This is the published version of a paper published in *Journal of Applied Probability*.

Weakly Approaching Sequences of Random Distribution

Laws with Applications to Resampling

Yuri K. Belyaev

Weakly approaching sequences of random distribution laws with applications to resampling

Yuri K. Belyaev

Department of Mathematical Statistics
Umeå University
S-901 87 Umeå, Sweden

Abstract

The standard definition of weak convergence of distribution laws (d.l.'s) is generalized to a definition of weak approach of d.l.'s to each other. For the case of random d.l.'s a definition is given for a sequence of random d.l.'s to weakly approach a sequence of non-random d.l.'s. The main result (Theorem 2) concerns this case providing necessary and sufficient conditions for weak approach in terms of characteristic functions of random and non-random d.l.'s. Applications of the theory are given to scheme of series of independent random variables and to linear regression.[1]

Key words: random distribution laws, weak convergence, characteristic functions, resampling, Central Limit Resampling Theorem, linear regression

1991 AMS subject classification: Primary 62F12; secondary 62G09

---

1 Introduction

Asymptotic methods are often used in the analysis of collected statistical data. A typical problem is to obtain an asymptotically exact approximation of an "unobserved" sequence of distribution laws (d.l.'s) by the use of some other "observed" sequence of d.l.'s when the volume of collecting data is increasing. For example one may wish to determine the d.l.'s of deviations of point estimators from some estimated parameters. Intensive computer methods, using bootstrap and resampling for example, can help to solve such problems. Usually it is assumed or implied under some assumptions, that the "unobserved" sequence of d.l.'s weakly converges to some limit d.l. which we call the *attractor d.l.* The existence of attractor d.l.'s can be a too strong restriction on statistical models and can limit the use of known statistical methods in the analysis of real statistical data. Moreover, the reduction of a problem to one with an attractor d.l. inevitably involves the use of stabilizing transformations, in the form of normalizing constants or matrices, for example, and then their unknown parameters need to be estimated. This can be difficult if the collected data have different structures and hence different probability distributions, Belyaev, and Gadjiev [3], White [11]. In this paper we suggest a way to avoid the use of attractor d.l.'s by defining what it means for sequences of d.l.'s to weakly approach each other. Sequences of random d.l.'s can also be considered, providing a justification of the use of resampling methods. The new type of convergence of sequences of d.l.'s is a natural generalization of the usual weak convergence of d.l.'s to some limit d.l., i.e. to some attractor d.l. It is shown that this generalized weak convergence has similar properties to standard weak convergence. The well-known continuity theorem (see e.g. Shiryaev [10], page 318) states relations between the weak convergence of d.l.'s and the pointwise convergence of characteristic functions (c.f.'s). One variant of the continuity theorem within the framework of ordinary weak convergence was given in Knight [6]. Theorems 1 and 2 in Section 2 are generalisations of the continuity theorem in the case of sequences of d.l.'s that weakly approach each other. These theorems can be generalised to the case of sequences of d.l.'s of vector - valued random variables (r.v.'s) that weakly approach each other, see forthcoming report Belyaev, and Sjöstedt [4]. Theorem 2 can be used to justify the resampling methods in many typical problems involving the point estimation of unknown parameters, and in some other problems. This paper is an extended version of Section 2 in the research report Belyaev [2]. The basic definitions and results in Belyaev [2] are restated. Here, the proofs are given only for new results.

Capital letters are used for r.v.'s, small letters for values of r.v.'s and for constants, Greek letters for parameters of d.l.'s, bold letters for vectors, contour fonts (such as $\mathbb{B}$) for collections of sets, and matrices, $\mathbb{X}^T$ for the transposed matrix $\mathbb{X}$, calligraphic (bold) letters for sets and d.l.'s, $\mathcal{L}(X)$ denotes the d.l. (probability measure) of a r.v. X, $N_1(\mu, \sigma^2)$ denotes the one dimensional normal d.l. with

2

mean value μ and variance σ^2, $\mathcal{C}(\mu, \theta)$ denotes the Cauchy d.l. centered at μ and with the scale parameter θ, := means definition by expression, $f(\cdot)$ denotes the "whole" function f and $f(x)$ denotes its value at x, $I(A)$ is an indicator function of an event A, $\mathbb{Z}^+ = \{1, 2, ...\}$, A^c is the complement of a set A. We use $\xrightarrow{P}$ to denote the convergence in probability, and $\xrightarrow{d}$ for convergence in distribution.

We now outline a class of problems to which Theorem 2 is applicable. This is a scheme of series of observations of independent r.v.'s, for which we are interested in approximating the d.l.'s of sums of the r.v.'s.

Let $\mathbf{U}_n := (U_{1n}, ..., U_{nn})$ be an array of collections of real valued r.v.'s U_{in}, $EU_{in} = 0$, $i = 1, ..., n$, $n = 1, 2, ...$. For each r.v. U_{in} we have only one observation u_{in}. We will consider $\mathbf{u}_n = (u_{1n}, ..., u_{nn})$ as the observed statistical data, and denote by $\mathcal{L}(U_{\cdot n})$ the d.l. of $U_{\cdot n} := U_{1n} + ... + U_{nn}$. We seek to understand how, and under what assumptions, it is possible to approximate $\mathcal{L}(U_{\cdot n})$, asymptotically accurately as $n \to \infty$, given only $\mathbf{u}_n$. In Section 2 we cite the so-called Central Limit Resampling Theorem (CLRT) obtained in Belyaev [1]. The case with non i.i.d. sequence of r.v.'s has been considered in Liu [7], where the uniform convergence of d.l.'s was used.

2 Weakly approaching sequences of distribution laws

Distribution laws (probability measures) are points in some functional space (the space of all probability measures or all distribution functions on $\mathbf{R}^1$). Estimators of d.l.'s should be considered as random d.l.'s (random probability measures). It is necessary to define what is to be meant by the statement that two sequences of d.l.'s approach each other. Our definition of convergence is a modification of the standard notion of weak convergence. We restate the definitions and Lemmas 1 - 3, and 5 from Belayev [2]. Let $\mathcal{C}_b(\mathbf{R}^1)$ be the set of all bounded, continuous, real valued functions on $\mathbf{R}^1$.

Definition 1 *Two sequences of d.l.'s $(\mathcal{L}(X_n))_{n \geq 1}$, $(\mathcal{L}(Y_n))_{n \geq 1}$ of real valued random variables $(X_n)_{n \geq 1}$, and $(Y_n)_{n \geq 1}$ are said to weakly approach each other if for any function $c(\cdot) \in \mathcal{C}_b(\mathbf{R}^1)$*

$$Ec(X_n) - Ec(Y_n) \to 0 \quad as \ n \to \infty. \tag{1}$$

Correspondingly $(X_n)_{n \geq 1}$, and $(Y_n)_{n \geq 1}$ are said to approach each other in distribution.

We denote this convergence by $\mathcal{L}(X_n) \overset{wa}{\leftrightarrow} \mathcal{L}(Y_n)$ or $X_n \overset{ad}{\leftrightarrow} Y_n$, $n \to \infty$.

Remark. If $\mathcal{L}(Y_n) = \mathcal{L}(Y_0)$ for some r.v. Y_0, and $n = 1, 2, ...$, and (1) holds, then the sequence $(\mathcal{L}(X_n))_{n \geq 1}$ is said to *weakly converge* to $\mathcal{L}(Y_0)$, $\mathcal{L}(X_n) \overset{w}{\to} \mathcal{L}(Y_0)$

3

or $X_n \xrightarrow{d} Y_0$, as $n \to \infty$. This is the case usually considered in the literature (see e.g. Dudley [5], and Shiryaev [10]). □

There is also a need to consider random d.l.'s. We restrict ourselves to the following situation. We assume that the d.l.'s of $(X_n)_{n\geq 1}$ are random d.l.'s in the sense that there are r.v.'s Z_n, taking values in some space $\mathcal{Z}_n$, such that the d.l. of X_n is the ordinary regular conditional d.l. given Z_n. We denote it by $\mathcal{L}(X_n \mid Z_n)$. In this case, we can consider conditional expectations $\mathrm{E}(c(X_n) \mid Z_n)$, $c(\cdot) \in \mathcal{C}_b(\mathbf{R}^1)$, $n = 1, 2, ...$, which are real valued r.v.'s.

Definition 2 *Let $(X_n, Z_n)_{n\geq 1}$, and $(Y_n)_{n\geq 1}$ be two sequences of r.v.'s, where $X_n, Y_n \in \mathbf{R}^1$, $Z_n \in \mathcal{Z}_n$, the r.v.'s X_n, Z_n being defined on the same (basic) probability space $(\Omega, \mathfrak{B}(\Omega), \mathrm{P})$ for each n. The sequence of conditional d.l.'s $(\mathcal{L}(X_n \mid Z_n))_{n\geq 1}$ given Z_n is said to weakly approach $(\mathcal{L}(Y_n))_{n\geq 1}$ in probability (almost surely) along $(Z_n)_{n\geq 1}$, if for any $c(\cdot) \in \mathcal{C}_b(\mathbf{R}^1)$*

$$\mathrm{E}(c(X_n) \mid Z_n) - \mathrm{E}c(Y_n) \to 0 \quad \text{as } n \to \infty, \tag{2}$$

in probability (almost surely). We denote this convergence by

$$\mathcal{L}(X_n \mid Z_n) \xleftrightarrow{wa(\mathrm{P})} \mathcal{L}(Y_n) \quad (\mathcal{L}(X_n \mid Z_n) \xleftrightarrow{wa(as)} \mathcal{L}(Y_n)) \quad \text{as } n \to \infty.$$

Remark. Here we combine weak convergence and convergence in probability. If $\mathcal{L}(Y_n) \equiv \mathcal{L}(Y_0)$, the sequence $(\mathcal{L}(X_n \mid Z_n))_{n\geq 1}$ is said to *weakly converge to* $\mathcal{L}(Y_0)$ *in probability (almost surely) along* $(Z_n)_{n\geq 1}$ as $n \to \infty$, if (2) holds almost surely. Note that P can also depend on n. □

We recall that a sequence of real valued r.v.'s $(X_n)_{n\geq 1}$ is said to be *uniformly tight* if for any small $\varepsilon > 0$ there is a $k_\varepsilon < \infty$ such that

$$\mathrm{P}(\mid X_n \mid > k_\varepsilon) < \varepsilon, \quad n = 1, 2, ... $$

Definition 3 *Let $(X_n \mid Z_n)_{n\geq 1}$ be a sequence of r.v.'s X_n given Z_n, $n = 1, 2, ...$ The sequence $(X_n \mid Z_n)_{n\geq 1}$ is called uniformly tight along $(Z_n)_{n\geq 1}$ in probability if for any small $\varepsilon > 0$, and $\delta > 0$ there is a constant $k_{\varepsilon\delta} < \infty$, such that*

$$\mathrm{P}(\mathrm{P}(\mid X_n \mid > k_{\varepsilon\delta} \mid Z_n) > \varepsilon) < \delta, \quad n = 1, 2, ... \tag{3}$$

Lemma 1 *If $(Y_n)_{n\geq 1}$ is uniformly tight and $\mathcal{L}(X_n) \xleftrightarrow{wa} \mathcal{L}(Y_n)$ as $n \to \infty$, then $(X_n)_{n\geq 1}$ is also a uniformly tight sequence of r.v.'s.*

Proof. See Belyaev [2].

Lemma 2 *Let $(X_n, V_n)_{n\geq 1}$ be a sequence of pairs of real valued r.v.'s defined on the same probability space, and $V_n \xrightarrow{\mathrm{P}} 0$, $n \to \infty$. Assume also that $(Y_n)_{n\geq 1}$ is uniformly tight real valued r.v.'s and that $\mathcal{L}(X_n) \xleftrightarrow{wa} \mathcal{L}(Y_n)$ as $n \to \infty$. Then $\mathcal{L}(X_n + V_n) \xleftrightarrow{wa} \mathcal{L}(Y_n)$ as $n \to \infty$.*

Proof. See Belyaev [2].

Definition 4 *A sequence of real valued random variables $(X_n)_{n \geq 1}$ is said to be properly random if there is no sequence of real numbers $(a_n)_{n \geq 1}$ such that $X_n - a_n \xrightarrow{\text{P}} 0$ as $n \to \infty$.*

If $\mathcal{L}(X_n) = N_1(0, \sigma_n^2)$, $0 < \sigma_-^2 \leq \sigma_n^2 \leq \sigma_+^2 < \infty$, then $(X_n)_{n \geq 1}$ is properly random.

Lemma 3 *Let $(X_n)_{n \geq 1}$ be uniformly tight. Then $(X_n)_{n \geq 1}$ is properly random iff for any sequence of real numbers $(a_n)_{n \geq 1}$, there exists $c(\cdot) \in \mathcal{C}_b(\mathbf{R}^1)$ such that $\mathrm{E}c(X_n) - c(a_n) \not\to 0$ as $n \to \infty$.*

Proof. See Belyaev [2].

   Note. If $(Y_n)_{n \geq 1}$ is uniformly tight, properly random and $\mathcal{L}(X_n) \overset{wa}{\leftrightarrow} \mathcal{L}(Y_n)$, as $n \to \infty$, then $(X_n)_{n \geq 1}$ is also uniformly tight and properly random. $\quad\square$

   Let $G_n(s) = \mathrm{E}I(Y_n \leq s)$, $F_n(s) = \mathrm{E}I(X_n \leq s)$, and $F_n(s \mid Z_n) = \mathrm{E}(I(X_n \leq s) \mid Z_n)$ be distribution functions (d.f.'s) of the r.v.'s Y_n, X_n, and X_n given Z_n. The sequence of d.f.'s $(G_n(\cdot))_{n \geq 1}$ is uniformly continuous iff

$$\sup_{n \geq 1} \sup_s \mid G_n(s+h) - G_n(s) \mid \to 0 \quad \text{as} \quad h \to 0 \, .$$

The following lemma and its corollary show that under natural assumptions, weakly approaching sequences of r.v.'s have uniformly approaching d.f.'s.

Lemma 4 *Let $(X_n)_{n \geq 1}$, $(Y_n)_{n \geq 1}$ be two sequences of r.v.'s with d.f.'s $(F_n(\cdot))_{n \geq 1}$, $(G_n(\cdot))_{n \geq 1}$. If the sequence (Y_n) is uniformly tight, $(G_n(\cdot))_{n \geq 1}$ is uniformly continuous, and $\mathcal{L}(X_n) \overset{wa}{\leftrightarrow} \mathcal{L}(Y_n)$, $n \to \infty$, then*

$$\sup_s \mid F_n(s) - G_n(s) \mid \to 0 \quad \text{as} \quad n \to \infty \, . \tag{4}$$

Proof. The assumption of uniform continuity $(G_n(\cdot))_{n \geq 1}$ implies that for any s, and for any arbitrarily small $\varepsilon > 0$ we can find a $h_\varepsilon = h_\varepsilon(s) > 0$ such that

$$\sup_s (G_n(s + h_\varepsilon) - G_n(s - h_\varepsilon)) < \frac{\varepsilon}{10}, \quad n = 1, 2, \dots . \tag{5}$$

For any s we define two bounded continuous functions of u

$$c_{0,\varepsilon}(u, s) := I(u \leq s + h_\varepsilon) - ((u-s)/h_\varepsilon)I(s < u \leq s + h_\varepsilon) \, ,$$
$$c_{-,\varepsilon}(u, s) := I(u \leq s) - ((u - (s - h_\varepsilon))/h_\varepsilon)I(s - h_\varepsilon < u \leq s) \, .$$

The assumption $\mathcal{L}(X_n) \overset{wa}{\leftrightarrow} \mathcal{L}(Y_n)$, $n \to \infty$, and (5) imply that for all sufficiently large $n \geq n_\varepsilon(s)$ $\mid \mathrm{E}c_{i,\varepsilon}(X_n, s) - \mathrm{E}c_{i,\varepsilon}(Y_n, s) \mid \leq \varepsilon/10$, $i = 0, -$. Hence,

$$
\begin{aligned}
F_n(s) - G_n(s) &\leq \mathrm{E}c_{0,\varepsilon}(X_n, s) - \mathrm{E}c_{-,\varepsilon}(Y_n, s) \\
&= \mathrm{E}c_{0,\varepsilon}(X_n, s) - \mathrm{E}c_{0,\varepsilon}(Y_n, s) + \mathrm{E}c_{0,\varepsilon}(Y_n, s) - \mathrm{E}c_{-,\varepsilon}(Y_n, s) \\
&\leq \frac{\varepsilon}{10} + G_n(s + h_\varepsilon) - G_n(s - h_\varepsilon) \leq \frac{\varepsilon}{5},
\end{aligned}
\tag{6}
$$

and

$$
\begin{aligned}
F_n(s) - G_n(s) &\geq \mathrm{E}c_{-,\varepsilon}(X_n, s) - \mathrm{E}c_{0,\varepsilon}(Y_n, s) \\
&= \mathrm{E}c_{-,\varepsilon}(X_n, s) - \mathrm{E}c_{-,\varepsilon}(Y_n, s) + \mathrm{E}c_{-,\varepsilon}(Y_n, s) - \mathrm{E}c_{0,\varepsilon}(Y_n, s) \\
&\geq -\frac{\varepsilon}{10} + G_n(s - h_\varepsilon) - G_n(s + h_\varepsilon) \geq -\frac{\varepsilon}{5}, \tag{7}
\end{aligned}
$$

It follows, that

$$
\mid F_n(s) - G_n(s) \mid \leq \frac{\varepsilon}{5}, \quad \text{if} \quad n \geq n_\varepsilon(s). \tag{8}
$$

By Lemma 1 both sequences $(X_n)_{n \geq 1}$, and $(Y_n)_{n \geq 1}$ are uniformly tight. Therefore we can find sufficiently large $k_\varepsilon > 0$ such that

$$
\mathrm{P}(\mid X_n \mid > k_\varepsilon) < \varepsilon, \quad \mathrm{P}(\mid Y_n \mid > k_\varepsilon) < \varepsilon, \quad n = 1, 2, \dots. \tag{9}
$$

We consider a finite set

$$
S_{m,\varepsilon} = \left(s_{lm} : \ s_{lm} = \frac{l k_\varepsilon}{2^m}, \quad l = -2^m, -2^m + 1, \dots, 2^m - 1, 2^m \right),
$$

with points s_{lm}, $\mid s_{l+1,m} - s_{lm} \mid \leq h_\varepsilon$. From (8) if follows that

$$
\max_{s_{lm} \in S_{m,\varepsilon}} \mid F_n(s_{lm}) - G_n(s_{lm}) \mid \leq \frac{\varepsilon}{5}, \quad \text{if} \quad n \geq n_\varepsilon(m, k_\varepsilon), \tag{10}
$$

$n_\varepsilon(m, k_\varepsilon) := \max(n_\varepsilon(s) : \ s \in S_{m,\varepsilon})$. We also have for s, $s_{lm} \leq s \leq s_{l+1,m}$, s_{lm}, and $s_{l+1,m} \in S_{m,\varepsilon}$, by using (5) and (10), that

$$
\begin{aligned}
\mid F_n(s) - G_n(s) \mid \ &\leq \ F_n(s) - F_n(s_{lm}) + G_n(s) - G_n(s_{lm}) \\
&+ \ \mid F_n(s_{lm}) - G_n(s_{lm}) \mid \\
&\leq \ F_n(s_{l+1,m}) - F_n(s_{lm}) + G_n(s_{l+1,m}) - G_n(s_{lm}) \\
&+ \ \mid F_n(s_{lm}) - G_n(s_{lm}) \mid \\
&\leq \ \mid F_n(s_{l+1,m}) - G_n(s_{l+1,m}) \mid + 2(G_n(s_{l+1,m}) - G_n(s_{lm})) \\
&+ \ 2 \mid F_n(s_{lm}) - G_n(s_{lm}) \mid < \varepsilon.
\end{aligned}
$$

It follows that

$$
\max_{s_{lm}, s_{l+1,m} \in S_{m\varepsilon}} \sup_{s_{lm} \leq s \leq s_{l+1,m}} \mid F_n(s) - G_n(s) \mid < \varepsilon, \quad \text{if} \quad n \geq n_\varepsilon(m, k_\varepsilon).
$$

This inequality together with (9) imply

$$
\sup_s \mid F_n(s) - G_n(s) \mid < \varepsilon, \quad \text{if} \quad n \geq n_\varepsilon(m, k_\varepsilon),
$$

and the desired result follows. $\qquad\square$

Corollary 1 If $(Y_n)_{n \geq 1}$ is the same as in Lemma 4, and $\mathcal{L}(X_n \mid Z_n) \overset{wa(\mathrm{P})}{\longleftrightarrow} \mathcal{L}(Y_n)$, $n \to \infty$, then for any small $\varepsilon > 0$

$$\mathrm{P}\left(\sup_s \mid F_n(s \mid Z_n) - G_n(s) \mid > \varepsilon\right) \to 0, \quad n \to \infty. \tag{11}$$

Proof. Let $\varepsilon > 0$ and $\delta > 0$ be any small numbers. We can find a sufficiently large $k_{\varepsilon\delta} > 0$ such that (3) holds with $\delta/2$ instead of δ, and such that

$$G_n(-k_{\varepsilon\delta}) + 1 - G_n(k_{\varepsilon\delta}) < \varepsilon/2. \tag{12}$$

We can write (3) in the equivalent form

$$\mathrm{P}(F_n(-k_{\varepsilon\delta}- \mid Z_n) + 1 - F_n(k_{\varepsilon\delta} \mid Z_n) < \varepsilon) > 1 - \delta/2. \tag{13}$$

Let $k_\varepsilon = k_{\varepsilon\delta}$, m, and $S_{m\varepsilon}$ be as in Lemma 4. From the assumption $\mathcal{L}(X_n \mid Z_n) \overset{wa(\mathrm{P})}{\longleftrightarrow} \mathcal{L}(Y_n)$, $n \to \infty$, and inequalities similar to (6) and (7) with $s \in S_{m\varepsilon}$, and with $F_n(s \mid Z_n)$ and $\mathrm{E}(c_{i\varepsilon}(X_n, s) \mid Z_n)$ instead of $F_n(s)$ and $\mathrm{E}c_{i\varepsilon}(X_n, s)$, we have

$$\mathrm{P}\left(\mid F_n(s_{lm} \mid Z_n) - G_n(s_{lm}) \mid \leq \frac{\varepsilon}{5}\right) \geq 1 - \frac{\delta}{2(2^{m+1}+1)},$$

if n is sufficiently large, say $n \geq n'_\varepsilon(m, k_{\varepsilon\delta})$.

Further, by similar arguments to those in the proof of Lemma 4, we obtain

$$\mathrm{P}\left(\max_{s_{lm}, s_{l+1,m} \in S_{m\varepsilon}} \sup_{s_{l+1,m} \leq s \leq s_{lm}} \mid F_n(s \mid Z_n) - G_n(s) \mid < \varepsilon\right) > 1 - \delta/2, \tag{14}$$

if $n \geq n'_\varepsilon(m, k_\varepsilon)$. Inequalities (12), (13) and (14) imply

$$\mathrm{P}\left(\sup_s \mid F_n(s \mid Z_n) - G_n(s) \mid < \varepsilon\right) > 1 - \delta, \tag{15}$$

for all sufficiently large n. If $\varepsilon > 0$ is given we can choose $\delta > 0$ arbitrarily small. Hence, inequality (15) implies the assertion of Corollary 1. $\qquad\square$

Remark. Under natural assumptions, weakly approaching sequences of r.v.'s approach each other in the Levy metric. Let $\mathcal{C}_n = (s: G_n(s-) = G_n(s+))$, i.e. $\mathcal{C}_n$ contains all points of continuity of $G_n(\cdot)$. If $\mathcal{C}_n = \mathcal{C}_0$, i.e. the sets $\mathcal{C}_n$ are the same, and at any $s \in \mathcal{C}_0$ for any $\varepsilon > 0$ exists $h_\varepsilon > 0$ such that

$$\sup_n \mid G_n(s + h_\varepsilon) - G_n(s - h_\varepsilon) \mid < \varepsilon,$$

then the Levy distance

$$d_L(\mathcal{L}(X_n \mid Z_n), \mathcal{L}(Y_n)) \overset{\mathrm{P}}{\to} 0, \quad n \to \infty. \tag{16}$$

If $G_n(\cdot) = G_0(\cdot)$, $n = 1, 2, \ldots$, then (16) is valid. $\qquad\square$

Together with the sequences of r.v.'s $(X_n)_{n \geq 1}$ and $(Y_n)_{n \geq 1}$, we consider their characteristic functions (ch.f.) $f_n(t) := \mathrm{E}e^{itX_n}$, and $g_n(t) := \mathrm{E}e^{itY_n}$, $n \in \mathbb{Z}^+$. Here $i = \sqrt{-1}$. We restate the well-known continuity theorem (see Shiryaev [10], page 320).

Proposition 1 (Continuity theorem)

(i) *If $\mathcal{L}(Y_n) \overset{w}{\to} \mathcal{L}(Y_0)$ as $n \to \infty$, then $g_n(t) \to g_0(t)$ as $n \to \infty$, for each $t \in \mathbf{R}^1$ and $g_0(t) = \mathrm{E}e^{itY_0}$.*

(ii) *If $g(t) = \lim_{n \to \infty} g_n(t)$ for each t, and $g(\cdot)$ is continuous at $t = 0$, then $g(\cdot)$ is the ch.f. of some r.v. Y_0, i.e. $g(\cdot) = g_0(\cdot)$, and $\mathcal{L}_n(Y_n) \overset{w}{\to} \mathcal{L}(Y_0)$ as $n \to \infty$.*

The following theorem is an extension of the above cited continuity theorem.

Theorem 1 (Continuity theorem for sequences of d.l.'s that weakly approach each other) *Let $(X_n)_{n \geq 1}$, and $(Y_n)_{n \geq 1}$ be two sequences of r.v.'s, and suppose $(Y_n)_{n \geq 1}$ is uniformly tight. Then*

$$\mathcal{L}(X_n) \overset{wa}{\leftrightarrow} \mathcal{L}(Y_n) \quad as \ n \to \infty, \tag{17}$$

iff for each $t \in \mathbf{R}^1$

$$f_n(t) - g_n(t) \to 0 \quad as \ n \to \infty. \tag{18}$$

Proof. (Necessity of (18)). We note that $\sin x$ and $\cos x$ are continuous and bounded functions. Hence, for any $t \in \mathbf{R}^1$ from (17) it follows that

$$\mid f_n(t) - g_n(t) \mid \leq \mid \mathrm{E}\cos(tX_n) - \mathrm{E}\cos(tY_n) \mid + \mid \mathrm{E}\sin(tX_n) - \mathrm{E}\sin(tY_n) \mid \to 0$$

as $n \to \infty$.
(Sufficiency of (18)). See Belyaev [2].

Lemma 5 *If $(Y_n)_{n \geq 1}$ is a uniformly tight sequence of real valued r.v.'s, then the sequence of corresponding ch.f.'s $(g_n(\cdot))_{n \geq 1}$, $g_n(t) := \mathrm{E}e^{itY_n}$, $t \in \mathbf{R}^1$, is uniformly equicontinuous, i.e. for any $\varepsilon > 0$ there is a $\delta > 0$ such that*

$$\sup_{n \in \mathbb{Z}^+} \sup_{h:\, |h| \leq \delta} \sup_{t \in \mathbf{R}^1} \mid g_n(t+h) - g_n(t) \mid \leq \varepsilon. \tag{19}$$

Proof. See Belyaev [2].

Let $(\mathcal{L}(X_n \mid Z_n))_{n \geq 1}$ and $(\mathcal{L}(Y_n))_{n \geq 1}$ be two sequences of d.l.'s as introduced in Definition 2. We denote by $f_n(t \mid Z_n) := \mathrm{E}(e^{itX_n} \mid Z_n)$ the ch.f. of X_n given Z_n. As before let $g_n(t) := \mathrm{E}e^{itY_n}$.

Lemma 6 *If $(Y_n)_{n \geq 1}$ is uniformly tight and if for any $t \in \mathbf{R}^1$*

$$\mid f_n(t \mid Z_n) - g_n(t) \mid \overset{\mathrm{P}}{\to} 0 \quad as \ n \to \infty, \tag{20}$$

then the sequence of r.v.'s $(X_n \mid Z_n)_{n \geq 1}$ is uniformly tight along $(Z_n)_{n \geq 1}$ in probability.

8

Proof. We recall that for a r.v. V with ch.f. $h(\cdot)$, and any $u > 0$, the following inequality is valid

$$P\left(\mid V \mid > \frac{1}{u}\right) \leq \frac{7}{u}\int_0^u (1 - \operatorname{Re} h(t))dt\,,$$

see e.g. Shiryaev [10], page 324.

This inequality implies that a.s. for any given Z_n

$$P\left(\mid X_n \mid > \frac{1}{u} \mid Z_n\right) \leq \frac{7}{u}\int_0^u (1 - \operatorname{Re}f_n(t \mid Z_n))dt \leq \frac{7}{u}\int_0^u \mid 1 - g_n(t) \mid dt$$
$$+ \frac{7}{u}\int_0^u \mid g_n(t) - f_n(t \mid Z_n) \mid dt\,. \tag{21}$$

Let ε be any small value ($0 < \varepsilon < 1$) and let $a_\varepsilon > 0$ be large enough such that

$$P(\mid Y_n \mid > a_\varepsilon) < \frac{\varepsilon}{42}. \tag{22}$$

We take $u_\varepsilon > 0$ such that $3u_\varepsilon a_\varepsilon \leq \varepsilon/42$. Then it follows

$$
\begin{aligned}
I_n(\varepsilon) \;:=\; & \sup_{0 \leq t \leq u_\varepsilon} E \mid 1 - e^{itY_n} \mid I(\mid Y_n \mid \leq a_\varepsilon) \\
\leq \; & \sup_{0 \leq t \leq u_\varepsilon} E\left(2\sin^2\left(\frac{tY_n}{2}\right) + \mid \sin(tY_n) \mid\right) I(\mid Y_n \mid \leq a_\varepsilon) \\
\leq \; & 2\sin^2\left(\frac{u_\varepsilon a_\varepsilon}{2}\right) + \mid \sin(u_\varepsilon a_\varepsilon) \mid \; \leq 3u_\varepsilon a_\varepsilon \leq \frac{\varepsilon}{42}.
\end{aligned}
$$

By using (22) we have

$$
\begin{aligned}
\sup_{0 \leq t \leq u_\varepsilon} \mid 1 - g_n(t) \mid \; \leq \; & I_n(\varepsilon) + \sup_{0 \leq t \leq u_\varepsilon} E(\mid 1 - e^{itY_n} \mid I(\mid Y_n \mid > a_\varepsilon)) \\
\leq \; & \frac{\varepsilon}{42} + 2P(\mid Y_n \mid > a_\varepsilon) \leq \frac{\varepsilon}{14}\,. \tag{23}
\end{aligned}
$$

The inequality $\mid g_n(t) - f_n(t \mid Z_n) \mid \leq 2$, assumption (20) and Lebesgue's Dominated Convergence Theorem give

$$E\int_0^{u_\varepsilon} \mid g_n(t) - f_n(t \mid Z_n) \mid dt = \int_0^{u_\varepsilon} E \mid g_n(t) - f_n(t \mid Z_n) \mid dt \to 0$$

as $n \to \infty$.

It follows that

$$J_n(\varepsilon) := \frac{7}{u_\varepsilon}\int_0^{u_\varepsilon} \mid g_n(t) - f_n(t \mid Z_n) \mid dt = o_p(1), \quad n \to \infty. \tag{24}$$

From relations (21), and (23) we have

$$P\left(\mid X_n \mid > \frac{1}{u_\varepsilon} \mid Z_n\right) \leq 7 \sup_{0 \leq t \leq u_\varepsilon} \mid 1 - g_n(t) \mid + J_n(\varepsilon) \leq \frac{\varepsilon}{2} + J_n(\varepsilon)\,.$$

Therefore, by (24) it follows that

$$\mathrm{P}\left(\mathrm{P}\left(\mid X_n\mid > \frac{1}{u_\varepsilon}\mid Z_n\right) > \varepsilon\right) \le \mathrm{P}\left(\frac{\varepsilon}{2} + J_n(\varepsilon) > \varepsilon\right) = \mathrm{P}\left(J_n(\varepsilon) > \frac{\varepsilon}{2}\right) \to 0$$

as $n \to \infty$. This relation holds for arbitrarily small $\varepsilon > 0$ and, with regard to Definition 3, for $k_\varepsilon = 1/u_\varepsilon$. $\qquad\square$

Lemma 7 *If* $((X_n \mid Z_n))_{n\ge 1}$ *is uniformly tight along* $(Z_n)_{n\ge 1}$ *in probability and*

$$w_n(\alpha \mid Z_n) := \sup_{|h|\le\alpha}\ \sup_{t\in\mathbf{R}^1}\mid f_n(t+h\mid Z_n) - f_n(t\mid Z_n)\mid,$$

then for arbitrarily small $\varepsilon' > 0$

$$\sup_n \mathrm{P}(w_n(\alpha \mid Z_n) > \varepsilon') \to 0 \quad as \ \alpha \to 0. \tag{25}$$

Proof. Let ε, δ be arbitrarily small positive numbers, and let $c_{\varepsilon\delta}$ be such that

$$\sup_n \mathrm{P}(\mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta}\mid Z_n) > \varepsilon) < \delta. \tag{26}$$

We consider the events

$$A_{\varepsilon\delta n} = (\omega:\ \mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta}\mid Z_n) > \varepsilon),$$

$$A^c_{\varepsilon\delta n} = (\omega:\ \mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta}\mid Z_n) \le \varepsilon).$$

(26) implies that $\sup_n \mathrm{P}(A_{\varepsilon\delta n}) \le \delta$, and we have

$$\begin{aligned}
\mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta}) &= \mathrm{E}(\mathrm{E}(I(\mid X_n\mid > c_{\varepsilon\delta})\mid Z_n))\\
&= \mathrm{E}(\mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta}\mid Z_n)I(A_{\varepsilon\delta n}))\\
&\quad + \mathrm{E}(\mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta}\mid Z_n)I(A^c_{\varepsilon\delta n}))\\
&\le \mathrm{P}(A_{\varepsilon\delta n}) + \varepsilon\mathrm{P}(A^c_{\varepsilon\delta n}) \le \delta + \varepsilon
\end{aligned} \tag{27}$$

for any $n = 1, 2, \dots$.

If $\mid X_n\mid\le c_{\varepsilon\delta}$ and $h \le 1/c_{\varepsilon\delta}$, then

$$\begin{aligned}
u_n(h\mid Z_n) &:= \sup_{t\in\mathbf{R}^1}\mid f_n(t+h\mid Z_n) - f_n(t\mid Z_n)\mid\\
&\le \mathrm{E}(\mid e^{ihX_n}-1\mid (I(\mid X_n\mid\le c_{\varepsilon\delta}) + I(\mid X_n\mid > c_{\varepsilon\delta}))\mid Z_n)\\
&\le 2\sin^2(h\,c_{\varepsilon\delta}/2) + \mid\sin(h\,c_{\varepsilon\delta})\mid + 2\mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta}\mid Z_n)\\
&\le 2(h\,c_{\varepsilon\delta} + \mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta}\mid Z_n)).
\end{aligned} \tag{28}$$

If $\mid h\mid\le \alpha \le c^{-1}_{\varepsilon\delta}$, then from (27), (28) and the Markov inequality, with any $\varepsilon' > 0$, it follows that

$$\begin{aligned}
\mathrm{P}(w_n(\alpha\mid Z_n) > \varepsilon') = \mathrm{P}\left(\sup_{|h|\le\alpha} u_n(h\mid Z_n) > \varepsilon'\right) &\le \frac{1}{\varepsilon'}\mathrm{E}w_n(\alpha\mid Z_n)\\
&\le \frac{1}{\varepsilon'}2(c_{\varepsilon\delta}\alpha + \mathrm{P}(\mid X_n\mid > c_{\varepsilon\delta})) \le \frac{2}{\varepsilon'}(c_{\varepsilon\delta}\alpha + \delta + \varepsilon).
\end{aligned} \tag{29}$$

10

For any given $\varepsilon' > 0$ we can choose so small $\varepsilon > 0$, and $\delta > 0$, that $\delta + \varepsilon \le \varepsilon'^2/4$. Then we have $c_{\varepsilon\delta} < \infty$ and we can take $\alpha \le \varepsilon'^2/(4c_{\varepsilon\delta})$. Hence (29) gives $P(w_n(\alpha \mid Z_n) > \varepsilon') \le \varepsilon'$ for all $n = 1, 2, \dots$. We thus have

$$\sup_n P(w_n(\alpha \mid Z_n) > \varepsilon') \le \varepsilon'. \tag{30}$$

Simce ε' can be chosen arbitrarily small (30) is equivalent to (25), and so the lemma is proved. $\qquad\square$

Corollary 2 *There is a sequence* $(\alpha_n)_{n \ge 1}$, $\alpha_n \downarrow 0$, $n \to \infty$, *such that almost surely*

$$w_n(\alpha_n \mid Z_n) \to 0 \quad as \ n \to \infty. \tag{31}$$

Proof. Let $\varepsilon_r \to 0$, $r \to \infty$ and let $\alpha_0 = 1$. By using Lemma 7 for any fixed $r = 1, 2, \dots$ we can find a small $\alpha_r > 0$ such that $\alpha_r < \alpha_{r-1}$

$$\sup_n P(w_n(\alpha_r \mid Z_n) > \varepsilon_r) < \frac{1}{2^r}.$$

Then, of course,

$$Q_r = P(w_r(\alpha_r \mid Z_r) > \varepsilon_r) < \frac{1}{2^r},$$

and $\sum_{r=1}^{\infty} Q_r < \infty$. By the Borel - Cantelli lemma almost surely there exists an $r(\omega)$ such that for all $r \ge r(\omega)$

$$0 \le w_r(\alpha_r \mid Z_r) \le \varepsilon_r,$$

$\qquad\square$

and (31) follows.

The following theorem can be applied to justify certain resampling schemes used to obtain consistent estimation of the d.l.'s of sums of r.v.'s, and in many other applications.

Theorem 2 (Continuity theorem for sequences of random d.l.'s that weakly approach deterministic sequences of d.l.'s in probability) *Let* $(X_n, Z_n)_{n \ge 1}$, *and* $(Y_n)_{n \ge 1}$ *be two sequences of r.v.'s as in Definition 2, and let* $(Y_n)_{n \ge 1}$ *be uniformly tight. Then*

$$\mathcal{L}(X_n \mid Z_n) \overset{wa(P)}{\longleftrightarrow} \mathcal{L}(Y_n) \quad as \ n \to \infty, \tag{32}$$

iff for each $t \in \mathbf{R}^1$

$$f_n(t \mid Z_n) - g_n(t) \overset{P}{\to} 0 \quad as \ n \to \infty. \tag{33}$$

11

Proof. Assume that (32) holds. For any $t \in \mathbf{R}^1$ both functions $a_t(x) = \cos(tx)$ and $b_t(x) = \sin(tx)$, $x \in \mathbf{R}^1$, are in $\mathcal{C}_b(\mathbf{R}^1)$. Therefore, for each $t \in \mathbf{R}^1$ we have as $n \to \infty$

$$\mathrm{E}(a_t(X_n) \mid Z_n) - \mathrm{E}a_t(Y_n) \xrightarrow{\mathrm{P}} 0, \quad \mathrm{E}(b_t(X_n) \mid Z_n) - \mathrm{E}b_t(Y_n) \xrightarrow{\mathrm{P}} 0. \quad (34)$$

We have $f_n(t \mid Z_n) = \mathrm{E}(a_t(X_n) \mid Z_n) + i\mathrm{E}(b_t(X_n) \mid Z_n)$, and $g_n(t) = \mathrm{E}a_t(Y_n) + i\mathrm{E}b_t(Y_n)$. Hence, (34) implies (33).

Now assume that (32) does not hold. Then we can find $c_0(\cdot) \in \mathcal{C}_b(\mathbf{R}^1)$ such that

$$D_n = \mathrm{E}(c_0(X_n) \mid Z_n) - \mathrm{E}c_0(Y_n) \overset{\mathrm{P}}{\nrightarrow} 0 \quad \text{as } n \to \infty. \quad (35)$$

Hence there is a subsequence $(n_k)_{k \geq 1}$ such that for some small $\varepsilon > 0$

$$\gamma = \inf_{n_k} \mathrm{P}(\mid D_{n_k} \mid \geq \varepsilon) > 0. \quad (36)$$

Let $a_n(Z_n) = \mathrm{E}(c_0(X_n) \mid Z_n)$, $b_n = \mathrm{E}c_0(Y_n)$, $n \in \mathbb{Z}^+$. We have $a_{n_k}(Z_n)$, $b_{n_k} \in [-\inf_{t \in \mathbf{R}^1} c_0(t), \sup_{t \in \mathbf{R}^1} c_0(t)]$, $k \in \mathbb{Z}^+$. Let b_0 be an accumulation point of $(b_{n_k})_{k \geq 1}$, e.g. we can choose one such point with the largest absolute value. Then there is a subsequence $(n_{k(l)})_{l \geq 1}$ of the sequence $(n_k)_{k \geq 1}$, along which $b_{n_{k(l)}} \to b_0$, as $l \to \infty$. The sequence of r.v.'s $(Y_{n_{k(l)}})_{l \geq 1}$ is still uniformly tight. Hence, by Prohorov's theorem (see Shiryaev [10], page 318) we can find a r.v. Y_0 and a new subsequence $(\tilde{n}_m)_{m \geq 1}$, $\tilde{n}_m = n_{k(l_m)}$, with the property $\mathcal{L}(Y_{\tilde{n}_m}) \xrightarrow{w} \mathcal{L}(Y_0)$ as $m \to \infty$, where Y_0 is a r.v. By the continuity theorem (see Proposition 1 (i)) for each $t \in \mathbf{R}^1$

$$g_{\tilde{n}_m}(t) \to g_0(t) \quad \text{as } m \to \infty. \quad (37)$$

(33), and (37) imply for each $t \in \mathbf{R}^1$

$$f_{\tilde{n}_m}(t \mid Z_n) - g_0(t) \xrightarrow{\mathrm{P}} 0 \quad \text{as } m \to \infty. \quad (38)$$

For any fixed t it is possible to find a subsequence of $(f_{\tilde{n}_m}(t \mid Z_{\tilde{n}_m}))_{m \geq 1}$ that converges almost surely to $g_0(t)$ (see Shiryaev [10], p. 258). For any given $t_0 > 0$ and any $r \in \mathbb{Z}^+$, we consider $\mathcal{T}_r(t_0) := \{t_{kr} : t_{kr} = \frac{k}{2^r}t_0, \quad k \in \{0, \pm 1, \pm 2, \pm 3, ..., \pm 2^r\}\}$, and $\mathcal{T}_r(t_0) \subset \mathcal{T}_{r+1}(t_0)$. We can take a subsequence $(n_r(m))_{m \geq 1}$ of the sequence $(\tilde{n}_m)_{m \geq 1}$ such that

$$\max_{t = t_{kr} \in \mathcal{T}_r(t_0)} \mid f_{n_r(m)}(t \mid Z_{n_r(m)}) - g_0(t) \mid \to 0 \quad \text{a.s. as } m \to \infty. \quad (39)$$

We can select these subsequences $(n_r(m))_{m \geq 1}$ in the following way. We start with $r = 1$ and find $(n_1(m))_{m \geq 1}$. For $r = 2$ we find $(n_2(m))_{m \geq 1}$ as a subsequence of $(n_1(m))_{m \geq 1}$ and so on. By Cantor's method, taking one term $n_r(r)$ from each subsequence $(n_r(m))_{m \geq 1}$, $r \in \mathbb{Z}^+$, we construct the universal subsequence $(n_r(r))_{r \geq 1}$ for which we have

$$\max_{t \in \mathcal{T}_{r_0}(t_0)} \mid f_{n_r(r)}(t \mid Z_{n_r(r)}) - g_0(t) \mid \overset{a.s.}{\to} 0 \quad \text{as } r \to \infty,$$

where r_0 is any fixed number, $r_0 \in \mathbb{Z}^+$. For each $u \in [-t_0, t_0]$ we can find a $t_r(u)$ such that $| t_{r_0}(u) - u | = \min_{t \in \mathcal{T}_{r_0}} | t - u |$. For any r_0, and r we will use the notation:

$$p_{r_0,r}(t_0) = \max_{t \in \mathcal{T}_{r_0}(t_0)} | f_{n_r(r)}(t \mid Z_{n_r(r)}) - g_0(t) |,$$

$$q_{r_0}(t_0) = \max_{-t_0 \leq u \leq t_0} | g_0(t_{r_0}(u)) - g_0(u) |.$$

We have

$$\max_{-t_0 \leq u \leq t_0} | f_{n_r(r)}(u \mid Z_{n_r(r)}) - g_0(u) |$$
$$\leq \max_{-t_0 \leq u \leq t_0} | f_{n_r(r)}(u \mid Z_{n_r(r)}) - f_{n_r(r)}(t_{r_0}(u) \mid Z_{n_r(r)}) |$$
$$+ \quad p_{r_0,r}(t_0) + q_{r_0}(t_0) \leq w_{n_r(r)}(t_0/2^{r_0} \mid Z_{n_r(r)})$$
$$+ \quad p_{r_0,r}(t_0) + q_{r_0}(t_0). \tag{40}$$

By Corollary 2, for any small given number $\beta > 0$, and for almost every $\omega \in \Omega$, we can find an $r_0' = r_0'(\omega)$ such that for all $r \geq r_0'(\omega)$

$$w_{n_r(r)}(t_0/2^{r_0'} \mid Z_{n_r(r)}) < \beta/3.$$

The uniform continuity of the ch.f. $g_0(\cdot)$ implies the existence of an r_0''', such that $q_r(t_0) < \beta/3$ for all $r \geq r_0'''$. From the construction of the sequence $(n_r(r))_{r \geq 1}$ there almost surely exists $r_0'' = r_0''(\omega)$ such that for all $r \geq r_0''(\omega)$, we have $p_{r_0'',r}(t_0) < \beta/3$. We denote $r_0^*(\omega) = \max(r_0''', r_0'(\omega))$. Then for almost all $\omega \in \Omega$ for any $r \geq \max(r_0^*(\omega), r_0''(\omega))$ the inequality

$$\max_{-t_0 \geq u \leq t_0} | f_{n_r(r)}(u \mid Z_{n_r(r)}) - g_0(u) | \leq \beta$$

holds for arbitrary small $\beta > 0$. Hence for any $t \in [-t_0, t_0]$

$$f_{n_r(r)}(t \mid Z_{n_r(r)}) - g_0(t) \overset{a.s.}{\to} 0 \quad \text{as} \quad r \to \infty. \tag{41}$$

By the same argument we can find a subsequence $(n_r(r,2))_{r \geq 1}$ of the sequence $(n_r(r))_{r \geq 1}$ satisfying (41) on $[-2t_0, 2t_0]$. We can then find a subsequence $(n_r(r,3))_{r \geq 1}$ of $(n_r(r,2))_{r \geq 1}$ satisfying (41) on $[-4t_0, 4t_0]$, and so on. We take the subsequence $(n_r(r,r))_{r \geq 1}$ to be "universal" in the sense that, for any $t \in \mathbf{R}^1$, along almost every sequence $(Z_{n_r(r,r)})_{r \geq 1}$

$$f_{n_r(r,r)}(t \mid Z_{n_r(r,r)}) - g_0(t) \to 0 \quad \text{as} \quad r \to \infty. \tag{42}$$

We stress that (42) holds for any $t \in \mathbf{R}^1$ when we have $(Z_{n_r(r,r)}(\omega))_{r \geq 1}$, i.e. we almost surely don't need to change ω when we consider all $t \in \mathbf{R}^1$. Therefore, we can use Theorem 1, and along almost every sequence $(Z_{n_r(r,r)})_{r \geq 1}$ we have

$$\mathcal{L}(X_{n_r(r,r)} \mid Z_{n_r(r,r)}) \overset{w}{\to} \mathcal{L}(Y_0) \quad \text{as} \quad r \to \infty.$$

13

This implies that

$$a_{n_r(r,r)}(Z_{n_r(r,r)}) - b_{n_r(r,r)} \overset{a.s.}{\to} 0 \quad \text{as} \quad n \to \infty .$$

Since this contradicts (36), it follows that (35) cannot hold. Thus the theorem is proved. $\square$

The following auxiliary r.v.'s will be used. Let $J_{1n}^\star, ..., J_{nn}^\star$ be i.i.d. r.v.'s, $P^\star(J_{in}^\star = h) = 1/n$, $h = 1, ..., n$, $M_{hn}^\star := \sum_{j=1}^n I(J_{jn}^\star = h)$, $M_{\cdot n}^\star = M_{1n}^\star + \cdots + M_{nn}^\star = n$. We use $P^\star(X^\star \in A)$ and $E^\star X^\star$ for probability expectation when a random variable $X^\star$ is obtained via the r.v.'s $J_{1n}^\star, ..., J_{nn}^\star$. We have

$$E^\star M_{hn}^\star = \sum_{j=1}^n E^\star I(J_{jn}^\star = h) = 1 , \tag{43}$$

$$\begin{aligned} & E^\star (M_{hn}^\star - 1)^2 \\ = & \sum_{j=1}^n E^\star I(J_{jn}^\star = h) + \sum\sum_{j_1 \neq j_2} E^\star I(J_{j_1,n}^\star = h) E^\star I(J_{j_2,n}^\star = h) - 1 \\ = & \; 1 + n(n-1)/n^2 - 1 = 1 - 1/n , \end{aligned} \tag{44}$$

$$\begin{aligned} E^\star (M_{hn}^\star - 1)(M_{gn}^\star - 1) & = \sum\sum_{j_1 \neq j_2} E^\star I(J_{j_1,n}^\star = h) E^\star I(J_{j_2,n}^\star = g) - 1 \\ & = n(n-1)/n^2 - 1 = -1/n , \quad h \neq g, \; h, g \in \{1, 2, ..., n\} . \end{aligned} \tag{45}$$

We also let $M_{hrn}^\star := \sum_{j=1}^r I(J_{jn}^\star = h)$, and have by similar calculations

$$E^\star M_{hrn}^\star = \frac{r}{n}, \quad E^\star \left(M_{hrn}^\star - \frac{r}{n} \right)^2 = \frac{r}{n} \left(1 - \frac{1}{n} \right) , \tag{46}$$

$$E^\star \left(M_{h_1 rn}^\star - \frac{r}{n} \right) \left(M_{h_2 rn}^\star - \frac{r}{n} \right) = -\frac{r}{n^2} , \quad h_1 \neq h_2 . \tag{47}$$

One can regard $J_{1n}^\star, .., J_{nn}^\star$ as the results of n independent multinomial experiments with n possible outcomes $\{1, 2, ..., n\}$. In fact, $M_{hn}^\star$ is the number of repetitions of outcome h after n independent multinomial experiments, and $M_{hrn}^\star$ after the first r independent multinomial experiments. We can use the r.v.'s $M_{hn}^\star$, $M_{hrn}^\star$, $h = 1, ..., n$ to define others r.v.'s.

Now we recall one result of Belyaev [1] related to the problem of the asymptotically exact approximation of $\mathcal{L}(U_{\cdot n})$ mentioned in the Introduction. The following assumptions will be used:

$AI:$ $U_{in}, i = 1, ..., n,$ are independent r.v.'s for each $n = 1, 2, ...,$

$AL(2 + \delta):$ for $\tilde{U}_{in} := \sqrt{n} U_{in}, \quad E \, | \, \tilde{U}_{in} \, |^{2+\delta} \leq c(2 + \delta) < \infty ,$
$$i = 1, ..., n, \; n = 1, 2, ...$$

14

The assumptions $AVar$ in Belyaev [1] can be dropped.

Let us define the r.v. $U_{\cdot n}^{\star} := \sum_{h=1}^{n} M_{hn}^{\star} u_{hn}$. It is possible to interpret $U_{\cdot n}^{\star}$ as a sum of values, n times randomly sampled with replacement from the components in the statistical data $\mathbf{u}_n = (u_{1n}, ..., u_{nn})$. After centering $U_{\cdot n}^{\star}$ by $u_{\cdot n} = \sum_{i=1}^{n} u_{in}$ we obtain

$$U_{\cdot n}^{\star 0} := \sum_{h=1}^{n} M_{hn}^{\star} u_{hn} - u_{\cdot n} = \sum_{h=1}^{n} (M_{hn}^{\star} - 1) u_{hn}. \tag{48}$$

In fact, the d.l. of the r.v. $U_{\cdot n}^{\star 0}$ defined in (48) is the conditional d.l. given $\mathbf{U}_n = \mathbf{u}_n$, and we denote it by $\mathcal{L}(U_{\cdot n}^{\star 0} \mid \mathbf{u}_n)$. We use the same notation for $U_{\cdot n}^{\star 0}$ if we consider in (48) the r.v.'s U_{hn} instead of their fixed (observed) values u_{hn}, $h = 1, ..., n$. In this case we write $\mathcal{L}(U_{\cdot n}^{\star 0} \mid \mathbf{U}_n)$ for the d.l. of $U_{\cdot n}^{\star 0}$. Now we can formulate the following theorem, see Belyaev [1].

Theorem 3 (The Central Limit Resampling Theorem (CLRT)) *If assumptions $AI, AL(2 + \delta)$ hold then*

$$\mathcal{L}(U_{\cdot n}^{\star 0} \mid \mathbf{U}_n) \overset{wa(\mathrm{P})}{\longleftrightarrow} \mathcal{L}(U_{\cdot n}), \quad n \to \infty,$$

along the sequence of r.v.'s $(\mathbf{U}_n)_{n \geq 1}$.

The d.l.'s of the r.v.'s $U_{\cdot rn}^{\star 0} = \sqrt{\frac{n}{r}} \sum_{h=1}^{n} \left(M_{hrn}^{\star} - \frac{r}{n} \right) u_{hn}$ can also be used to approximate the d.l.'s $\mathcal{L}(U_{\cdot})$ if $r = r(n) \to \infty$, and $r = o(n)$, as $n \to \infty$. Theorem 3 can be applied to obtain approximations of the d.l.'s of deviations of ordinary least squares estimators (OLS-estimators) from true parameters of heteroscedastic linear regression.

Let $Y_i = \sum_{j=1}^{k} x_{ij} \beta_{j0} + W_i$ be an observed random response in the ith experiment, with a vector $\mathbf{x}_i = (x_{i1}, ..., x_{ik})^T$ of factors, a vector $\boldsymbol{\beta}_0 = (\beta_{10}, ..., \beta_{k0})^T$ of unknown true parameters, and with an unobserved value of random error W_i, $i = 1, ..., n$. We assume that $W_1, ..., W_n$ are independent r.v.'s, $EW_i = 0$, $EW_i^2 = \sigma_i^2$, where σ_i^2 are unknown and can be different for different i. Let $\mathbf{Y}_n = (Y_1, ..., Y_n)^T$, $\mathbb{X}_n = (x_{ij})$, $i = 1, ..., n$, $j = 1, ..., k$, and further let $(\mathbb{X}_n^T \mathbb{X}_n)^+$ be the pseudoinverse matrix for $(\mathbb{X}_n^T \mathbb{X}_n)$ and $\mathrm{tr}(\mathbb{X}_n^T \mathbb{X}_n)^+$ be the trace of $(\mathbb{X}_n^T \mathbb{X}_n)^+$, Rao [8]. We write $\mathbb{A} \preceq \mathbb{B}$ if $\mathbb{B} - \mathbb{A}$ is a non-negatively defined matrix. Let $\boldsymbol{\beta}_0$ be estimable, i.e. $\boldsymbol{\beta}_0 = (\mathbb{X}_n^T \mathbb{X}_n)^+ (\mathbb{X}_n^T \mathbb{X}_n) \boldsymbol{\beta}_0$. Then the unbiased OLS-estimator $\hat{\boldsymbol{\beta}}_n = (\mathbb{X}_n^T \mathbb{X}_n)^+ \mathbb{X}_n^T \mathbf{Y}_n$ exists, Searly [9]. Let $\mathbf{c} = (c_1, ..., c_k)^T$ be a given vector, $\| \mathbf{c} \| = (c_1^2 + \cdots + c_k^2)^{1/2} = 1$, $q_n(\mathbf{c}) := (\mathbf{c}^T (\mathbb{X}_n^T \mathbb{X}_n)^+ \mathbf{c})^{1/2}$, $a_{in}(\mathbf{c}) := \mathbf{c}^T (\mathbb{X}_n^T \mathbb{X}_n)^+ \mathbf{x}_i / q_n(\mathbf{c})$. The typical problem, and one of great interest in heteroscedastic linear regression theory, is to determine the d.l. of the weighted deviations of the OLS-estimators from the true parameters, this is given by the r.v.

$$U_n(\mathbf{c}) = \mathbf{c}^T (\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}_0) / q_n(\mathbf{c}).$$

15

Let $Y_i - \hat{Y}_{in}$ be the residuals, $\hat{Y}_{in} = \mathbf{x}_i^T(\mathbb{X}_n^T\mathbb{X}_n)^+\mathbb{X}_n^T\mathbf{Y}_n$, $i = 1, ..., n$. We define the random variable

$$U_n^\star(\mathbf{c}) = \sum_{i=1}^n a_{in}(\mathbf{c})(Y_i - \hat{Y}_{in})(M_{in}^\star - 1),$$

where the r.v.'s $M_{in}^\star$, $i = 1, ..., n$ are as before.

Theorem 4 (Belyaev [1]) *Assume that*

(i) $\sup_n \operatorname{tr}(\mathbb{X}_n^T\mathbb{X}_n)^+/q_n(\mathbf{c}) < \infty$,   *and* $\operatorname{tr}(\mathbb{X}_n^T\mathbb{X}_n)^+ \to 0$ *as* $n \to \infty$;

(ii) $\sqrt{n} \sup_i \mathbf{c}^T(\mathbb{X}_n^T\mathbb{X}_n)^+\mathbf{x}_i/q_n(\mathbf{c}) < \infty$;

(iii)   *for some* $0 < \sigma_- \le \sigma_+ < \infty$ *and any* $n = 1, 2, ...$

$$\sigma_-^2(\mathbb{X}_n^T\mathbb{X}_n) \preceq \sum_{i=1}^n \sigma_i^2\mathbb{X}_i\mathbb{X}_i^T \preceq \sigma_+^2(\mathbb{X}^T\mathbb{X}).$$

Then as $n \to \infty$ *along* $(\mathbf{Y}_n)_{\ge 1}$

$$\mathcal{L}(U_n^\star(\mathbf{c}) \mid \mathbf{Y}_n) \stackrel{wa(\mathrm{P})}{\longleftrightarrow} \mathcal{L}(U_n(\mathbf{c})).$$

Note. The d.l. of $U_n^\star(\mathbf{c})$ given $\mathbf{Y}_n$ can be obtained by a simulation of the r.v.'s $M_{in}^\star$, $i = 1, ..., n$. Actually we obtain this d.l. by the random resampling from the centered weighted residuals $a_{in}(\mathbf{c})(Y_i - \hat{Y}_{in})$, $i = 1, ..., n$, (compare with recommendations in Wu [12]). The left hand-side inequality in Theorem 4 (iii) can be dropped.

3   Acknowledgements

References

[1] Belyaev, Yu.K. (1995) Bootstrap, Resampling and Mallows Metric. Lecture Notes **1**, Institute of Mathematical Statistics, Umeå University.

[2] Belyaev, Yu.K. (1996) Central Limit Resampling Theorems for $m-$Dependent Heterogeneous Random Variables. *Research Report 1996-5, Department of Mathematical Statistics, Umeå University, Sweden.*

[3] Belyaev, Yu.K. and Gadjiev, A. G. (1979) Statistical Analysis of Simulated Results. *Proc. of Academy of Sciences USSR*, ser. *Technical Cybernetics*, 4, 87-93 (in Russian).

[4] Belyaev, Yu.K. and Sjöstedt, S. (1996) Resampling Theorems for Vector Valued Heterogenious Random Variables. *Research Report 1996-9, Department of Mathematical Statistics, Umeå University, Sweden.* (In print)

[5] Dudley, R. (1989) Real analysis and Probability. *Wadsworth and Brooks/Cole.*

[6] Knight, K. (1989) On the bootstrap of the sample mean in the infinite variance case. *Annals of Statistics*, **17**, 1168-1175.

[7] Liu, R.Y. (1988) Bootstrap procedures under some non-i.i.d. models. *Annals of Statistics*, **16**, 1696-1708.

[8] Rao, R.C. (1973) Linear Statistical Inference and Its Applications. 2nd ed. *Wiley*, N.Y.

[9] Searly, S.R. (1987) Linear Models for Unbalanced Data. *Wiley*, N.Y.

[10] Shiryaev, A.N. (1996) Probability. 2nd ed., *Springer Verlag*, Berlin.

[11] White, H. (1980) A Heteroskedasticity Consistent Covariance Matrix Estimator and Direct Test of Heteroskedasticity. *Econometrica*, **48**, 817-838.

[12] Wu, C.F.J. (1986) Jacknife bootstrap and other resampling methods in regression analysis. *Annals of Statistics*, **14**, 1261-1295.

Research Reports published by the

Department of Mathematical Statistics
Umeå University
S-901 87 Umeå, Sweden

1996