ASYMPTOTIC (BETA) SUB D JOINT NORMALITY OF SAMPLE QUANTILES.
ASYMPTOTIC JOINT NORMALITY OF SAMPLE QUANTILES* 

By 
Sadao Aka 
Department of Economics, Soka University 
and 
Department of Statistics, Purdue University 
Mimeograph Series, B1-20 

Department of Statistics 
Division of Mathematical Sciences 
June 1981 

*) The writing of this paper was facilitated while the author was visiting Purdue University in part under the ONR Contract N00014-75-C-0455.
ASYMPTOTIC $(\mu)_d$ JOINT NORMALITY OF SAMPLE QUANTILES*

By

Sadao Ikeda
Dept. of Economics, Soka Univ., and Dept. of Statistics, Purdue Univ.

Summary

Notions of asymptotic equivalence of probability distributions and some of their properties are briefly presented. By applying the results on type $(\mu)_d$ asymptotic equivalence, asymptotic $(\mu)_d$ joint normality of a set of increasing number of sample quantiles are discussed, which improves and refines the previous work by Ideda and Matsunawa (1972).

1. Introduction

Consider two sequences of random variables $(X_t; t \to 0)$ and $(Y_t; t \to 0)$, $X_t$ and $Y_t$ belonging to $\mathcal{P}(R_t, \mathcal{B}_t)$ for each $t$, where $t$ is a parameter taking values in a given metric space, $R_t$ is any given abstract space, $\mathcal{B}_t$ is a $\sigma$-field of subsets of $R_t$, and finally $\mathcal{P}(R_t, \mathcal{B}_t)$ designates the class of all random variables distributed over the measurable space $(R_t, \mathcal{B}_t)$.

Let us consider a sequence, $(C_t; t \to 0)$, of subclasses of corresponding $R_t$'s, and denote it by $(c)$.

*The writing of this paper was facilitated while the author was visiting Purdue University in part under the ONR Contract N00014-75-C=0455.

Keywords: Asymptotic theory, order statistics, normal distribution.
Under the above situation, two kinds of notions are defined:

Definition 1.1 (a) Two sequences of random variables \( \{X_t; t \gg t_0\} \) and \( \{Y_t; t \gg t_0\} \), are said to be asymptotically equivalent in the sense of type \((\text{C})_d\), and briefly denoted as

\[
X_t \sim Y_t (\text{C})_d, (t \gg t_0),
\]

provided that the following condition holds:

\[
(1.2) \quad \delta_d(X_t, Y_t; C_t) = \sup_{E \in C_t} |P^t(E) - P^t(E)| < 0, (t \gg t_0).
\]

(b) The two sequences are said to be asymptotically equivalent in the sense of type \((\text{C})_r\), and denoted as

\[
X_t \sim Y_t (\text{C})_r, (t \gg t_0),
\]

provided that

\[
(1.4) \quad \delta_r(X_t, Y_t; C_t) = \sup_{E \in C_t} |P^t(E) - P^t(E)| < 0, (t \gg t_0).
\]

In case where \((R_t, B_t)\) and \(C_t\) are fixed independently of \(t\), a weaker notion of asymptotic equivalence is defined corresponding to each types of the above definition. Let \((R_t, B_t) = (R, B)\) and \(C_t = C\).

Definition 1.2. (a) \( \{X_t; t \gg t_0\} \) and \( \{Y_t; t \gg t_0\} \) are said to be asymptotically equivalent in the sense of type \((\text{C})_d\), and denoted as

\[
X_t \sim Y_t ((\text{C})_d), (t \gg t_0),
\]

provided that

\[
(1.6) \quad |P^t(E) - P^t(E)| < 0, (t \gg t_0)
\]

for each \(E\) belonging to \(C\).
(b) The sequences are said to be asymptotically equivalent in the sense of type \(((c))_r\), and denoted as
\[
X_t \sim Y_t \ ((c)), \ (t \to t_0),
\]
provided that
\[
\left| \frac{P^X(E)}{P^Y(E)} - 1 \right| \to 0, \ (t \to t_0)
\]
for each \(E\) belonging to \(C\).

Each of the four types of asymptotic equivalence defines the corresponding type of notions of asymptotic independence and convergence. (see Ikeda 1963, 1968).

Among various types of asymptotic equivalence defined above, type \((\alpha)_d\) asymptotic equivalence enjoys very nice properties. By Def. 1.1, \(X_t \sim Y_t \ ((\alpha)_d), \ (t \to t_0)\), if and only if
\[
\delta_d(X_t; Y_t: B_t) = \sup_{E \in B_t} \left| P^X(E) - P^Y(E) \right| \to 0, \ (t \to t_0).
\]

First we state the following:

Proposition 1.1 Let \(\varphi_t(z)\) be any given measurable transformation from \((R^t, \Theta_t)\)
to \((\bar{R}^t, \bar{\Theta}_t)\) for each \(t\), and put \(\bar{x}_t = \varphi_t(x_t)\) and \(\bar{y}_t = \varphi_t(y_t)\). Then, \(X_t \sim Y_t \ ((\alpha)_d)\)
implies \(\bar{x}_t \sim \bar{y}_t \ ((\alpha)_d)\), as \(t \to t_0\). If, in particular, both the spaces are identical
and \(\varphi_t\) is non-singular, then \(X_t \sim Y_t \ ((\alpha)_d)\) is equivalent to \(\bar{x}_t \sim \bar{y}_t \ ((\alpha)_d)\), as \(t \to t_0\).

Proposition 1.2 Let \(h_t(z)\) be a real, single valued function defined over \(R_t\),
for each \(t\). If \(E|h_t(x_t)|^{1+\delta}\) and \(E|h_t(y_t)|^{1+\delta}\) are both uniformly bounded in \(t\)
for some \(\delta > 0\), then the condition \(X_t \sim Y_t \ ((\alpha)_d), \ (t \to t_0)\), implies that
\[
E| h_t(x_t) - h_t(y_t) | \to 0, \ (t \to t_0),
\]
where in general \(E h(X) = \int_{R} h \ d P^X\).

Some useful criteria for type \((\alpha)_d\) asymptotic equivalence have been given:

Proposition 1.3 (a) For \(X_t \sim Y_t \ ((\alpha)_d), \ (t \to t_0)\), it is sufficient that either one
of the following conditions holds:

\[(1.10) \quad \min\{I(X_t;Y_t;A_t), I(Y_t;X_t;A_t)\} \to 0, \quad (t\to t_0).\]

\[(1.11) \quad \min\{W(X_t;Y_t;A_t), W(Y_t;X_t;A_t)\} \to 0, \quad (t\to t_0), \quad \text{where \(\{A_t; t\to t_0\}\) is a sequence of subsets such that \(P^t(A_t) \to 1\) or \(P^t(A_t) \to 1\) as \(t \to t_0\), and \(I\) and \(W\) are defined by}

\[(1.12) \quad I(X_t;Y_t;A_t) = \int_{A_t} f_t \log \frac{f_t}{g_t} d\nu_t,
\]

\[(1.13) \quad W(X_t;Y_t;A_t) = \int_{A_t} g_t (-1)^2 d\nu_t.
\]

(b) For \(X_t \sim Y_t (\theta)_{d'} \quad (t\to t_0), \quad \text{it is necessary and sufficient that}

\[(1.14) \quad \rho(X_t,Y_t;\theta) \to 0, \quad (t\to t_0)
\]

where \(\{A_t, t\to t_0\}\) is the same as in (a) above.

Evaluation of approximation error, \(\delta_d(X_t,Y_t;\theta_t)\), can be done by using either one of the following inequalities:

**Proposition 1.4**

\[(1.15) \quad 0 \leq 1 - \rho(X_t,Y_t;R_t) \leq \delta_d(X_t,Y_t;\theta_t) \leq (1 - \rho(X_t,Y_t;R_t))^{1/2}
\]

\[\leq \left(\min\{I(X_t;Y_t;R_t), I(Y_t;X_t;R_t)\}\right)^{1/2}
\]

\[\leq \left(\min\{\log(1+W(X_t;Y_t;R_t)), \log(1+W(Y_t;X_t;R_t))\}\right)^{1/2}
\]

\[(1.16) \quad \delta_d(X_t,Y_t;\theta_t) \leq \left\{\frac{3}{4} \left(1 - \frac{4}{3} \min\{I(X_t;Y_t;R_t), I(Y_t;X_t;R_t)\}\right)^{1/2} - 1\right\}^{1/2}
\]

\[\leq \left(\frac{3}{8} \min\{I(X_t;Y_t;R_t), I(Y_t;X_t;R_t)\}\right)^{1/2}.
\]
Recently, Matsunawa and Ikeda (1981) have obtained a necessary and sufficient condition for $X_t \sim Y_t (A_d)_{t \to t_0}$ to hold in terms of the quantity $I(X_t; Y_t; A_t)$.

In the following section, we shall specialize the notions of asymptotic equivalence of type $(c)_d$ and of type $((c))_d$ to a special set of subclasses in real case, which are of common interest in statistics.

2. Asymptotic equivalence in real case.

Let $(X_s(n_s); s \to \infty)$ and $(Y_s(n_s); s \to \infty)$ be two sequences of real random variables, $X_s(n_s)$ and $Y_s(n_s)$ belonging to $\mathcal{P}(R_{n_s}, \mathcal{B}_{n_s})$, the class of all random variables distributed over the $n_s$-dimensional Euclidean space, $R_{n_s}$, with the usual Borel field $\mathcal{B}_{n_s}$.

We take the following sequences of subclasses:

$$
\mathcal{A}(n) = \{ Z(n) \mid -\infty < Z_i < b_i; i = 1,2,\ldots, n; b_i: \text{extended real} \},
$$

$$
\mathcal{S}(n) = \{ Z(n) \mid a_i \leq Z_i < b_i; i = 1,2,\ldots, n; a_i, b_i: \text{extended real} \},
$$

$$
\mathcal{N} = \{ \sum_{j=1}^{N} E_j (\text{disjointsum}) E_j \in \mathcal{S}(n); j = 1, \ldots, N; N: \text{any positive integer} \},
$$

$$
\mathcal{O} = \text{class of all open subsets of } R_{n}.
$$

Among the five types of asymptotic equivalence, $(\mathcal{M})_d$, $(\mathcal{S})_d$, $(\mathcal{A})_d$, $(\mathcal{N})_d$ and $(\mathcal{O})_d$, corresponding to the sequences in (2.1), the following implication relations hold:
In case of general basic spaces,

\[(2.3) \quad \{(\theta)_d, (A)_d, (\omega)_d\} \xrightarrow{\rightarrow} \{(\theta)_d\} \xrightarrow{\rightarrow} \{(\gamma)_d\},\]

and in case of identical (or equal) basic spaces where \((R(n_s), \theta(n_s)) = (R(n), \theta(n))\) for all \(s,\)

\[(2.4) \quad \{(\theta)_d, (A)_d, (\omega)_d\} \xrightarrow{\rightarrow} \{(\theta)_d, (\gamma)_d\}\]

where \{\} designates a group of equivalent notions, \(\rightarrow\) "imply", and \(\leftrightarrow\) "not necessarily imply, counter example shown".

Among the five weaker types of notions, the following implication diagram is obtained:

\[(2.5) \quad \{(\theta)_d\} \rightarrow \{(\omega)_d\} \xrightarrow{\rightarrow} \{(A)_d, (\gamma)_d\}\]

Some conditional implication relations have been also obtained (see Ikeda 1968). Note that type \((\gamma)_d\) convergence is equivalent to the usual in law convergence, provided the limiting distribution is of continuous type, in which case, it is also shown that \((\gamma)_d\) and \((\gamma)_d\) are mutually equivalent, as in the Central Limit Theorem. Also, in many cases of in law convergence, it turns out that the convergence is of type \((\theta)_d\).

In the following section, we shall apply the type \((\theta)_d\) asymptotic equivalence notion to the asymptotic normality of a set of increasing number of selected order statistics.

3. Asymptotic \((\theta)_d\) joint normality of sample quantiles.

Mosteller (1946) has shown, under mild conditions, that given a spacing \(0 < \lambda_1 < \ldots < \lambda_k < 1\), the corresponding set of order statistics,
$X_{n,n_1}, X_{n,n_2}, \ldots, X_{n,n_k}, n_i = [\lambda_i n] + 1$, are asymptotically jointly normally distributed, in a sense of type $((\eta_0))_d$ in our present terminology, with mean vector $(F^{-1}(\lambda_1), \ldots, F^{-1}(\lambda_k))$ and covariance matrix

$$
\frac{1}{n} \begin{bmatrix}
\frac{\lambda_1(1-\lambda_1)}{f_1^2} & \frac{\lambda_1(1-\lambda_2)}{f_1 f_2} & \cdots & \frac{\lambda_1(1-\lambda_k)}{f_1 f_k} \\
\frac{\lambda_2(1-\lambda_2)}{f_2^2} & \frac{\lambda_2(1-\lambda_3)}{f_2 f_3} & \cdots & \frac{\lambda_2(1-\lambda_k)}{f_2 f_k} \\
\frac{\lambda_3(1-\lambda_3)}{f_3^2} & \frac{\lambda_3(1-\lambda_4)}{f_3 f_4} & \cdots & \frac{\lambda_3(1-\lambda_k)}{f_3 f_k} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\lambda_k(1-\lambda_k)}{f_k^2} & & & \frac{\lambda_k(1-\lambda_1)}{f_k f_1}
\end{bmatrix}
$$

Later on, Weiss (1969) has tried to derive an asymptotic distribution, in a strong sense, of a set of increasing number of sample quantiles, and has proved a result, which is a special case of Ikeda and Matsunawa (1972).

Now, we begin with the case of uniform distribution over the unit interval, $(0,1)$. Let $U_{n,1}, U_{n,2}, \ldots, U_{n,n}$ be order statistics based on a random sample of size $n$ drawn from a uniform distribution over $(0,1)$. Select $k$ order statistics, $U_{n,n_1}, U_{n,n_2}, \ldots, U_{n,n_k}$, and put

$$
U_n(k) = (U_{n,n_1}, U_{n,n_2}, \ldots, U_{n,n_k})';
$$

where $k$ and $(n_1, n_2, \ldots, n_k)$ may depend on $n$ as $n \to \infty$.

By applying the criterion (1.10), Ikeda and Matsunawa (1972) have shown that the following theorem holds:
Theorem 3.1  If the condition
\[ w_n = \frac{k}{\min(n_i - n_{i-1})} \to 0, (n \to \infty), \]
holds, then \( U_n(k) \) and \( Z_n(k) \) are asymptotically equivalent in the sense of type \((\ast)_d\) as \( n \to \infty \), where \( Z_n(k) \) stands for a normal random variable with mean vector
\[ \varepsilon_n(k) = (\varepsilon_{n1}, \varepsilon_{n2}, \ldots, \varepsilon_{nk})', \]
with \( \varepsilon_{ni} = \frac{n_i}{n+1} \), \( i=1,2,\ldots,k \),
and covariance matrix
\[ L_n(k) = \frac{1}{n+2} \begin{bmatrix} \varepsilon_{n1}(1-\varepsilon_{n1}) & \varepsilon_{n1}(1-\varepsilon_{n2}) & \cdots & \varepsilon_{n1}(1-\varepsilon_{nk}) \\ \varepsilon_{n2}(1-\varepsilon_{n1}) & \varepsilon_{n2}(1-\varepsilon_{n2}) & \cdots & \varepsilon_{n2}(1-\varepsilon_{nk}) \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{nk}(1-\varepsilon_{n1}) & \varepsilon_{nk}(1-\varepsilon_{n2}) & \cdots & \varepsilon_{nk}(1-\varepsilon_{nk}) \end{bmatrix}. \]
Here we have taken a convention \( n_0 = 0, n_{k+1} = n+1 \).

In this case an upper bound for the quantity \( \varepsilon_d(U_n(k);Z_n(k);R(k)) \) is given by (1.16):
\[ \varepsilon_d(U_n(k);Z_n(k);R(k)) \leq \left( \frac{3}{8} I(U_n(k);Z_n(k);R(k)) \right)^{1/2}, \]
with
\[ I(U_n(k);Z_n(k);R(k)) = w_n + o(w_n). \]

The result obtained in Theorem 3.1 is, of course, extended to a uniform distribution over any given finite interval, by virtue of Prop.1.1.
Ikeda and Matsunawa (1972) have also obtained a result in case of general basic distribution. However, the conditions there are somewhat gloomy and not convenient to practical use. Moreover, the conditions include a stronger condition for the spacing of \((n_1, n_2, \ldots, n_k)\) than (3.3):

\[
\frac{k^2}{\min(n_i - n_{i-1})} \to 0, \quad (n \to \infty).
\]

Recently, Ikeda and Nonaka (1981) have obtained a refined result, which improves the earlier one.

Let \(X_{n,1} < X_{n,2} < \ldots < X_{n,n}\) be order statistics from a continuous distribution over the real line, whose pdf. and cdf. being given by \(f(x)\) and \(F(x)\), respectively. Choose \(k\) out of the order statistics, and put

\[
X_{n}(k) = (X_{n,n_1}, X_{n,n_2}, \ldots, X_{n,n_k}).
\]

We shall first consider the case where the support of \(f(x)\) is identical with the entire real line: \(D_f = (-\infty, \infty)\).

Then, the transformed variable

\[
F(X_{n}(k)) = \left(F(X_{n,n_1}), F(X_{n,n_2}), \ldots, F(X_{n,n_k})\right)
\]

is identically distributed with \(U_{n}(k)\) given in (3.2), or,

\[
F^{-1}(U_{n}(k)) = \left(F^{-1}(U_{n,n_1}), F^{-1}(U_{n,n_2}), \ldots, F^{-1}(U_{n,n_k})\right)
\]

is identically distributed with \(X_{n}(k)\).
Let us consider a truncation of $Z_{n(k)}$ over the domain $A_{(k)} = \{z(k)\}$, $0 \leq z_i \leq 1; i = 1,2,\ldots,k$, and denote it by $Z^*_n(k)$. It is then evident that, under the condition (3.3), $Z_{n(k)} \sim Z^*_n(k) (\mu)_d, (n \to \infty)$. Further, let us put

$$Y^*_n(k) = F^{-1}(Z^*_n(k)) = \left(F^{-1}(Z^*_n,1), F^{-1}(Z^*_n,2), \ldots, F^{-1}(Z^*_n,k)\right),$$

and finally let $Y_n(k)$ be a normal random variable with mean vector

$$s_n(k) = (s_{n1}, s_{n2}, \ldots, s_{nk}),$$

with $s_{ni} = F^{-1}(\xi_{ni}), i = 1,2,\ldots,k,$

and covariance matrix

$$S_n(k) = \frac{1}{n+2} \begin{bmatrix}
\kappa_{n1}(1-\kappa_{n1}) & \kappa_{n1}(1-\kappa_{n2}) & \cdots & \kappa_{n1}(1-\kappa_{nk}) \\
\frac{\kappa_{n1}^2}{n1} & \frac{\kappa_{n1}\kappa_{n2}}{n1n2} & \cdots & \frac{\kappa_{n1}\kappa_{nk}}{n1nk} \\
\kappa_{n2}(1-\kappa_{n2}) & \frac{\kappa_{n2}^2}{n2} & \cdots & \frac{\kappa_{n2}\kappa_{nk}}{n2nk} \\
\kappa_{nk}(1-\kappa_{nk}) & \frac{\kappa_{nk}^2}{nk} & \cdots & \frac{\kappa_{nk}\kappa_{nk}}{nk}\end{bmatrix},$$

where we have put $f_{ni} = f(s_{ni}) = f(F^{-1}(\xi_{ni})), i = 1,2,\ldots,k.$

The following diagram indicate the relations among the variables thus defined:

$$U_n(k) \sim d Z_n(k) \sim d Z^*_n(k)$$

$$F^{-1} \vee Y^*_n(k) \sim F^{-1} \sqrt{Y^*_n(k)}$$

$$X_n(k) \sim d Y_n(k) \sim d Y^*_n(k)$$
Under the condition (3.3), it holds that $U_n(k) \sim Z^*_n(k) (\beta)_d$ and therefore, by Prop.1.1, $X_n(k) \sim Y^*_n(k) (\beta)_d$, as $n \to \infty$. Hence, if one can show that $Y_n(k) \sim Y^*_n(k) (\beta)_d$, (n \to \infty), with possibly some additional conditions, then it holds that $X_n(k) \sim Y_n(k) (\beta)_d$, i.e., $X_n(k)$ would be distributed as normal with mean vector $S_n(k)$ and covariance matrix $S_n(k)$.

Ikeda and Nonaka (1981) investigated the quantity $I(Y_n(k); Y^*_n(k); R(k))$ for criticizing the type $(\beta)_d$ asymptotic equivalence between $Y_n(k)$ and $Y^*_n(k)$, to get the following theorem.

**Theorem 3.2** Suppose that the following assumptions are fulfilled:

(i) The support of $f(x)$ is identical to the entire real line: $D_f = (-\infty, \infty)$.

(ii) $f(x)$ is twice differentiable and $f''(x)$ is bounded and continuous over the entire real line.

(iii) The function, $\varphi(x) = \frac{(f(x)f''(x)-f'(x)^2)}{f(x)^2}$, is bounded uniformly for all $x$ in $(-\infty, \infty)$.

Then, in order that $X_n(k) \sim Y_n(k) (\beta)_d$, (n \to \infty), it is sufficient that the following conditions are satisfied simultaneously:

\begin{align*}
(3.16) \quad w_n \cdot \min_{1 \leq i \leq k+1} \frac{k}{\min(n_i-n_{i-1})} & \to 0, (n \to \infty), \\
(3.17) \quad w_n \cdot \max(\frac{4}{\sigma_n^2}, \frac{6}{n+2}) & \to 0, (n \to \infty),
\end{align*}

where we have put

\begin{align*}
(3.18) \quad \sigma_n^2 & = \max \sigma_{n1}^2, \quad \sigma_{ni}^2 = \frac{\xi_i n_i (1-k_{ni})}{f_{ni}^2}, \quad i=1,2,\ldots,k.
\end{align*}
In case where \( k \) and \( \varepsilon_{n_1} \)'s are fixed, the conditions (3.16) and (3.17) are automatically fulfilled, in which case, however, the assumptions (i)-(iii) happen to be slightly stronger than those by Mosteller (1946). In this case, more direct calculation would be possible, which will be left open.

Ikeda and Nonaka (1981) have given a more general result than the above theorem, which states the following result:

**Theorem 3.3** Suppose that the following assumptions are fulfilled:

(i) The support of \( f(x) \) is identical to an open interval: \( D_f = (a,b) \), where \( a \) and \( b \) are extended real.

(ii) \( f(x) \) is twice differentiable and \( f''(x) \) is bounded and continuous over \((a,b)\).

(iii) The function, \( \varphi(x) = \left\{f(x)f''(x) - f'(x)^2\right\}/f(x)^2 \), is bounded uniformly for all \( x \) in \((a,b)\).

Then, in order that \( X_n(k) \sim Y_n(k) \) \( (a) \), \( (n \to \infty) \), it is sufficient that the following conditions are satisfied simultaneously:

\[
(3.19) \quad w_n = \frac{k}{\min(n_1, n_1, \ldots, n_1)} \to 0, \quad (n \to \infty).
\]

\[
(3.20) \quad w_n \left\{\sigma_n, \frac{\sigma_n^4}{\sqrt{n+2}}, \frac{\sigma_n^6}{n+2}\right\} \to 0, \quad (n \to \infty).
\]

and

\[
(3.21) \quad w_n \max_{1 \leq i \leq k} \left[\frac{\sigma_{ni}^4}{\min(a-s_{ni}^2, b-s_{ni}^2)}\right] \to 0, \quad (n \to \infty).
\]

It is evident that this theorem implies the result in the preceding theorem. Also, it should be noted that, in order to obtain an error estimation
for $\delta_d(X_n(k), Y_n(k); \beta(k))$, an evaluation of the K-L information $I(X_n(k); Y_n(k); R(k))$ should be done directly.
References


Notions of asymptotic equivalence of probability distributions and some of their properties are briefly presented. By applying the results on type $(\alpha)_d$ asymptotic equivalence, asymptotic $(\beta)_d$ joint normality of a set of increasing number of sample quantiles are discussed, which improves and refines the previous work by Ikeda and Matsunawa (1972).