

Hence,

$$\sqrt{n}\ (\hat{\varphi}_n - \varphi^o) \xrightarrow{L} N_2\left[0, Q(\varphi^o|\theta^o, \eta_1, \eta_2)\right]$$

where

$$Q(\varphi^o|\theta^o, \eta_1, \eta_2) = B^{-1}(\varphi^o|\theta^o, \eta_1, \eta_2)\ K(\varphi^o|\theta^o, \eta_1, \eta_2)\ B^{-1}(\varphi^o|\theta^o, \eta_1, \eta_2)$$

and the proof is completed.

Appendix

Two theorems on joint convergence of distributions are discussed in this appendix. These are used in the proofs of Theorems 1 and 2.

Let a random vector X_n have c.d.f $F_n(\cdot)$, and, let a random vector Y_n given $X_n = x$ have c.d.f $F_n(\cdot|x)$ for all x in the union of the closure of the ranges of X_n. Let $F_o(\cdot)$ be a c.d.f, and, let $F_o(\cdot|x)$ be a c.d.f. for all x. Let

$$F_n(x, y) = \int_{-\infty}^{x} F_n(y|z)d\,F_n(z),$$

$$F_o(x, y) = \int_{-\infty}^{x} F_o(y|z)d\,F_o(z).$$

Then $F_n(\cdot,\cdot)$ and $F_o(\cdot,\cdot)$ are c.d.f.'s and $F_n(\cdot,\cdot)$ is the joint c.d.f of (X_n, Y_n). In the following theorem we write $F_n(\cdot) \Rightarrow F_o(\cdot)$ to indicate that a sequence $F_n(\cdot)$ converges weakly to $F_o(\cdot)$.

Theorem 1. If (i) $F_o(y|x)$ is continuous in x and y separately for all (x, y); (ii) $F_n(\cdot) \Rightarrow F_o(\cdot)$; (iii) $F_n(\cdot|x) \Rightarrow F_o(\cdot|x)$ for all x; and, (iv) for each y, convergence in (iii) is uniform in x for x in the open sets of arbitrarily large probability under F_o, then

$$F_n(\cdot,\cdot) \Rightarrow F_o(\cdot,\cdot).$$

Proof.

$$F_n(x, y) - F_o(x, y) = Q_1 + Q_2$$

where

$$Q_1 = \int_{-\infty}^{x} \left[F_n(y|z) - F_o(y|z) \right] d\, F_n(z)$$

$$Q_2 = \int_{-\infty}^{x} F_o(y|z)\, d\, F_n(z) - \int_{-\infty}^{x} F_o(y|z)\, d\, F_o(z).$$

Fix y, and for given $\epsilon > 0$, let A be an open set with

$$\int_A d\, F_o > 1 - \frac{\epsilon}{3}$$

for which the uniform convergence holds. Then, for $n > n(y, \epsilon, A)$,

$$|F_n(y|x) - F_o(y|x)| < \frac{\epsilon}{3} \quad \text{for all } x \epsilon A.$$

But

$$F_n(\cdot) \Rightarrow F_o(\cdot)$$

if and only if

$$\lim_{n \to \infty} \inf \int_A d\, F_n \geq \int_A d\, F_o$$

for every open set A. Hence we can choose $n > N^* (\epsilon, A)$, so that

$$\int_A d\, F_n \geq 1 - \frac{2}{3}\, \epsilon\ .$$

Then

$$Q_1 = \int_{-\infty}^{x} \left[F_n(y|z) - F_o(y|z)\right] d\,F_n(z)$$

$$= \int_A \left[F_n(y|z) - F_o(y|z)\right] d\,F_n(z)$$

$$+ \int_{A^c} \left[F_n(y|z) - F_o(y|z)\right] d\,F_n(z)$$

$$< \epsilon \int_A d\,F_n(z) + \int_{A^c} d\,F_n(z) < \epsilon.$$

Moreover, $F_o(y|x)$ is a bounded and continuous function, and $F_n(\cdot) \Rightarrow F_o(\cdot)$ by (ii). Hence, it follows from the Helly-Bray lemma

$$\lim_{n \to \infty} \int F_o(y|x)\, d\,F_n(x) = \int F_o(y|x)\, d\,F_o(x).$$

Therefore $Q_2 \longrightarrow 0$ as $n \longrightarrow x$, and thus,

$$F_n(x, y) \Rightarrow F_o(x, y).$$

Theorem 2. For each n, let (U_n, W_n) be a pair of random vectors. Let (X_o, Y_o) be a pair of independent random vectors and let u_o be a real vector and $w_o(u)$ a real valued function on the space of u_o. If (i) $U_n \xrightarrow{p} u_o$; (ii) $X_n = \sqrt{n}\,(U_n - u_o) \xrightarrow{L} X_o$; moreover, for all u in some neighborhood S of u_o, if (iii) the conditional distribution of $W_n|U_n = u$ is defined for all $u \in S$; (iv) $W_n|U_n = u \xrightarrow{p} w_o(u)$; (v) $\sqrt{n}\, a(u)\, \big(W_n - w_o(u)\big) \mid U_n = u \xrightarrow{L} Y_o$; (vi) the convergence of the c.d.f of $W_n|U_n = u$ is uniform in u for $u \in S$ at each continuity point of the c.d.f of Y_o, then

$$(X_n, Y_n) \xrightarrow{L} (X_o, Y_o)$$

where

$$Y_n = \sqrt{n}\ a(U_n)(W_n - w_o(u_n)).$$

Proof.

At first we will show that $Y_n \xrightarrow{L} Y_o$.

Let $G_o(y) = P(Y_o \leq y)$ be the limiting conditional c.d.f. of Y_o which does not depends on u. Let y be any fixed continuity point of G_o.

$$| P(Y_n \leq y) - G_o(y) | = | \int P(Y_n \leq y|U_n = u)\ d\ F_{u_n}(u)$$

$$- \int P(Y_o \leq y|U_n = u)\ d\ F_{u_n}(u)|$$

$$= |\int P(Y_n \leq y|U_n = u)\ d\ F_{u_n}(u)$$

$$- \int P(Y_o \leq y)\ d\ F_{u_n}(u)|$$

$$\leq \int |P(Y_n \leq y|U_n = u) - P(Y_o \leq y)|d\ F_{u_n}(u).$$

For given $\epsilon > 0$ and a neighborhood S of u_o, if $n > n_o'(\epsilon, S)$, we have

$$\int_S d\ F_{u_n}(u) > 1 - \frac{2}{3}\epsilon.$$

Since the convergence of $W_n | U_n = u$ is uniform in $u \in S$ by (vi)

$$|P(Y_n \leq y | U_n = u) - P(Y_o \leq y)| \leq \frac{\epsilon}{3}$$

for all $u \in S$ and $n > n_o''(\epsilon, S)$. Therefore if $n > \max(n_o', n_o'')$

$$|P(Y_n \geq y) - G_o(y)| \leq \int_S |P(Y_n \leq y | U_n = u) - P(Y_o \leq y)| d F_{u_n}(u)$$

$$+ \int_{S^c} |P(Y_n \leq y | U_n = u) - P(Y_o \leq y)| d F_{u_n}(u)$$

$$< \frac{1}{3} \epsilon \int_S d F_{u_n}(u) + \int_{S^c} d F_{u_n}(u) < \epsilon.$$

Now, it remains to show the asymptotic independence of X_n and Y_n.

$$P(X_n \leq x, Y_n \leq y) = \int P(X_n \leq x, Y_n \leq y | U_n = u) d F_{u_n}(u)$$

$$= \int_{-\infty}^{u_o + x/\sqrt{n}} P(Y_n \leq y | U_n = u) d F_{u_n}(u) = P(X_o \leq x) G_o(y)$$

$$+ \int_{-\infty}^{u_o + x/\sqrt{n}} \left[P(Y_n \leq y | U_n = u) - G_o(y) \right] d F_{u_n}(u)$$

$$\xrightarrow[n]{} P(X_o \leq x) G_o(y)$$

Uniformly in u. Thus,

$$(X_n, Y_n) \xrightarrow{L} (X_o, Y_o).$$

References

[1] AITCHSON, J. and SILVEY, S.D. (1958). Maximum likelihood estimation of parameters subject to restraints. Ann. Math. Statist. 29, 813-828.

[2] BERGSRÖM, H. (1945). On the general limit theorem in the space of R_k, $k > 1$. Skandinavisk Actuarietidskrift, 28, 106-127.

[3] BILLINGSLEY, P. (1961). Statistical Inference for Markov Processes. Univ. Chicago Press, Chicago.

[4] CRAMÉR, H (1958). Mathematical Methods of Statistics. Princeton Univ. Press, Princeton.

[5] HALPERIN, M. (1952). Maximum likelihood estimation in truncated samples. Ann. Math. Statist. 23, 226-238.

[6] LOÉVE, M. (1960). Probability Theory (2nd ed.). D. Van Nostrand, New York.