On Gaussian Random Measures Generated by Empirical Distributions of Independent Random Variables

Włodzimierz Bryc*
Department of Mathematics
University of Cincinnati
PO Box 210025
Cincinnati, OH 45221-0025
bryc@ucbeh.san.uc.edu

Włodzimierz Smolenski†
Department of Mathematics
Siedlce University
Orlicz-Dreszera 19/21
08-110 Siedlce, Poland

March 1, 1994

Abstract

Normalized fluctuations of empirical measures converge to a law of a random measure if and only if the underlying random variable is purely discrete with square-root-summable probabilities.

1 Introduction

Let $\xi, \xi_1, \xi_2, \dots$ be a sequence of i.i.d. real random variables with the law μ. The associated empirical measures are defined by $\frac{1}{n}\sum_{k=1}^n \delta_{\xi_k}$, where δ_x denotes the point mass at $x \in \mathbb{R}$. Various aspects of asymptotic theory for empirical measures have been developed in the literature; see e.g. [3, Section 3.2] (large deviations), [12] (central limit theory) and, in a more general setup, [6] (hydrodynamic limits).

In this note we consider normalized fluctuations of empirical measures, given by random signed measures

$$X_n = \frac{1}{\sqrt{n}}\sum_{k=1}^n (\delta_{\xi_k} - \mu).$$

*Partially Supported by C.P. Taft Memorial Fund
†

Key Words: Normalized fluctuations of empirical measures, random measures.
AMS (1991) Subject Classifications: Primary: 60G57 Secondary: 60F05
Short title: Gaussian Random Measures

It was pointed out by various authors (e.g., [8, 9]) that, in general, we cannot expect the weak limit of X_n to exist as a random measure on $\mathbb{R}$. Therefore, larger spaces are considered; those play little role bellow and are used only to make sense of the objects considered; the reader not comfortable with the space $\mathcal{S}'$ of Schwartz distributions (see [13]), may safely substitute for $\mathcal{S}'$ a separable Hilbert subspace H of $\mathcal{S}'$. When X_n are considered as random variables taking values in the Schwartz space $\mathcal{S}'$, then it is known that X_n converge weakly to a random tempered distribution X (see [9]). The law of X is a symmetric tight Gaussian measure Γ on $\mathcal{S}'$. Its covariance functional $C(f,g)$ for rapidly decreasing $f, g \in \mathcal{S}$ is given by

$$C(f,g) = E\{(f(\xi) - Ef(\xi))(g(\xi) - Eg(\xi))\}, \tag{1}$$

i.e., $C(f,g)$ is equal to the covariance of the real random variables $f(\xi), g(\xi)$.

Conversely, given a real random variable ξ, formula (1) defines a continuous symmetric positive definite bilinear form on $\mathcal{S}$ which, therefore, is a covariance functional of a Gaussian random element X with values in $\mathcal{S}'$, the (topological) dual of $\mathcal{S}$; the law of X is a symmetric tight Gaussian measure Γ on $\mathcal{S}'$. In the terminology of [4], the map $L : \mathcal{S} \ni f \to \langle X, f \rangle$ is called the centered noise random linear functional. Since the space $\mathcal{M}(\mathbb{R})$ of signed measures of finite variation is a Borel linear subspace of the Schwartz space $\mathcal{S}'$, by the zero-one law of Kallianpur [10], $Prob(X \in \mathcal{M}(\mathbb{R}))$ is either 0 or 1. Our main results answers when the probability is one. The result complements [5, Theorem 3.1], who consider the discrete case only; our proof is also very elementary (writing a series expansion that trivially converges or diverges).

Theorem 1 *Let ξ be a real r.v. and let X be a Gaussian $\mathcal{S}'$-valued r.v. with the covariance given by (1). Then $Prob(X \in \mathcal{M}(\mathbb{R})) = 1$ if and only if the following two conditions are fulfilled.*

$$\xi \text{ is discrete;} \tag{2}$$

$$\sum_x \sqrt{Prob(\xi = x)} < \infty. \tag{3}$$

Moreover, if (2) and (3) hold, then X takes values in the set $\mathcal{M}_\mu(\mathbb{R})$ of measures absolutely continuous with respect to μ.

Random variables X_n are measure valued and from Theorem 1 it is clear that we can expect X_n to converge weakly in $\mathcal{M}(\mathbb{R})$, i.e., $\mathcal{M}(\mathbb{R})$-valued i.i.d. random variables δ_{ξ_k} to satisfy the central limit theorem in the Banach space $\mathcal{M}(\mathbb{R})$ of signed measures with bounded variation topology, only if conditions (2) and (3) are fulfilled. This indeed is the case as shown by the following corollary, see Durst & Dudley [5, Theorem 3.1].

Corollary 1 *If ξ satisfies (2) and (3), then r.v. δ_{ξ_k} satisfy the central limit theorem in $\mathcal{M}(\mathbb{R})$.*

Proof: Since μ is discrete, $\mathcal{M}_\mu(\mathbb{R})$ with the induced (total variation) norm topology is isomorphic to the space ℓ_1 of all absolutely summable sequences. It is well known that ℓ_1 is a Banach space of cotype 2 (see, e.g., [1, p. 188]). Therefore (cf., e.g., [1, p. 194]) to prove the theorem it is enough to check that there is a Gaussian $\mathcal{M}_\mu(\mathbb{R})$-valued r.v. X with the covariance given by (1). Clearly, X from Theorem 1 satisfies the requirements. □

Remark 1 *Corollary 1 does not assume any integrability properties of ξ; for related CLT results that assume conditions on tails of ξ, see Gine & Zinn [7].*

It is also of interest to point out that in general, distribution valued r.v. X with the covariance (1) has a series expansion $X = \sum \nu_n \gamma_n$, where ν_n are deterministic measures which are absolutely continuous with respect to μ and γ_n are real i.i.d. N(0, 1) r.v. This fact is a direct consequence of the theory of reproducing kernel Hilbert spaces associated with a Gaussian measure (see, e.g. [11]) and of the following result.

Proposition 1 *For each μ, the reproducing kernel Hilbert space H_X of X is contained in $\mathcal{M}_\mu(\mathbb{R})$.*

In one of the proofs we shall use the following folklore result, which we prove for completeness in a more general form than what is needed below.

Proposition 2 *If Γ_1, Γ_2 are two tight Gaussian measures on a locally convex space E such that their reproducing Hilbert spaces satisfy $H_{\Gamma_1} \subset H_{\Gamma_2}$, then for each Borel subspace L of E,*

$$\Gamma_2(L) \leq \Gamma_1(L).$$

2 Proofs

We shall use the following "abstract" results about Gaussian vectors.

(A) If ν_n are deterministic measures such that for i.i.d. N(0, 1) r.v. γ_n the series $X = \sum \nu_n \gamma_n$ converges in the variation norm $\|\cdot\|$ on $\mathcal{M}(\mathbb{R})$, then $Prob(X \in \mathcal{M}(\mathbb{R})) = 1$.

(B) (see [10]) If Γ is a tight Gaussian measure on a locally convex quasi-complete space E, then for each Borel subspace L of and every vector $v \in E$, $\Gamma(L + v)$ is either 0 or 1.

(C) (see [2]) If Γ_1, Γ_2 are two tight Gaussian measures on a locally convex quasi-complete space E such that their reproducing Hilbert space norms satisfy $|\cdot|_{\Gamma_2} \leq K|\cdot|_{\Gamma_1}$, then there exists a symmetric Gaussian measure Γ_0 and a constant $c > 0$ such that

$$\Gamma_2(cA) = \Gamma_1 * \Gamma_0(A)$$

for all measurable sets A.

2.1 Proof of Proposition 2

We claim that the inclusion $H_{\Gamma_1} \subset H_{\Gamma_2}$ is a continuous embedding . Indeed, let K_i denote the unit ball of $H_{\Gamma_i}, i = 1, 2$. Both sets K_i are compact subsets of E (c.f. [2]) and, since the embedding $H_{\Gamma_1} \subset E$ is continuous, their intersection $K = K_1 \cap K_2$, being closed in E, is closed in H_{Γ_1} . Moreover, H_{Γ_1} is the union of sets $nK, n \geq 1$. Indeed, $H_{\Gamma_i} = \bigcup_{n\geq 1} nK_i$ and for sequences of non-decreasing sets A_n and B_n one has

$$\bigcup_{n\geq 1} A_n \cap \bigcup_{n\geq 1} B_n = \bigcup_{n\geq 1} A_n \cap B_n.$$

By the Baire Theorem, $K \subset n_0 K \subset n_0 K_2$ for some n_0, proving that the embedding is continuous.

Since the inclusion $H_{\Gamma_1} \subset H_{\Gamma_2}$ is continuous, therefore by (C) we have

$$\Gamma_2(L) = \int_E \Gamma_1(L - x)\Gamma_0(dx). \tag{4}$$

By symmetry $\Gamma_1(L - x) = \Gamma_1(L + x)$ and for $x \in L$, sets $L + x$ and $L - x$ are disjoint affine subspaces of E. Therefore from (4) it follows that $\Gamma_1(L - x) < 1$ and by the zero-one law (for Borel affine subspaces) $\Gamma_1(L - x) = 0$ for $x \in L$. This shows that $\Gamma_2(L) = \Gamma_1(L)\Gamma_0(L) \leq \Gamma_1(L)$. □

2.2 Proof of Theorem 1

(Sufficiency) Suppose (2) and (3) hold. Denote by r_n the values of ξ and put $p_n = Prob(\xi = r_n) = \mu(r_n)$. Let (γ_n) be a sequence of independent standard normal N(0, 1) r. v. The series $X = \sum \sqrt{p_n}(\delta_{r_n} - \mu)\gamma_n$ converges in the variation norm and hence X is an $\mathcal{M}_\mu(\mathbb{R})$-valued Gaussian random variable. A direct computation shows that for $f \in \mathcal{S}$ we have $E\langle X, f\rangle^2 = E(\sum \sqrt{p_n}(f(r_n) - \int f d\mu)\gamma_n)^2 = \sum p_n(f(r_n) - \int f d\mu)^2$, which matches (1). □

(Necessity) Suppose that either (2) or (3) fails. We shall show that this contradicts $Prob(X \in \mathcal{M}(\mathbb{R})) = 1$.

Let γ_0 be a normal N(0, 1) random variable independent of X and put

$$Y = X + \gamma_0 \mu. \tag{5}$$

The covariance of Y is given by

$$E\langle Y, f\rangle^2 = \int f^2 d\mu. \tag{6}$$

Since μ is in $\mathcal{M}(\mathbb{R})$, therefore the events $X \in \mathcal{M}(\mathbb{R})$ and $Y \in \mathcal{M}(\mathbb{R})$ are identical and hence $Prob(X \in \mathcal{M}(\mathbb{R})) = Prob(\mathrm{Y} \in \mathcal{M}(\mathbb{R}))$. We shall show that $Prob(Y \in \mathcal{M}(\mathbb{R})) = 0$. If either (2) or (3) fails, then one can find disjoint open intervals J_n such that

$$\mu(J_n) = \mu(\overline{J_n}) = q_n,$$

and

$$\sum \sqrt{q_n} = \infty.$$

Let ν_n be a restriction of μ to J_n, i.e. ν_n are supported on $\overline{J_n}$ and

$$\int f d\nu_n = \int_{J_n} f d\mu$$

for all bounded measurable f. Denote by $|\cdot|$ the associated Hilbert space norm in H_Y. The formula

$$|\sum a_n \nu_n|^2 = \sum a_n^2 q_n. \tag{7}$$

gives the explicit expression for the reproducing kernel Hilbert space norm of a finite linear combination of measures (ν_n). Indeed,

$$|\sum a_n \nu_n| = \sup\{\sum a_n \int f d\nu_n : f \in \mathcal{S}, \int f^2 d\mu \leq 1\}$$

$$= \sup\{\sum a_n b_n \int f_n d\nu_n : \int f_n^2 d\mu \leq 1, \sum b_n^2 \leq 1\}$$

$$= \sup\{(\sum a_n^2 (\int_{J_n} f_n d\mu)^2)^{1/2} : \int f_n^2 d\mu \leq 1\}.$$

This proves (7), since

$$\sup\{(\int_{J_n} f d\mu)^2 : \int f^2 d\mu \leq 1\} = \mu(J_n) = q_n.$$

From (7) it follows that $(\frac{1}{\sqrt{q_n}}\nu_n)_{n\geq 1}$ is an orthonormal sequence in H_Y. Let $Z = \sum \frac{1}{\sqrt{q_n}}\nu_n \gamma_n$, where (γ_n) are i.i.d. N(0, 1) random variables. The reproducing kernel Hilbert space of Z lies in the reproducing kernel Hilbert space of Y, as its conjugate Hilbert space norm on $\mathcal{S}$ is smaller (Jensen's inequality). Applying Proposition 2 to $E = \mathcal{S}'$ and its linear subspace $L = \mathcal{M}(\mathbb{R})$ we have

$$Prob(Z \in \mathcal{M}(\mathbb{R})) \geq Prob(\mathrm{Y} \in \mathcal{M}(\mathbb{R})).$$

However, the variation norm $\| \sum \frac{1}{\sqrt{q_n}} \nu_n \gamma_n \| = \sum \frac{1}{\sqrt{q_n}} |\gamma_n|$ diverges, i.e., if C_N is a ball of radius N in the variation norm in $\mathcal{M}(\mathbb{R})$, then $Prob(Z \in C_N) = 0$ for all $N \geq 1$. Therefore $0 \leq Prob(X \in \mathcal{M}(\mathbb{R})) = Prob(\mathrm{Y} \in \mathcal{M}(\mathbb{R})) \leq Prob(\mathrm{Z} \in \mathcal{M}(\mathbb{R})) = 0$. $\square$

2.3 Proof of Proposition 1

Let Y be the Gaussian random measure on $\mathcal{S}'$ defined by (5). Then the reproducing kernel Hilbert space of X is contained in the reproducing kernel Hilbert space H_Y of Y, since the covariance (6) dominates (1). We shall show that H_Y is contained in $\mathcal{M}(\mathbb{R})$. To this end notice that by definition H_Y consists of the distributions $T \in \mathcal{S}'$ such that $\sup\{\langle T, f\rangle : \int f^2 d\mu \leq 1\} < \infty$. Each such T is actually given by $< T, f >= \int f(x)g(x)d\mu(x)$ for some $g \in L_2(\mathbb{R}, \mathrm{d}\mu)$. Since μ is a probability measure, therefore $L_2(\mathbb{R}, \mathrm{d}\mu)$ is contained in $L_1(\mathbb{R}, \mathrm{d}\mu)$ and hence each T is a μ-absolutely continuous measure with the density $\frac{dT}{d\mu} = g$.

Acknowledgement. We would like to thank A. Jakubowski, S. Kwapień, and W. Niemiro for helpful comments and encouragement.

References

[1] Araujo A. & Giné E., *The Central Limit Theorem for Real and Banach Valued Random Variables*, Wiley, New York 1980.

[2] Borel C., Gaussian Radon measures on locally convex spaces, *Math. Scand.* 38 (1976), 265–284.

[3] Deuschel, J. D. & Stroock, D. W., *Large Deviations*, Pure and Applied Mathematics vol. 137, Academic Press, Inc, Harcourt Brace Jovanovich, Publishers, Boston, 1989.

[4] Dudley R. M., Gaussian processes on several parameters, *Ann. Math. Statist.* 36 (1965), 771–788.

[5] Durst M. & Dudley R. M., Empirical processes, Vapnik-Chervonenkis classes and Poisson processes, *Prob. Math. Stat.* (Wrocław) 1 (1980), 109–115.

[6] De Masi A. & Presutti E., *Mathematical Methods for Hydrodynamic Limits*, Lecture Notes in Math. v. 1501, Springer, 1991.

[7] Giné E. & Zinn J., Empirical processes indexed by Lipschitz functions, *Ann. Probab.* 14 (1986), 1329–1338.

[8] Holley R. & Stroock D., *Generalized Ornstein-Uhlenbeck processes as Limits of Interacting systems*, Lecture Notes in Math. v. 851, Springer, 1981, 152–168.

[9] Itô K., Distribution-Valued Processes Arising from Independent Brownian Motions, *Math. Zeitschr.* 182 (1983), 17–33.

[10] Kallianpur G., Zero one laws for Gaussian processes, *Trans. Amer. Math. Soc.* 149 (1970), 199–211.

[11] Kuo H. H., *Gaussian Measures in Banach Spaces*, Lecture Notes in Math. v. 463, Springer 1975.

[12] Pollard D., *Empirical processes: Theory and Applications*, NSF-CMBS Regional Conference Series in Probability and Statics Vol. 2, Institute of Math. Statist. 1990.

[13] Schwartz L., *Théorie des distributions*, Hermann, Paris 1966.