⟨Research Paper⟩

Journal of the Korean
Statistical Society
Vol. 14, No. 2, 1985

Laplace's Method for General Integrals with Applications to Statistical Mechanics

Nae Hyun Park* and Jong Woo Jeon**

ABSTRACT

This paper extends the results of Ellis and Rosen (1982 a) to some more general integrals and applies our main theorem to compute the specific free energy of some models in statistical mechanics. The general integrals of this paper mean the integrals with respect to the probability measures induced by the sample mean of n i.i.d. random variables taking values in a separable Banach space.

1. Introduction

Classical Laplace's method for Rieman integrals on R^1 (see Murray(1984)) yields

$$\lim_{n\to\infty} (1/n) \log\int_{-\infty}^{\infty} f(x)e^{-ng(x)}dx = -\inf_{-\infty<x<\infty}\{g(x)\} \tag{1.1}$$

for some continuous functions $f(x)$ and $g(x)$ on R^1. Schilder (1966) and Pincus (1968) extended (1.1) to the Gaussian integrals on $C[0,1]$ (the space of real-valued continuous functions on $[0,1]$). Donsker and Varadhan (1976) considered the Laplace's method for Gaussian integrals on a separable Banach space. Recently Ellis and Rosen (1982a) dealt with the Laplace's method for Gaussian integrals on $C[0,1]$ of the form

$$A_n = \int_{C[0,1]} e^{-nF_n(y)} dQ_n(\sqrt{n}y)$$

where Q_n is a mean zero Gaussian measure on $C[0,1]$ such that $Q_n \Rightarrow Q$ (Q_n converges weakly to Q) for a mean zero Gaussian measure Q on $C[0,1]$ and $\{F_n\}$ is a sequence of suitably bounded and continuous real-valued functions on $C[0,1]$ which tend in some

* Department of Computer Science and Statistics, Chung-nam National University, Dae-jeon 300-31, Korea.
** Department of Computer Science and Statistics, Seoul National University, Seoul 151, Korea.

sense to a function F on $C[0,1]$. Their main theorem which is an extension of the results of Schilder (1966) and Pincus (1968) states that

$$(1/n)\ \log A_n \longrightarrow -\inf_{y\in C[0,1]} \{F(y)+I_Q(y)\} \tag{1.2}$$

as $n\to\infty$, where I_Q is the entropy functional of Q which is defined in the next section.

The purpose of this paper is to extend the Ellis and Rosen's main theorem to some more general integrals. Our main theorem is an extension of (1.2) and unify (1.1) and (1.2). The assumptions of our main theorem are simpler and shorter than those of the Ellis and Rosen's main theorem. We also produce some limit theorems which are related to the main theorem. Finally we apply our main theorem to compute the SFE (specific free energy) of some models in statistical mechanics.

2. Main Results

Let V be a separable Banach space equipped with its Borel σ-field $\mathscr{B}$, $\{\nu_n\}$ be a sequence of p.m.'s on $(V, \mathscr{B})$ such that $\nu_n \Rightarrow \nu$ for some p.m. ν on $(V, \mathscr{B})$ and for all $a>0$,

$$\sup_{n\geq 1}\int_V e^{a\|x\|} d\nu_n(x) < \infty \tag{2.1}$$

where $\|\cdot\|$ denotes a norm in V.

The entropy functional of ν is defined by

$$I_\nu(y) = \sup_{\theta\in V^*} \{\theta(y) - \psi_\nu(\theta)\} \tag{2.2}$$

for $y\in V$, where V^* is the dual of V and

$$\psi_\nu(\theta) = \log\int_V e^{\theta(x)} d\nu(x). \tag{2.3}$$

When $V=C[0,1]$ and $\nu=Q$, it can be shown

$$I_Q(y) = \begin{cases} \frac{1}{2}\langle y, \Lambda^{-1}y\rangle, & \text{for } y\in\mathscr{D} \text{ (domain of } \Lambda^{-1}) \\ \infty & \text{, otherwise} \end{cases} \tag{2.4}$$

where Λ is the covariance operator on $L^2[0,1]$ corresponding to Q and $\langle -, - \rangle$ is the L^2-inner product.

Let $X_1^n, \ldots X_n^n$ be n i.i.d. V-valued r.v.'s (random variables) with common p.m. ν_n and ν_n^n be the p.m. induced by $\overline{X}_n = (1/n)\sum_{j=1}^{n} X_i^n$.

The following theorem is our main result. The principal ingredients in the proof of the main theorem are a large deviations theorem due to Bolthausen (1984) and a limit theorem in Varadhan (1966).

Theorem 2.1 Assume the followings;

(A1) There exist positive constants α, β, γ, not depending on n, such that for all sufficiently large n and all $t \geq \gamma$, $\nu_n{}^n(y; \|y\|^2 > t) \leq \alpha e^{-n\beta t}$.

(A2) $\{F_n\}$ is a sequence of real-valued functions in V such that for all sufficiently large n and all $y \in V$, $-F_n(y) \leq a + b\|y\|^2$ for some finite constants a and b $(0 < b < \beta)$.

(A3) For any $y \in V$ with $I_\nu(y) < \infty$ and for any sequence $\{y_n\}$ with $y_n \to y$ as $n \to \infty$, $F_n(y_n) \to F(y)$ for some real-valued function F on V.

Then

$$\lim_{n\to\infty} (1/n) \log \int_V e^{-nF_n(y)} d\nu_n{}^n(y) = -\inf_{y \in V} \{F(y) + I_\nu(y)\}. \tag{2.5}$$

Proof. According to Bolthausen's large deviations theorem, Theorem 5.2 in Donsker and Varadhan (1976) and Theorem 3.4 in Varadhan (1966), it suffices to show that

$$\lim_{L\to\infty} \overline{\lim_{n\to\infty}} (1/n) \log \int_{\{y; -F_n(y) \geq L\}} e^{-nF_n(y)} d\nu_n{}^n(y) = -\infty \tag{2.6}$$

Since by (A 2) for all sufficiently large n

$$\int_{\{y; -F_n(y) \geq L\}} e^{-nF_n(y)} d\nu_n{}^n(y)$$
$$\leq \int_{\{y; \|y\|^2 \geq (L-a)/b\}} e^{n(a+b\|y\|^2)} d\nu_n{}^n(y)$$

it suffices to prove

$$\lim_{L\to\infty} \overline{\lim_{n\to\infty}} (1/n) \log \int_{\{y; \|y\|^2 \geq L\}} e^{nb\|y\|^2} d\nu_n{}^n(y) = -\infty.$$

Let $Q_n = \nu_n{}^n T^{-1}$, that is $Q_n(B) = \nu_n{}^n[T^{-1}(B)]$ for any $B \in \mathscr{B}$, where $T(y) = \|y\|^2$ for all $y \in V$ and let $G_n(t) = Q_n(-\infty, t] = \nu_n{}^n(y; \|y\|^2 \leq t)$.

Then

$$J_n = \int_{\{y; \|y\|^2 \geq L\}} e^{nb\|y\|^2} d\nu_n{}^n(y) = \int_L^\infty e^{nbt} dQ_n(t)$$
$$= \int_L^\infty e^{nbt} dG_n(t) = \int_L^\infty e^{nbt} d[-P(\|\overline{X}_n\|^2 > t)]$$
$$= -e^{nbt} P(\|\overline{X}_n\|^2 > t) \Big|_L^\infty + nb \int_L^\infty e^{nbt} P(\|\overline{X}_n\|^2 > t) dt.$$

The second equality is by change of variable and the last equality follows from the integration by parts. Since

$0 \leq \lim_{t\to\infty} e^{nbt} P(\|\overline{X}_n\|^2 > t) \leq \lim_{t\to\infty} \alpha e^{n(b-\beta)t} = 0$ and

$e^{nbL}P(\|\overline{X}_n\|^2>L)\leq\alpha\ e^{n(b-\beta)L}$ by (A 1), we have

$$J_n\leq\alpha\ e^{n(b-\beta)L}+nb\int_L^\infty e^{nbt}P(\|\overline{X}_n\|^2>t)\,dt$$

$$\leq\alpha\ e^{n(b-\beta)L}+nb\alpha\int_L^\infty e^{n(b-\beta)t}dt$$

$$=[\alpha+b\alpha/(\beta-b)]e^{n(b-\beta)L}$$

for all sufficiently large L. Therefore

$$\overline{\lim_{n\to\infty}}\ (1/n)\log J_n\leq(b-\beta)L\to-\infty \text{ as } L\to\infty$$

since $0<b<\beta$ and $\alpha>0$. Q.E.D.

Corollary Let μ_n be the p.m. induced by the sample mean of n i.i.d. V-valued r.v.'s whose common p.m. is μ on $(V,\mathscr{B})$. Assume that (A 2) and (A 3) of Theorem 2.1 hold. Suppose (A 1) holds when we replace $\nu_n{}^n$ with μ^n. Then

$$\lim_{n\to\infty}(1/n)\ \log\int_V e^{-nF_n(y)}d\mu^n(y)=-\inf_{y\in V}\{F(y)+I_\mu(y)\}. \tag{2.7}$$

Remark 2.1 If $V=C[0,1]$, $\nu_n=Q_n$ and $\nu=Q$, then $(1/\sqrt{n})\sum_{i=1}^n X_i{}^n$ and $X_1{}^n$ have the same probability distribution Q_n and $\nu_n{}^n(y;\|y\|^2>t)=P(\|X_1{}^n\|^2>nt)$. Thus (A1) is satisfied by Bolthausen (1984) or Ellis and Rosen (1982 a). Also for any $B\in\mathscr{B}$, $\nu_n{}^n(B)=P(X_1{}^n\in\sqrt{n}B)=Q_n(\sqrt{n}B)$. Therefore (2.5) reduces to (1.2) so that Theorem 2.1 is an extension of Ellis and Rosen's main theorem. It should be noted that Theorem 2.1 does not need Hypothesis 3.3 in the Ellis and Rosen (1982 a).

Theorem 2.2 Let ν_n, $\nu_n{}^n$, ν, F and $\{F_n\}$ satisfy the hypotheses of Theorem 2.1. Define

$$W_n(S)=\frac{\int_S e^{-nF_n(y)}d\nu_n{}^n(y)}{\int_V e^{-nF_n(y)}d\nu_n{}^n(y)} \tag{2.8}$$

for any Borel subset S of V. Assume that A is a closed subset of V such that for some $\delta>0$,

$$\inf_{y\in A}\{F(y)+I_\nu(y)\}-\inf_{y\in V}\{F(y)+I_\nu(y)\}\geq\delta.$$

Then $W_n(A)\leq e^{-n\delta}$ for all sufficiently large n.

Proof. Let $d\lambda_n(y)=e^{-nF_n(y)}d\nu_n{}^n(y)$. Then by Theorem 3.5 in Varadhan (1966),

$$\overline{\lim_{n\to\infty}}(1/n)\ \log\lambda_n(A)\leq-\inf_{y\in A}\{F(y)+I_\nu(y)\}.$$

Hence by Theorem 2.1,

$$\overline{\lim_{n\to\infty}}(1/n)\ \log W_n(A)=\overline{\lim_{n\to\infty}}(1/n)\ \log\lambda_n(A)-\overline{\lim_{n\to\infty}}(1/n)\ \log\lambda_n(V)$$

$$\leq -\inf_{y\in A}\{F(y)+I_\nu(y)\}+\inf_{y\in V}\{F(y)+I_\nu(y)\}\leq -\delta.$$

Therefore $(1/n)\log W_n(A)\leq -\delta$ for all sufficiently large n. Q.E.D.

Theorem 2.2 is a generalization of Theorem 1.3 in Ellis and Rosen (1982a).

Theorem 2.3 Let Y_n be a V-valued r.v. whose probability distribution is W_n in Theorem 2.2 and let $G(y)=F(y)+I_\nu(y)$ have a unique minimum point y^*. Then Y_n converges in probability to y^*.

Proof. For any bounded real-valued uniformly continuous function g on V,

$$\frac{\int_V g(y)e^{-nF_n(y)}d\nu_n{}^n(y)}{\int_V e^{-nF_n(y)}d\nu_n{}^n(y)}\longrightarrow g(y^*)=\int_V g(y)\,d\delta y^*$$

as $n\to\infty$ by Theorem 3.6 in Varadhan (1966), where δ_x denotes the degenerate distribution at x. Hence Y_n converges in probability to y^* by Theorem 2.1 in Billingsley (1968). Q.E.D.

The above theorem is a generalization of Theorem 2.5 in Ellis and Rosen (1982b). If $F=0$, then nY_n is distributed like the sum of n i.i.d. r.v.'s whose common probability distribution is ν_n. This sum is an important quantity in a model in statistical mechanics.

Ellis and Rosen (1982b) have studied in great detail the central limit theorems about Y_n when ν_n is a mean zero Gaussian measure on a Hilbert space and $F_n=F$ for all n. Studies on central limit theorems about Y_n for a general p.m. ν_n on $(V,\mathscr{B})$ seem interesting. However this article does not consider this topic which seems difficult to solve.

3. Applications to Statistical Mechanics

In this section we compute the SFE of the models in statistical mechanics by applying Theorem 2.1.

3.1. Generalized Curie-Weiss Model

Let ρ be a p.m. on R^1 satisfying

$$\int_{-\infty}^{\infty} e^{cx^2}d\rho(x)<\infty \tag{3.1}$$

for all $c>0$ and let $\{X_i^{(n)};\ i=1,\dots,n\}$ be a triangular array of dependent and identically distributed r.v.'s with the joint distribution given by

$$z_n^{-1}\exp[n\psi_\mu\{(x_1+\cdots+x_n)/n\}]\prod_{i=1}^{n}d\rho(x_i) \tag{3.2}$$

where $\psi_\mu(\cdot)$ is the cumulant generating function of some r.v. Y whose probability distribution is μ on $(R^1, \mathscr{B})$ and

$$z_n=\int_{R^n}\exp[n\ \psi_\mu\{(x_1+\cdots+x_n)/n\}]\prod_{i=1}^{n}d\rho(x_i) \tag{3.3}$$

is the normalization constant. $X_i^{(n)}$ represents the spin or magnetic moment of the individual atom at i-th site in a magnetic crystal. The Curie-Weiss model is the special case of (3.2) when Y is standard normal r.v.. Note that it can be easily verified that

$$z_n=\int_{-\infty}^{\infty}\exp[-n\{-\psi_\rho(y)\}]d\mu^n(y). \tag{3.4}$$

We define the physical quantity

$$f(\rho)=-\lim_{n\to\infty}(1/n)\ \log\ z_n \tag{3.5}$$

known as the SFE of the model.

For mean zero bounded r.v. Y we can show μ^n satisfies (A 1) in Theorem 2.1 by the result of Prokohrov (1968). Since for any $\tau>0$

$$xy\leq(\tau/2)y^2+(x^2/2\tau)\ \text{ for any } x,y \text{ in } R^1,$$

$$\psi_\rho(y)\leq(\tau/2)y^2+\log\int_{-\infty}^{\infty}e^{x^2/2\tau}d\rho(x)$$

and thus $F_n(y)=-\psi_\rho(y)$ satisfies (A 2) in Theorem 2.1. Hence for mean zero bounded r.v. Y,

$$\begin{aligned} f(\rho)&=-\lim_{n\to\infty}(1/n)\ \log\int_{-\infty}^{\infty}e^{n\psi_\rho(y)}d\mu^n(y)\\ &=\inf_{-\infty<y<\infty}\{I_\mu(y)-\psi_\rho(y)\} \end{aligned} \tag{3.6}$$

by applying Theorem 2.1.

Example 3.1 Let Y have the symmetric Bernoulli distribution $\mu=(1/2)\delta_{-1}+(1/2)\delta_1$. Then the joint distribution in the generalized Curie-Weiss model becomes

$$z_n^{-1}[\cosh\{(x_1+\cdots+x_n)/n\}]^n\prod_{i=1}^{n}d\rho(x_i). \tag{3.7}$$

Since $I_\mu(y)=\begin{cases}(1/2)\{(1+y)\log(1+y)+(1-y)\log(1-y)\}, & \text{if } |y|<1\\ \infty & \text{, otherwise}\end{cases}$

$$,f(\rho)=\inf_{-1<y<1}[(1/2)\{(1+y)\log(1+y)+(1-y)\log(1-y)\}-\psi_\rho(y)]$$

by (3.6).

3.2 Circle Model

Ellis and Rosen's main theorem was inspired by the circle model. We first define the circle model, then compute the *SFE* by applying Theorem 2.1.

Let $\Delta=\{y\in C[0,1]\ ;\ y(0)=y(1)\}$ and let Q, $\{Q_n\}$ be mean zero Gaussian measures on Δ with covariance functions $\sigma(s,t)$, $\{\sigma_n(s,t)\}$ respectively such that $Q_n\Rightarrow Q$. The circle model is defined by

$$z_n^{-1}\exp[(1/2)\sum_{j=1}^{n}\sum_{i=1}^{n}(1/n)\sigma_n(i/n,j/n)x_ix_j+\sum_{i=1}^{n}H_n(i/n)x_i]\prod_{i=1}^{n}d\rho(x_i), \quad (3.8)$$

where ρ is a p.m. on R^1 satisfying (3.1), $\{H_n\}$ is a sequence of functions on Δ such that $H_n\to H$ uniformly for some function H on Δ and

$$z_n=\int_\Delta\exp[(1/2)\sum_{j=1}^{n}\sum_{i=1}^{n}(1/2)\sigma_n(i/n,j/n)x_ix_j+\sum_{i=1}^{n}H_n(i/n)x_i]\prod_{i=1}^{n}d\rho(x_i).$$

Define

$$F_n(y)=-(1/n)\sum_{i=1}^{n}\psi_\rho[y(i/n)+H_n(i/n)].$$

Then

$$z_n=\int_{R^n}\int_\Delta\exp\Big[\sum_{i=1}^{n}\{(1/\sqrt{n})y(i/n)x_i+H_n(i/n)x_i\}\Big]dQ_n(y)\prod_{i=1}^{n}d\rho(x_i)$$

$$=\int_\Delta\int_{R^n}\exp\Big[\sum_{i=1}^{n}\{1/\sqrt{n})y(i/n)x_i+H_n(i/n)x_i\}\Big]\prod_{i=1}^{n}d\rho(x_i)dQ_n(y)$$

$$=\int_\Delta\exp[-nF_n(y)]dQ_n(\sqrt{n}y)$$

$$=\int_\Delta\exp[-nF_n(y)]dQ_n^n(y),$$

where Q_n is the p.m. induced by the sample mean of n i.i.d. r.v.'s whose common probability distribution is Q_n. By Remark 2.1 and a result in Bolthausen (1984, p428), Q_n^n satisfies (A1) of Theorem 2.1. Ellis and Rosen (1982a) showed that $-F_n(y)\leq a+b\|y\|^2$ for any $b>0$ and some $a>0$ for all n and hence F_n satisfies (A2). Finally F_n satisfies (A3) with $F(y)=-\int_0^1\psi_\rho[y(u)+H(u)]\,du$ [Ellis and Rosen (1982a)] and hence all the hypotheses of Theorem 2.1 are satisfied.

Therefore

$$f(\rho)=-\lim_{n\to\infty}(1/n)\log\int_\Delta e^{-nF_n(y)}dQ_n^n(y)$$

$$=\inf_{y\in\Delta}\{F(y)+I_Q(y)\}=\inf_{y\in\Delta}\{F(y)+(1/2)<y,\Lambda^{-1}y>\}$$

and this coincides with the result of Ellis and Rosen (1982 a). The last equality follows from (2.4).

References

(1) Billingsley, P. (1968). *Convergence of Probability Measures*, John Wiley & Sons, New York.

(2) Bolthausen, E. (1984). On the probaility of large deviations in Banach spaces, *Ann. Probability*, Vol. 12, 427∼435.

(3) Donsker, M.D. and Varadhan, S.R.S. (1976). Asymptotic evaluation for certain Markov process expectations for large time Ⅲ, *Comm. Pure Appl. Math.*, Vol. 29. 389∼461.

(4) Ellis, R.S. and Rosen, J.S. (1982 a). Laplace's method for Gaussian integrals with an application to statistical mechanics, *Ann. Probability*, Vol. 10, 47∼66.

(5) Ellis, R.S. and Rosen, J.S. (1982 b). Asymptotic analysis of Gaussian integrals Ⅰ. Isolated minimum points, *Trans. Amer. Math. Soc.*, Vol. 273, 447∼481.

(6) Murray, J.D, (1984). *Asymptotic Analysis*, Springer-Verlag, New York.

(7) Pincus, M. (1968). Gaussian processes and Hammerstein integral equations, *Trans. Amer. Math. Soc.*, Vol. 134, 193∼216.

(8) Prokhorov, Y.V. (1968). An extension of S.N. Bernstein's inequalities to multidimensional distributions, *Theor. Probability Appl.*, Vol. 13, 260∼280.

(9) Schilder, M. (1966). Some asymptotic formulas for Wiener integrals, *Trans. Amer. Math. Soc.*, Vol. 125, 63∼85.

(10) Varadhan, S.R.S. (1966). Asymptotic probability and differential equations, *Comm. Pure Appl. Math.*, Vol. 19, 261∼286.