Resenhas IME-USP 1999, Vol. 4, No. 1, 79 - 91.

Notes on Dominating Points and Large Deviations

Peter Ney

Abstract: In these notes (which supplement my SINAPE lecture) I will survey some early and some current results and applications of the dominating point construction in large deviation theory.

1. Dominating points and large deviation asymptotics.

Let $\Lambda(\alpha) : \mathbb{R}^d \to \mathbb{R}$ be a convex function, and let

$$\Lambda^*(x) = \sup_{\alpha \in \mathbb{R}^d} [< \alpha, x > -\Lambda(\alpha)], \quad x \in \mathbb{R}^d,$$

be its convex conjugate. ($< \cdot, \cdot >$ is inner product). Let $B \subset \mathbb{R}^d$, $B^0 =$ interior of B, $\overline{B}$ =closure of B, $\partial B =$ boundary of B. Let $\mathcal{D}(\Lambda) = \{\alpha \in \mathbb{R}^d : \Lambda(\alpha) < \infty\}$, assume that Λ is differentiable on $\mathcal{D}$, and let $\mathcal{R}(\nabla\Lambda) =$ the range of the map $\nabla\Lambda$.

Definition (N, 1983, 84). *A point v_B is called a dominating point of (Λ, B) if*

(i) $v_B \in \partial B$ and $\inf[\Lambda^*(v) : v \in B] = \Lambda^*(v_B)$.

(ii) $\nabla\Lambda(\alpha) = v_B$ *has a solution* $\alpha_B \in \mathcal{D}(\lambda)$, *and*

(iii) $B \subset \{x :< x, \alpha_B > \geq < v_B, \alpha_B >\} \stackrel{\text{def}}{=} H_{\alpha_B}$.

The situation that concerns us here is when μ is a probability measure on $\mathbb{R}^d$,

$$\hat{\mu}(\alpha) = \int e^{<\alpha, x>} \mu(dx), \ \alpha \in \mathbb{R}^d, \text{ and } \Lambda(\alpha) = \log \hat{\mu}(\alpha).$$

Let $\mathcal{S} =$ the closure of the convex hull of the support of μ. Then we have

Theorem 1.1 ([N2], 1984). *Assume that $\mathcal{D}(\lambda)$ contains a neighborhood of the origin, Λ is essentially smooth, B is convex with $[B \cap \mathcal{S}]^0 \neq \phi$ and $m = \int x\mu(dx) \notin \overline{B}$. Then a unique dominating point v_B for (Λ, B) exists.*

The construction is illustrated in Figure 1a.

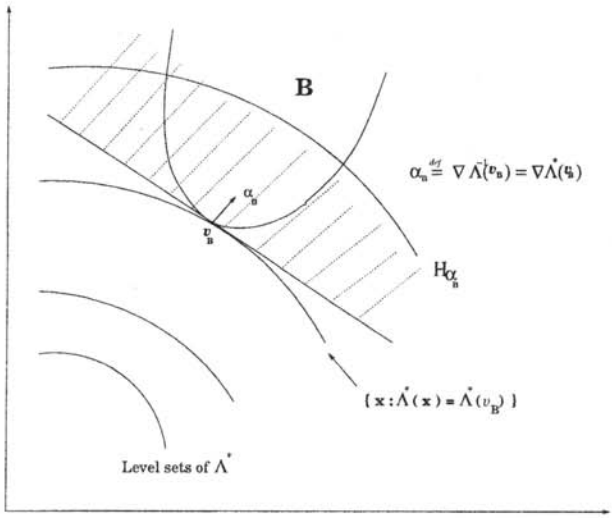

Figure 1a.

Remarks

(i) Note that

$$v_B = \nabla\Lambda(\alpha_B) = \frac{1}{\hat{\mu}(\alpha_B)} \int x e^{<\alpha, x>} \mu(dx).$$

(ii) Also note that for $v \in \mathcal{D}^0(\Lambda^*) = \mathcal{S}^0$

$$\nabla\Lambda^*(v) = \nabla\Lambda^{-1}(v)$$

and thus

$$\alpha_B = \nabla\Lambda^{-1}(v_B) = \nabla\Lambda^*(v_B),$$

as ilustrated in Figure 1. See e.g. Rockafellar [R] (1970) for such properties of convex functions.

(iii) A sufficient condition for essential smoothness is that $\mathcal{D}(\Lambda)$ be an open set.

Extensions of the above construction to infinite dimensional spaces have been treated by Bolthausen [Bo] (1984), Dinwoodie [Di] (1992), Einmahl and Kuelbs [E,K] (1996), and Kuelbs [K] (1998). A related concept, called an exposed point is discussed in Dembo and Zeitouni [D,Z] (1993).

We start with a sequence $X_0, X_1, \ldots$ of independent, identically distributed random variables (i.i.d.r.v.'s), taking values in $\mathbb{R}^d$, with probability law μ. Let Λ

be as above, and define the probability measure

$$\mu_\alpha(A) = e^{-\Lambda(\alpha)} \int e^{<\alpha,s>} \mu(ds),\ A \in \mathcal{R}^d, \alpha \in I\!R^d, \tag{1.3}$$

$\mathcal{R}^d$ = Borel sets. Let $S_n = X_1 + \ldots + X_n$. Take any $\alpha, v \in I\!R^d, B \in \mathcal{R}^d$. Then a straight forward manipulation and change of variable yields

$$I\!P\Big\{\frac{S_n}{n} \in B\Big\} = e^{-n[<\alpha,v>-\Lambda(\alpha)]} J_n(B,v,\alpha) \tag{1.4}$$

where

$$J_n(B,v,\alpha) = \int\limits_{n(B-v)} e^{-<\alpha,s>} \mu_\alpha^{*n}(ds+nv). \tag{1.5}$$

(μ^{*n} = the n fold convolution of μ).

Assume B is as in Theorem 1.1. We are free to choose α and v as we like. We take $v = v_B$ = the dominating point of B, and take

$$\alpha = \alpha_B = \text{ the solution of } \nabla\Lambda(\alpha) = v_B, \tag{1.6}$$

which exists by theorem.

Now we note two important properties. First

$$\int x\mu_{\alpha_B}(dx) = \nabla\Lambda(\alpha_B) = v_B \tag{1.7}$$

and hence

$$\mu_{\alpha\beta}^{*n}(\ \cdot\ + nv_B) \stackrel{\text{def}}{=} \tilde{\mu}_B^{*n}(\ \cdot\ )$$

is the p.m. of a sum of i.i.d.r.'s with mean zero.

Secondly

$$< \alpha_B, v_B > -\Lambda(\alpha_B) = \Lambda^*(v_B). \tag{1.8}$$

We can thus write

$$P\Big\{\frac{S_n}{n} \in B\Big\} = e^{-n\Lambda^*(v_B)} J_n(B), \tag{1.9}$$

where in the notation of (1.5)

$$J_n(B) = J_n(B, v_B, \alpha_B) = \int\limits_{n(B-v_B)} e^{-<\alpha_B,x>} \tilde{\mu}_B^{*n}(dx). \tag{1.10}$$

We call (1.9) the *representation formula* for $P\{\frac{S_n}{n} \in .\}$.

Now from property (iii) in the definition of dominating point

$$< \alpha_B, x > \geq 0 \text{ for } x \in n(B - v_B), \tag{1.11}$$

and hence we immediately see that $J_n(B) \le 1$ (for convex B), and we have the upper bound

$$P\Big\{\frac{S_n}{n} \in B\Big\} \le e^{-n\Lambda^*(B)}, \tag{1.12}$$

(where we write $\Lambda^*(B) = \inf\{\Lambda^*(v) : v \in B\}$).

By a standard covering argument one can then extend (1.12) to get the classical Cramer large deviation upper bound, namely: for all compact F

$$\limsup \frac{1}{n} \log P\Big\{\frac{S_n}{n} \in F\Big\} \le -\Lambda^*(F). \tag{1.13}$$

If also $O \in \mathcal{D}^0(\Lambda)$, then (1.13) also holds for all closed F.

By applying elementary (central limit) estimates of $\tilde{\mu}_B^{*n}$ one gets lower bound estimates on $J_n(B)$, yielding the large deviation lower bound

$$\liminf \frac{1}{n} \log P\Big\{\frac{S_0}{n} \in G\Big\} \ge -\Lambda^*(G) \tag{1.14}$$

for open G. (See e.g. [N1]).

But the real value of the representation formula is that sharper *central* limit estimates (as can be found in Bhatacharya and Rao [B,R] (1976)) can now be used in (1.9) to obtain asymptotic expansions for $P\{\frac{S_n}{n} \in \Gamma\}, \Gamma \in I\!R^d$. These finer results depend on smoothness properties of Γ in the neighborhood of the dominating point v_Γ, provided the later exists. (Dominating points, or even dominating sets, may also exits for non-convex sets). These ideas were developed in [N1] (1983) and more throughly by Ilts [I1] (1995). Up to first order one finds that

$$P\{S_n \in n\Gamma\} = n^\gamma e^{-n\Lambda^*(\Gamma)}[c + O(n^{-\delta})], \tag{1.15}$$

where $-\infty < \gamma \le \frac{d-2}{2}$, and γ depends on the geometry of Γ. ($0 < \delta, c < \infty$ are constants). With the author's permission I reproduce some of the illustrations (for various γ's) from [I1] in Figures 1-6.

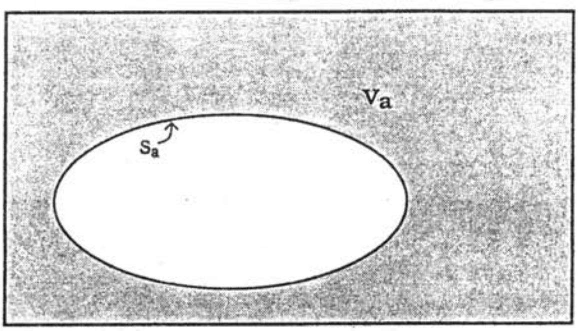

Figure 1. Case $i : \Gamma = V_a$ = Complement of the level set of the rate function. Here the dominating set S_a has dimension $d-1$ and $\gamma = (d-2)/2$.

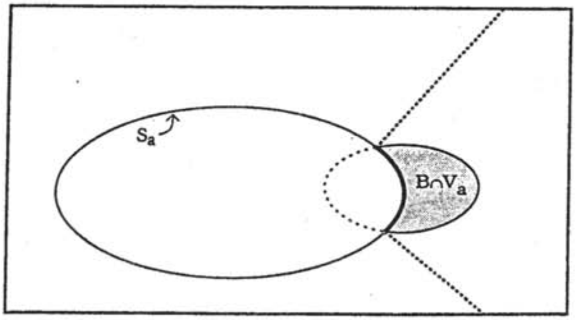

Figure 2. Case ii: Intersection of a convex set B with the complement of the level set of the rate function. Here the dominating set $S_a \cap B$ again has dimension $(d-1)$ so that $\gamma = (d-2)/2$ is maximal (as in case i).

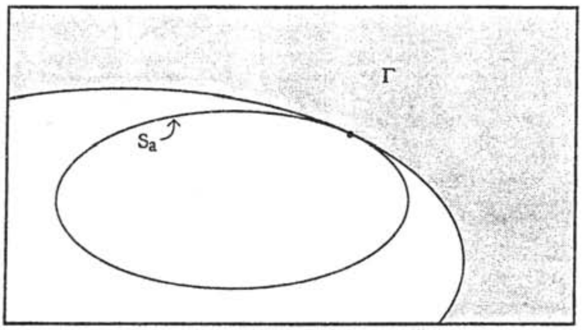

Figure 3. Cases iii and v: Set Γ with unique dominating point $v \in S_a$. Here $(d-2)/2 > \gamma \geq -\frac{1}{2}$

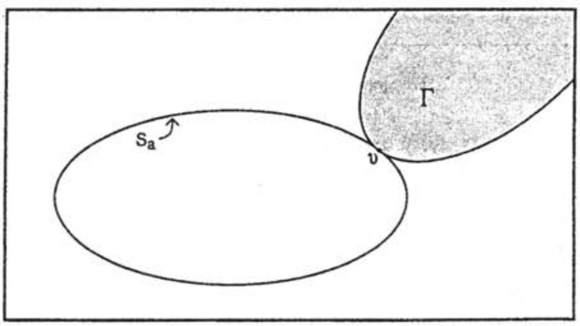

Figure 4. Cases iii and iv: Set Γ with unique dominating point $v \in S_a$. Here $(-\frac{d}{2} < \gamma \leq -\frac{1}{2}$ and $1 > \beta > \frac{1}{2})$.

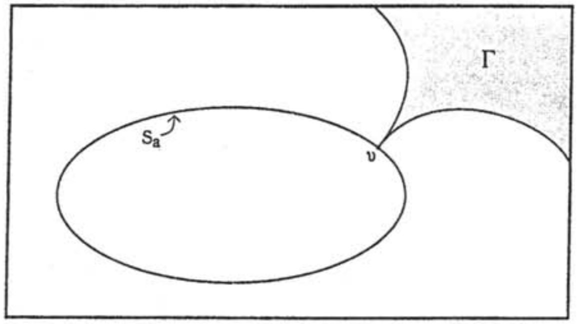

Figure 5. Case iv: Set Γ with unique dominating point $v \in S_a$ having a cusp at v (here $\gamma < -\frac{d}{2}$ and $\beta > 1$)

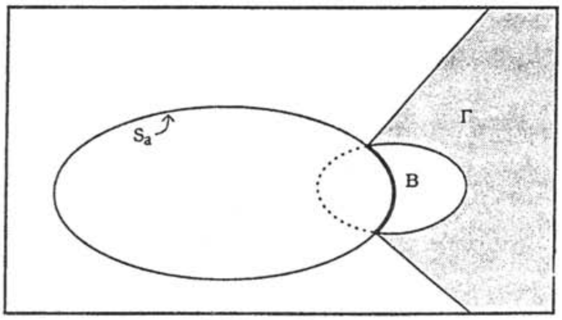

Figure 6. Example of a dominating set of dimension $d-2$ found by removing a convex set $B \subset \mathcal{R}^d$ from the set obtained by following the normals to the portion $B \cap S_a$ of the level surface S_a. (Compare with Figure 2 above). Here the dominating set is $\partial B \cap S_a$ where ∂B is the boundary of B. $\gamma = (d-4)/2$.

A similar situation prevails in the Markov case. Let $X_0, X_1, \ldots$ be an aperiodic, irreducible Markov chain, taking values in a state space S, which we take to be finite for simplicity. Let $\{p(x,y); x, y \in S\}$ denote the transition matrix of $\{X_n\}$, and $\pi(x)$ be its invariant measure. (Under suitable hypotheses on the MC, the following results also extend to general state spaces; see e. g. Iscoe, Ney and Nummelin [I,N,Nu] (1985)).

Let $f : S \to \mathbb{R}^d$ and $S_n = \sum_{i=1}^{n} f(X_i)$. We want analogues of the representation formula (1.9) and asymptotics like (1.15) for $\mathbb{P}_x\{X_n = y, S_n \in n\Gamma\}$.

Define the matrix

$$\{p_\alpha(x,y)\} = \{e^{<\alpha, f(y)>} p(x,y),\ x, y \in S,\ \alpha \in \mathbb{R}^d\}. \tag{1.17}$$

Then $\{p_\alpha\}$ will have a maximal, real (Perron-Frobenius) eigenvalue $\lambda(\alpha)$ with associated right and left eigenvectors $r_\alpha(x), l_\alpha(x), x \in S$.

Let $\pi_\alpha(x) = l_\alpha(x) r_\alpha(x), x \in S$, normalized so that $\sum_{x \in S} \pi_\alpha(x) = 1$.

We will see that $\lambda(\alpha)$ plays a role analogous to the generating function $\hat{\mu}(\alpha)$ in the i.i.d. case. Let $\Lambda = \log \lambda$, $\Lambda^* =$ convex conjugate of Λ. Now define the matrix

$$q_\alpha(x,y) = \frac{e^{<\alpha, f(y)>} p(x,y) r_\alpha(y)}{e^{\Lambda(\alpha)} r_\alpha(x)}, \; x, y \in S, \tag{1.18}$$

and note that this is a stochastic matrix (irreducible, aperiodic) with invariant probability measure π_α. Define the "Markov-additive" kernels

$$q_\alpha^{(n)}(x, y \times ds) = \frac{I\!P_x\{X_n = y, S_n \in ds\} e^{<\alpha, s>} r_\alpha(y)}{e^{n\Lambda(\alpha)} r_\alpha(x)}, \tag{1.19}$$

whose multiplication is governed by

$$q_\alpha^{(n+m)}(x, A \times \Gamma) = \sum_{y \in S} \int_{I\!R^d} q_\alpha^{(n)}(x, y \times ds) q_\alpha^m(y, A \times \Gamma - s), \; \Gamma \subset I\!R^d.$$

Now some algebra yields the analog of (1.4) and (1.5), namely for $\alpha, v \in I\!R^d$

$$I\!P_x\{X_n = y, S_n \in nB\} = e^{-n[<\alpha, v> - \Lambda(\alpha)]} I_n(B, v, \alpha) \tag{1.20}$$

where

$$I_n(B, v, \alpha) = \int_{n(B-v)} e^{-<\alpha, s>} q_\alpha^{(n)}(x, y \times (ds + nv)). \tag{1.21}$$

The existence of a dominating point for (Λ, B) follows by specializing Theorem 5.2 (i) of [I,N,Nu]. Namely we have

Lemma *Let B be convex with $[B \cap \mathcal{R}(\nabla\Lambda)]^0 \neq \phi$ and assume $E_\pi f(X_1) \notin \overline{B}$. Then there exists a unique dominating point v_B of (Λ, B). Define α_B as the solution of $\nabla\Lambda(\alpha) = v_B$, and take $\alpha = \alpha_B, v = v_B$ in (1.20) and (1.21). Abbreviate*

$$I_n(B, v_B, \alpha_B) = I_n(B).$$

Then we have

$$I\!P_x\{X_n = y, S_n \in nB\} = e^{-n\Lambda^*(B)} I_n(B). \tag{1.22}$$

Now $\{q_{\alpha_B}^{(n)}\}$ is "centered" in the sense that

$$\sum_x \int_{I\!R^d} \pi_{\alpha_B}(dx) s q_{\alpha_B}^{(n)}(x, y \times ds) = n v_B$$

and so central limit estimates for centered sums of functions of Markov chains can be applied to (1.22) to yield results like (1.15). Some details of such calculations are in [I,N,Nu] and [I2] (1998). Central limit estimates for Markov chains can be found e. g. in Nagaev [Na1], [Na2] (1957), (1961).

2. Dominating points and conditioned limit laws for independent random variables.

Let $X_0, X_1, \ldots$ be i.i.d.r.v.'s taking values in a measure space $(S, \mathcal{S}), P\{X_0 \in \cdot\,\} = \mu(\,\cdot\,), g : S \to I\!R^d, u : S \to I\!R^m$, measurable functions, $0 < d, m < \infty, G_n = \sum_{i=0}^{n-1} g(X_i), U_n = \sum_{i=0}^{n-1} u(X_i)$. We are interested in the limiting behavior (as $n \to \infty$) of conditioned measures of the form

$$I\!P\Big\{\frac{G_n}{n} \in \cdot\Big|\frac{U_n}{n} \in C\Big\},\ C \subset I\!R^m, \tag{2.1}$$

and of

$$I\!P\{(X_1, \ldots, X_k) \in \cdot\Big|\frac{U_n}{n} \in C\}. \tag{2.2}$$

Example: Let $X_1, X_2, \ldots$ be i.i.d. Bernoulli r. v.'s with $P\{X_i = +1\} = P\{X_i = -1\} = \frac{1}{2}$, and take $0 < a < 1$. Then

$$\lim_{n\to\infty} P\Big\{X_1 = +1\Big|\sum_{i=1}^{n} X_i \geq an\Big\} = \frac{1+a}{2}. \tag{2.3}$$

There is an extensive literature on this subject under the heading "Gibbs Conditioning principle". A basic paper is Csiszar [Cs] (1984). A general reference with extensive bibliography in Dembo and Zeitouni [D,Z1]. We focus here on an approach based on the dominating point construction. We follow an approach of Lehtonen and Nummelin [Le,Nu1] for i.i.d. case, and then discuss some extensions involving Markov chains.

First we need some definitions.

For a sequence og measures $\nu, \nu_1, \nu_2, \ldots$ on $(S, \mathcal{S})$, write $\nu_n \xrightarrow{b} v$ on $I\!R^d$ if $\varphi\nu_n \to \varphi\nu$ for all bounded measurable $\varphi : S \to I\!R^d$. (Abbreviate $\int_S \varphi(s)\nu(ds) = \varphi\nu$). For a sequence of r.v.'s $\{Z_n\} \subset I\!R^d$, write $Z_n \overset{\exp}{\to} z_0$ with respect to $\{I\!P_n\}$ as $n \to \infty$, if given any $\epsilon > 0$ there is an $a > 0$ such that

$$I\!P_n\{||Z_n - z_0|| > \epsilon\} \leq e^{-an}, n = 1, 2, \ldots. \tag{2.4}$$

For a sequence of random measures $\{\tilde{\mu}_n\}$ on $(S, \mathcal{S})$, write $\{\tilde{\mu}_n\} \overset{\exp}{\to} \mu$ w.r. to $\{I\!P_n\}$ if $\varphi\tilde{\mu}_n \overset{\exp}{\to} \varphi\mu$ w.r. to $\{I\!P_n\}$ for all bounded, measurable φ. Later we will use

Lemma 2.1 [Le,Nu1] *If $\tilde{\mu}_n \overset{\exp}{\to} \mu$ w. r. to $\{I\!P_n\}$ then*

$$I\!E_{I\!P_n}(\tilde{\mu}_n) \xrightarrow{b} \mu. \tag{2.5}$$

Let $S(\mu)$ = convex hull of support of μ, and $\Lambda_g(\alpha) = \log \int e^{<\alpha, g(x)>} \mu(dx)$. We can now state some conditioned limit laws.

Lemma 2.2. *Let g be bounded and $B \subset \mathbb{R}^d$ be open and convex. Let $S_g = \overline{S(\mu g^{-1})}$ and assume that $[B \cap S_g]^0 \neq \phi$. Assume that $\mathbb{E}g(X_0) \notin \overline{B}$. Let v_b be the dominating point of (B, Λ_g) and $\alpha_B \in \mathbb{R}^d$ be the solution of $\nabla \Lambda_g(\alpha) = v_b$. Then*

$$\frac{G_n}{n} \overset{\exp}{\longrightarrow} v_B = \int g(x) \mu_{\alpha_B}(dx) \tag{2.6}$$

with respect to the measures

$$\mathbb{P}_n = \mathbb{P}\Big\{ \cdot \Big| \frac{G_n}{n} \in B \Big\}. \tag{2.7}$$

Idea of proof. v_B is the unique point at which Λ^* achieves its minimum. Hence by its convexity, $\Lambda^*(v) > \Lambda^*(v_B) + \delta$ for all v outside an $\epsilon-$ball around v_B and some $\delta > 0$. Now applying the definition of conditional probability, and the fact that

$$\lim \frac{1}{n} \log P\Big\{ \frac{S_n}{n} \in B \Big\} = -\Lambda^*(v_B)$$

(by Cramer's large deviation theorem), it follows that

$$P\Big\{ \|\frac{S_n}{n} - v_B\| > \epsilon \Big| \frac{S_n}{n} \in B \Big\} \leq e^{-an} \tag{2.8}$$

for some $a > 0$.

From Lemma 2.2 one can go to a more general conditioning of the form (2.1), where the conditioning functions u differ from g. The idea is to apply the lemma to $f = (g, u) : S \to \mathbb{R}^{d+m}$, and $S_n = (B_n, U_n)$, where u, g, U_n and G_n are defined above (2.1); with $B = \mathbb{R}^d \times C,\ C \subset \mathbb{R}^m$. Let $\Lambda_u(\beta) = \log \int_S e^{<\beta, u(x)>} \mu(dx),\ \beta \in \mathbb{R}^m$ and $\mu_\beta(A) = e^{-\Lambda_u(\beta)} \int_A e^{<\beta, u(x)>} \mu(dx)$. Then one obtains

Lemma 2.3 [Le,Nu2] *Let u and g be bounded, $C \subset \mathbb{R}^m$ be open and convex with $[C \cap S_u]^0 \neq \emptyset,\ Eu(X_0) \notin \overline{C}$. Let v_C be the dominating point of (Λ_u, C), and $\beta_C \in \mathbb{R}^m$ be the solution of $\nabla \Lambda_u(\beta) = v_C$. Then*

$$\frac{G_n}{n} \overset{\exp}{\to} \int g(x) \mu_{u, \beta_C}(dx) \tag{2.9}$$

with respect to

$$\mathbb{P}_n = \mathbb{P}\{ \cdot \Big| \frac{U_n}{n} \in C \}. \tag{2.10}$$

Now let $\tilde{P}_n$ denote the empirical measure of $\{X_n\}$:

$$\tilde{P}_n(A) = \frac{1}{n}\sum_{i=0}^{n-1} \mathbf{1}_A(X_i),\ A \in \mathcal{S},$$

and note that

$$\frac{G_n}{n} = \int_S g(x)\tilde{P}_n(dx) = g\tilde{P}_n. \tag{2.11}$$

Then if u and C are as in Lemma 2.3, one can conclude that

$$g\tilde{P}_n \overset{\exp}{\to} g\mu_{\beta_C} \ \text{ w. r. to } \mathbb{P}_n \text{ (in (2.10))} \tag{2.12}$$

for all bounded measurable g. Hence by Lemma 2.1

$$\mathbb{E}_{\mathbb{P}_n}(\tilde{P}_n) \overset{b}{\to} \mu_{\beta_C}, \tag{2.13}$$

or writing this out

$$\frac{1}{n}\sum_{i=0}^{n-1} P\Big\{X_i \in \ \cdot \Big| \frac{U_n}{n} \in C\Big\} \overset{b}{\to} \mu_{\beta_C}(\cdot). \tag{2.14}$$

But by simmetry, the terms in the above summand are all equal. Hence we have

Theorem 2.1 [Le,Nu2] *Let u, C and β_C be as Lemma 2.3. Then*

$$\mathbb{P}\{X_1 \in \cdot \Big| \frac{U_n}{n} \in C\} \overset{b}{\to} \mu_{\beta_C} \ \text{on } \mathbb{R}^d. \tag{2.15}$$

One can carry the above argument further to conclude

Corollary 2.1. *Under the above hypotheses on u, C and β_C*

$$\mathbb{P}\{(X_1, \ldots, X_K) \in \cdot \ \Big| \frac{U_n}{n} \in C\} \overset{b}{\to} \mu^k_{\beta_C} \qquad (2.,16)$$

where μ^k_β is the k-fold product measure of μ_β. Thus in the sense of (2.16), $X_1, X_2, \ldots$ are "asymptotically independent".

Csiszar [Cs] proved a stronger convergence for (2.7) under somewhat different conditioning. The case when $k = k_n \to \infty$ has been studied in Dembo and Zeitoni [D,Z2] (1996) and Dembo and Kuelbs [DK] (1997).

3. Conditioned limits for Markov chains.

We briefly consider the case when $X_0, X_1, \ldots$ is a (time homogeneous) Markov chain. To avoid technicalities we describe the simplest situation namely when the state space S is finite, and $\{X_n\}$ is irreducible and aperiodic. The general state space case has been studied in Schroeder [Sc] and Meda and Ney M,N1] and [M,N2].

Let $P = \{p(x,y), x, y \in S\}$ be the transition matrix of $\{X_n\}$. Let $g : S \times S \to I\!R^d$, $u : S \times S \to I\!R^m$, $G_n = \sum_{i=0}^{n-1} g(X_i, X_{i+1})$, $U_n = \sum_{i=0}^{n-1} u(X_i, X_{i+1})$. (We use the same symbols u and g as before, but they are now functions on $S \times S$). Consider the matrix

$$P_g(\alpha) = \{p_{g,\alpha}(x,y)\} = \{e^{<\alpha, g(x,y)>} p(x,y);\ x, y \in S\},\ \alpha \in I\!R^d. \tag{3.1}$$

Let $\lambda_g(\alpha) =$ the maximal, real (Perron-Frobenius) eigenvalue of $P_g(\alpha)$, with associated right (left) eigenvector $r_{g,\alpha}(x)(l_{g,\alpha}(x))$. Let $\Lambda_g = \log \lambda_g$. Let $\pi_{g,\alpha}(x) = l_{g,\alpha}(x)\ r_{g,\alpha}(x)$, normalized so that $\sum_{k \in S} \pi_{g,\alpha}(x) = 1$, and let $\{q_{g,\alpha}(x,y)\} = \{e^{-\Lambda_g(\alpha)} (\frac{r_{g,\alpha}(y)}{r_{g,\alpha}(x)}) p_{g,\alpha}(x,y)\}$. This is a stochastic matrix. Let $P_u(\beta)$, $\beta \in I\!R^m$, and corresponding quantities be defined similarly for u. (When S is finite and $\{X_n\}$ is irreducible these quantities always exist, but for general state spaces further hypotheses are needed). Let $\Lambda^* : I\!R^d \to I\!R =$ the convex conjugate of Λ and $\mathcal{D}(\Lambda^*) = \{x \in I\!R^d : \Lambda^*(x) < \infty\}$. Then we have the following analogues of Lemmas 2.2 and 2.3.

Lemma 3.1 *Let $B \subset I\!R^d$ be open and convex, with $[B \cap \mathcal{D}(\Lambda_g^*)]^0 \neq \emptyset$. Let $v_B =$ the dominating point of (Λ_g, B) and $\alpha_B =$ the solution of $\nabla \Lambda_g(\alpha) = v_B$. Assume $E_{\pi_{g,\alpha_B}} g(X_0, X_1) \notin \overline{B}$. Then*

$$\frac{G_n}{n} \overset{\exp}{\to} v_B = \sum_{x,y \in S} g(x,y) \pi_{\alpha_B}(x) q_{g,\alpha_B}(x,y) \tag{3.2}$$

with respect to

$$I\!P_n = I\!P_x\{\ \cdot\ \Big| \frac{G_n}{n} \in B\}. \tag{3.3}$$

The proof again follows from properties of dominating points. The analog of Lemma 2.3 becomes

Lemma 3.2 *Let $C \subset I\!R^m$ be open and convex with $[C \cap \mathcal{D}(\Lambda_u^*)]^0 \neq \emptyset$. Let v_C be the dominating point of (Λ_u, C) and β_C be the solution of $\nabla \Lambda_u(\beta) = v_C$. Assume $I\!E_{\pi_{u,\beta_C}} u(X_0, X_1) \notin \overline{C}$. Then*

$$\frac{G_n}{n} \overset{\exp}{\to} \sum_{x,y \in S} g(x,y) \pi_{u,\beta_C}(x) q_{u,\beta_C}(x,y). \tag{3.4}$$

with respect to

$$\mathbb{P}_n = \mathbb{P}\{ \,\cdot\, \Big| \frac{U_n}{n} \in C\}.$$

From this we can argue as before that

Lemma 3.3 *Under the hypothesis of Lemma 3.2*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} P\Big\{X_i = x, X_{i+1} = y \Big| \frac{U_n}{n} \in C\Big\} = \pi_{u,\beta_C}(x) q_{u,\beta_C}(x,y). \tag{3.5}$$

This is the analog of (2.14). However, an analog of the symmetry argument used to get to (2.15) does not work here. In the finite state space case Csiszar, Cover and Choi [Cs,Co,Ch] (1987) use a careful counting argument to obtain a version of (2.15), but for general state spaces the question seems open. One can also extend (3.5) to

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbb{P}\{X_{i+1} = x_1, \ldots, X_{i+k} = x_k \Big| \frac{U_n}{n} \in C\} \tag{3.6}$$

$$= \pi_{u,\beta_C}(x_1) \prod_{i=1}^{k-1} q_{u,\beta_C}(x_i, x_{i+1}),$$

and thus in this Cesaro convergence sense, $\{X_i, i = 1, 2, \ldots\}$ conditioned on $\frac{U_n}{n} \in C$ is asymptotically Markov with transition matrix $\{q_{u,\beta_C}\}$. (See [M,N2] for details of the above calculations and extensions to general state spaces).

References

[B,R] Bhattacharya, R. N. and Rao, R. Ranga (1976). *Normal Approximation and Asymptotic Expansions.* John Wiley & Sons.

[Bo] Bolthausen, E. (1986). Laplace approximations for sums of independent random vectors Part I. *Prob. Th. Rel. Fields* (72) 305-318.

[Cs] Csiszar, I. (1984). Sanov property, generalized I-projection and a conditional limit theorem. *Ann. Prob.* (12) 768-793.

[D,K] Dembo, A. and Kuelbs, J. (1997). A Gibbs conditioning principle for certain infinite dimensional statistics. *Univ. of Wisconsin technical report.*

[D,Z1] Dembo, A. and Zeitouni, O. (1993). *Large Deviation Techniques.* Jones and Bartlett, Boston.

[D,Z2] Dembo, A. and Zeitouni, O. (1996). Refinements of the Gibbs conditioning principle. *Prob. Th. Relat. Fields* (104) 1-14.

[Di] Dinwoodie, I. H. (1992). Measure dominantes et théorème de Sanov. *Ann. Inst. H. Poincaré Prob. Stat.* (28) 365-373.

[E,K] Einmahl, U. and Kuelbs, J. (1996). Dominating points and large deviations for random vectors. *Pr4ob. Th. Relat. Fields* (105) 529-543.

[I1] Iltis, M. (1995). Sharp asymptotics of large deviations in $\mathbb{R}^d$. *J. of Th. Prob.* (8) 501-522.

[I2] Iltis, M. (1998). Asymptotics of large deviations for Markov-additive chains in $\mathbb{R}^d$. *Univ. of Wisconsin technical report.* (Also 1991 Ph. D. thesis).

[I,N,Nu] Iscoe, I., Ney, P. and Nummelin, E. (1985). Large deviations of uniformly recurrent Markov additive processes. *Adv. in Appl. Math.* (6) 373-412.

[K] Kuelbs, J. (1998). Large deviation probabilities and dominating points for open convex sets: non-logarithmic behavior. *Univ. of Wisconsin technical report.*

[Le,Nu1] Lehtonen, T. and Nummelin, E. (1988). On the convergence of empirical distributions under partial observations. *Ann. Acad. Sc. Fennicae. Ser. A. I.* (13) 219-223.

[Le,Nu2] Lehtonen, T. and Nummelin, E. (1990). Level I theory of large deviations in the ideal gas. *Int'l J. of Theoret. Phys.* (29) 621-635.

[M,N1] Meda, A. and Ney, P. (1998). A conditioned law of large numbers for Markov-additive ch ains. *Studia Sc. Math. Hungarica* to appear.

[M,N2] Meda, A. and Ney, P. (1998). The Gibbs conditioning principle for Markov chains. *Univ. of Wisconsin technical report.*

[Na1,Na2] Nagaev, S. V. (1957, 1961). Some limit theorems for stationary Markov chains. *Th. of Prob. and Appl's.* (2) 378-406; (6) 62-81.

[N1] Ney, P. (1983). Dominating points and the asymptotics of large deviations on $\mathbb{R}^d$. *Ann. Prob.*(11) 158-167.

[N2] Ney, P. (1984). Convexity and Large Deviations. *Annals of Prob.* (12) 903-906.

[Sc] Schroeder, C. (1993). *I*-projection and limit theorems for discrete parameter Markov chains. *Ann. Prob.* (21) 721-758.

Peter Ney
University of Wisconsin, Madison
ney@math.wisc.edu
U.S.A.