Equivalent conditions on the central limit theorem
for a sequence of probability measures on R

Toshio Mikami*

Hokkaido University

ABSTRACT

In this paper we give equivalent conditions on the central limit theorem in total variation norm for a sequence of probability measures on R. This generalizes Cacoullos, Papathanasiou and Utev's central limit theorem in L^1-norm for a sequence of probability density functions on R. We also give equivalent conditions on the central limit theorem in weak convergence and those on the local limit theorem.

Key words and phrases; central limit theorem, total variation norm.

1. Introduction.

Let f be a probability density function on R such that $\int_R yf(y)dy = \mu$ and $\int_R (y-\mu)^2 f(y)dy = \sigma^2 < \infty$. Cacoullos and Papathanasiou (1989) introduced the following function, called a covariance kernel or ω-function of f, to study the characterization of probability distributions:

$$\omega(x) \equiv \int_{-\infty}^{x} (\mu - y) f(y)dy/[\sigma^2 f(x)] \qquad (1).$$

on the set $\{y \in R : f(y) > 0\}$.

It is known that $\omega(x)f(x)$ is also a probability density function on R and that $f(x)$ is normal iff $\omega(x) \equiv 1$ (see Cacoullos, Papathanasiou and Utev, 1992).

Put

$$\begin{aligned} \phi(x) &\equiv (2\pi)^{-1/2} \exp(-x^2/2), \\ \Phi(x) &\equiv \int_{-\infty}^{x} \phi(y)dy. \end{aligned} \qquad (2).$$

Then the following was given in Cacoullos, Papathanasiou and Utev, 1994 (Theorem 1.2), which can also be stated as follows.

* Toshio Mikami, Dep. of Mathematics, Hokkaido University, Sapporo 060, Japan

Theorem 1.1. *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of probability density functions on R with $\int_R yf_n(y)dy = 0$ and $\int_R y^2 f_n(y)dy = 1$ $(n \geq 1)$ and with interval supports. Denoting by ω_n the ω-function of f_n $(n \geq 1)$, the following holds;*

$$\begin{aligned} &\lim_{n\to\infty} \int_R |\phi(x) - f_n(x)|dx = 0 \\ \text{iff} \quad &\lim_{n\to\infty} \int_{|\omega_n(x)-1|>\delta} f_n(x)dx = 0 \quad \text{for any } \delta > 0. \end{aligned} \tag{3}$$

Remark 1.1. It is easy to see that the following is also true (see Cacoullos, Papathanasiou and Utev, 1994, and the proof of Theorem 1.2 below):

$$\begin{aligned} &\lim_{n\to\infty} \int_R |\omega_n(x) - 1|f_n(x)dx = 0 \\ \text{iff} \quad &\lim_{n\to\infty} \int_{|\omega_n(x)-1|>\delta} f_n(x)dx = 0 \quad \text{for any } \delta > 0. \end{aligned} \tag{4}$$

In this paper we extend Theorem 1.1 to a sequence of Borel probability measures on $(R, \mathbf{B}(R))$, using a different method from Cacoullos, Papathanasiou and Utev (CPU), 1994. We also give equivalent conditions on central and local limit theorems.

In section 2 we state our results which will be proved in section 3.

2. Main results.

In this section we state our results.

Let us first introduce our assumption

(A.1). $\{P_n\}_{n=1}^{\infty}$ is a sequence of probability measures on $(R, \mathbf{B}(R))$ such that $\int_R xP_n(dx) = 0$ and that $\int_R x^2 P_n(dx) = 1$ $(n \geq 1)$.

Before we state Theorem 2.1, we give the following definition.

Definition 2.1. For a probability measure P on $(R, \mathbf{B}(R))$ such that $\int_R yP(dy) = 0$ and $\int_R y^2 P(dy) = 1$, put for $x \in R$

$$W(P)(x) \equiv \int_{-\infty}^{x} -yP(dy) \tag{5}$$

(cf. f^* in Lemma 2.2, CPU, 1994).

It is easy to see that $W(P)(x)$ is a probability density function on R since $\int_R yP(dy) = 0$ and $\int_R y^2 P(dy) = 1$. We would also like to point out that $W(P)(\cdot)$ is defined on the whole real line, though $\omega(\cdot)$ is not; in fact this necessitated the restriction to interval support of f in CPU's proofs of the CLT.

Theorem 2.1. *Suppose that (A.1) holds. Then (I), (II) and (III) are equivalent.*

(I). $\lim_{n\to\infty} P_n(dx) = \Phi(dx)$ weakly.
(II). $W(P_n)(\cdot)$ converges to $\phi(\cdot)$, as $n \to \infty$, uniformly on R.
(III). For any $g \in C_0^{\infty}(R; R)$

$$\lim_{n\to\infty}\int_R g(x)(W(P_n)(x)dx - P_n(dx)) = 0.$$

For two probability measures $P(dx)$ and $Q(dx)$ on $(R, \mathbf{B}(R))$, let

$$\rho(P,Q) \equiv \sup_{A\in\mathbf{B}(R)} |P(A) - Q(A)| \qquad (6).$$

denote the total variation distance. For the sake of simplicity, we also write

$$\rho(P,W(P)) \equiv \sup_{A\in\mathbf{B}(R)} |P(A) - \int_A W(P)(x)dx|. \qquad (7).$$

The following result plays a crucial role in the proof of Theorem 2.3.

Proposition 2.2. *For any probability measures $P(dx)$ such that $\int_R xP(dx) = 0$ and that $\int_R x^2P(dx) < \infty$ and $Q(dx) = q(x)dx$ on $(R, \mathbf{B}(R))$ and $r > 0$,*

$$\begin{aligned}&|\rho(P,W(P)) - \rho(P,Q)| \qquad (8).\\ &\le 2r \sup_{|x|\le r} |q(x) - W(P)(x)| + \int_{|x|\ge r} q(x)dx + \int_{|x|\ge r} x^2P(dx).\end{aligned}$$

Finally we state our main result. It turns out that the following equivalence can be shown via the weak convergence on the central limit theorem by way of Proposition 2.2.

Theorem 2.3. *Suppose that (A.1) holds. Then (I) and (II) are equivalent.*

(I). $\lim_{n\to\infty} \rho(P_n, W(P_n)) = 0$.
(II). $\lim_{n\to\infty} \rho(\Phi, P_n) = 0$.
Remark 2.1. Theorem 2.3 generalizes Theorem 1.1 in view of the following: for any probability density functions f and g on R

$$2\sup_{A\in\mathbf{B}(R)} |\int_A f(x)dx - \int_A g(x)dx| = \int_R |f(y) - g(y)|dy.$$

Let us state another assumption and definition to state our final result.
(A.2). $P_n(dx) = f_n(x)dx$ for $n \ge 1$.
Definition 2.2. For a probability density function f on R such that $\int_R yf(y)dy = 0$ and $\int_R y^2f(y)dy = 1$, put for $x \in R$,

$$W(f)(x) \equiv \int_{-\infty}^{x} -yf(y)dy. \qquad (9).$$

$W(f)$ is defined on R and $W(f)(x) = \omega(x)f(x)$ on the set $\{y \in R : f(y) > 0\}$ (cf. (1)).

Proposition 2.4. *Suppose that (A.1)-(A.2) hold. Then (I) and (II) are equivalent.*

(I). $\lim_{n\to\infty} f_n(\cdot) = \phi(\cdot)$, uniformly on every compact subsets of R.
(II). $\lim_{n\to\infty} W(f_n)(\cdot)/f_n(\cdot) = 1$, uniformly on every compact subsets of R.

3. Proof of results in section 2.

Let us first prove Theorem 2.1.

Proof of Theorem 2.1.

(Proof of (II) from (I)). For $x \in R$, take r for which $r > |x|$. Then

$$W(P_n)(x) = \int_{-r}^{x} -yP_n(dy) + \int_{-\infty}^{-r} -yP_n(dy) \equiv I_{n,r} + II_{n,r}.$$

From (I),

$$\begin{aligned} I_{n,r} &\to \int_{-r}^{x} -y\phi(y)dy \quad (\text{as } n \to \infty) \\ &\to \int_{-\infty}^{x} -y\phi(y)dy = \phi(x) \quad (\text{as } r \to \infty); \end{aligned}$$

and from (A.1), by Chebychev's inequality

$$II_{n,r} \le 1/r \to 0 \quad (\text{as } r \to \infty).$$

The convergence of $W(P_n)$ at each point implies the uniform convergence of $W(P_n)$ on R, since $W(P)(\cdot)$ is nondecreasing on $(-\infty, 0]$ and nonincreasing on $[0, \infty)$, and since $\phi(x)$ is continuous on R, and since $\lim_{|x|\to\infty} \phi(x) = 0$.

Q. E. D.

(Proof of (I) from (II)). By (A.1), $\{P_n\}_{n\ge1}$ is tight. Therefore there exist a probability measure Q on $(R, \mathbf{B}(R))$ and a subsequence $\{P_{n_k}\}_{k\ge1}$ which converges weakly to Q as $k \to \infty$.

In the same way as in Proof of (II) from (I), for $x \in R$, $W(P_{n_k})(x)$ converges to $W(Q)(x)$ as $k \to \infty$, and

$$W(Q)(x) = \int_{-\infty}^{x} -yQ(dy) = \phi(x),$$

from (II). This completes the proof.

Q. E. D.

(Proof of (III) from (I)). (II) implies that $W(P_n)(x)dx$ converges to $\phi(x)dx$ as $n \to \infty$, weakly. Since (I) and (II) are equivalent to each other, the proof is over.

Q. E. D.

(Proof of (I) from (III)). For any $\varphi \in C_0^\infty(R; R)$

$$\int_R [\varphi''(x) - x\varphi'(x)]P_n(dx) = \int_R \varphi''(x)[P_n(dx) - W(P_n)(x)dx] \to 0$$

as $n \to \infty$, from (III). This implies (I), since $\{P_n\}_{n\ge1}$ is tight from (A.1) and since $\phi(x)$ is a unique solution of the following ; for any $g \in C_0^\infty(R; R)$

$$\int_R [g''(x) - xg'(x)]f(x)dx = 0,$$

$$\int_R f(x)dx = 1.$$

Q. E. D.

Next let us prove Proposition 2.2.

Proof of Proposition 2.2.

For $r > 0$, put $U_r(o) \equiv \{y \in R; |y| \le r\}$. Then we only have to prove the following;

$$\begin{aligned} &\rho(W(P), Q) \qquad (10). \\ &\le 2r \sup_{|x|\le r} |W(P)(x) - q(x)| + \int_{|x|\ge r} x^2 P(dx) + Q(U_r(o)^c), \end{aligned}$$

since

$$|\rho(P, W(P)) - \rho(P, Q)| \le \rho(W(P), Q).$$

Let us prove (10). For any $A \in \mathbf{B}(R)$,

$$\begin{aligned} &\int_A W(P)(x)dx - Q(A) \\ &= \int_{A\cap U_r(o)} (W(P)(x) - q(x))dx + \int_{A\cap U_r(o)^c} W(P)(x)dx - Q(A \cap U_r(o)^c), \end{aligned}$$

and from the assumption on P,

$$\begin{aligned} \int_{|x|>r} W(P)(x)dx &= \int_{|x|>r} (\int_{-\infty}^{x} -yP(dy))dx \\ &= \int_r^\infty (\int_x^\infty yP(dy))dx + \int_{-\infty}^{-r} (\int_{-\infty}^{x} -yP(dy))dx \\ &= \int_r^\infty y(y-r)P(dy) + \int_{-\infty}^{-r} y(y+r)P(dy) \\ &\le \int_r^\infty y^2 P(dy) + \int_{-\infty}^{-r} y^2 P(dy). \end{aligned}$$

Q. E. D.

Finally we prove Theorem 2.3.

Proof of Theorem 2.3.

(I) and (II) implies (III) and (I) in Theorem 2.1, respectively, and henceforth (II) in Theorem 2.1.

This completes the proof. In fact, by Proposition 2.2, from (A.1), for $r > 0$

$$
\begin{aligned}
&|\rho(P_n, W(P_n)) - \rho(P_n, \Phi)| \qquad\qquad (12).\\
&\le r \sup_{|x|\le r} |\phi(x) - W(P_n)(x)| + \int_{|x|\ge r} \phi(x)dx + 1 - \int_{|x|\le r} x^2 P_n(dx)\\
&\to \int_{|x|\ge r} \phi(x)dx + 1 - \int_{|x|\le r} x^2\phi(x)dx\\
&\qquad \text{(as } n \to \infty\text{, from Theorem 2.1, (I) and (II))}\\
&\to 0 \quad (\text{as } r \to \infty).
\end{aligned}
$$

Q. E. D.

Finally we prove Proposition 2.4.

Proof of Proposition 2.4.

(Proof of (II) from (I)). From (I), by Theorem 2.1, $\lim_{n\to\infty} W(f_n)(\cdot) = \phi(\cdot)$, uniformly on R. This and (I) implies (II).

Q. E. D.

(Proof of (I) from (II)). By (II), for any $g \in C_0^\infty(R; R)$,

$$\lim_{n\to\infty} \int_R g(x)[W(f_n)(x) - f_n(x)]dx = 0.$$

Therefore by Theorem 2.1, $\lim_{n\to\infty} W(f_n)(\cdot) = \phi(\cdot)$, uniformly on R. Hence as $n \to \infty$,

$$f_n(\cdot) = W(f_n)(\cdot)\{W(f_n)(\cdot)/f_n(\cdot)\}^{-1} \to \phi(\cdot),$$

uniformly on every compact subsets of R from (II).

Q. E. D.

References

Bolthausen, E. (1984), An estimate of the remainder in a combinatorial central limit theorem, *Z. Wahrsch. Verw. Gebiete* **66**, 379-386.

Cacoullos, T. (1982), On upper and lower bounds for the variance of a function of a random variable, *Ann. Probab.* **10**, 799-809.

Cacoullos, T. (1989), Dual Poicaré-type inequalities via the Cramér-Rao and the Cauchy-Schwartz inequalities and related characterizations, in: Y. Dodge, ed., *Proc. Statistical Data Analysis and Inference* (North-Holland, Amsterdam) pp. 239-250.

Cacoullos, T. and Papathanasiou, V. (1989), Characterizations of distributions by variance bounds, *Statist. Probab. Lett.* **7**, 351-356.

Cacoullos, T., Papathanasiou, V. and Utev, S. A. (1992), Another characterization of the normal law and a proof of the central limit theorem connected with it, *Theory Probab. Appl.* **37**, 581-588.

Cacoullos, T., Papathanasiou, V. and Utev, S. A. (1994), Variational inequalities with examples and an application to the central limit theorem, *Ann. Probab.* **22**, 1607-1618.

Stein, C. M. (1972), A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, in: L. M. Le Cam, J. Neyman and E. L. Scott, eds., *Proc. Sixth Berkeley Symp. Math. Statist.Probab.* (Univ. California Press, Berkeley) Vol. 2, pp. 583-602.