Sanov's Theorem

Let E be a Polish space, and define $L_n : E^n \longrightarrow \mathbf{M}_1(E)$ to be the empirical measure given by $L_n(x) = \frac{1}{n}\sum_{m=1}^n \delta_{x_m}$ for $x = (x_1, \dots, x_n) \in E^n$. Given a $\mu \in \mathbf{M}_1(E)$, denote by $\tilde{\mu}_n$ the distribution of L_n under μ^n.

LEMMA 1. *For each $M \in (0,\infty)$ there is a compact set $\mathcal{K}_M \subseteq \mathbf{M}_1(E)$ such that*

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \tilde{\mu}_n(E \setminus \mathcal{K}_M) \le -M.$$

PROOF: Choose a non-decreasing sequence $\{K_j : j \ge 1\}$ of compact subsets of E so that $\mu(E\setminus K_j) \le e^{-2j}$, and set $E' = \sum_{j=1}^\infty K_j$ and $V = \sum_{j=1}^\infty \mathbf{1}_{E\setminus K_j}$. Then $V \le \ell$ on K_ℓ, and so

$$\mathbb{E}^\mu\big[e^V\big] = \int_{E'} e^V\,d\mu = \lim_{\ell\to\infty} \int_{K_\ell} e^V\,d\mu = \sum_{\ell=1}^\infty \int_{K_{\ell+1}\setminus K_\ell} e^V\,d\mu \le \frac{e}{e-1}.$$

At the same time, $V \ge \ell$ off K_ℓ, and so $\langle V,\nu\rangle \le R \implies \nu(E\setminus K_\ell) \le \frac{R}{\ell}$. In addition, because V is l.s.c., $\mathcal{K}_R = \{\nu : \langle V,\nu\rangle \le R\}$ is closed. Hence, for each $R > 0$, $\mathcal{K}_R$ is compact in $\mathbf{M}_1(E)$. Finally,

$$\tilde{\mu}_n\big(\mathcal{K}_R \complement\big) \le e^{-nR} \int e^{n\langle V,\nu\rangle}\,\tilde{\mu}_n(d\nu) = e^{-nR}\left(\int e^V\,d\mu\right)^n = e^{-n(R+A)},$$

where $A = \log \int e^V\,d\mu$. $\square$

LEMMA 2. *Set $\tilde{\Lambda}_\mu(\varphi) = \log \mathbb{E}^\mu\big[e^{\langle \varphi,\nu\rangle}\big]$ for $\varphi \in C_{\mathrm{b}}(E;\mathbb{R})$ and*

$$\tilde{\Lambda}^*_\mu(\nu) = \sup\big\{\langle\varphi,\nu\rangle - \tilde{\Lambda}_\mu(\varphi) : \varphi \in C_{\mathrm{b}}(E;\mathbb{R})\big\}.$$

Then $\tilde{\Lambda}^*_\mu$ is a rate function and

$$\varlimsup_{n\to\infty} \frac{1}{n}\log\tilde{\mu}_n(A) \le -\inf_{\overline{A}} \tilde{\Lambda}^* \quad \textit{for all } A \in \mathcal{B}_{\mathbf{M}_1(E)}.$$

PROOF: Let $B(\nu,r)$ denote the Lévy ball of radius r around $\nu \in \mathbf{M}_1(E)$. Because of Lemma 1, it suffices to show that

$$\lim_{r\searrow 0}\varlimsup_{n\to\infty}\frac{1}{n}\log\tilde{\mu}_n\big(B(\nu,r)\big) \le -\tilde{\Lambda}^*(\nu) \text{ for each } \nu \in \mathbf{M}_1(E).$$

But, for each φ, there exist $\epsilon_\varphi(r) \searrow 0$ such that

$$\tilde{\mu}_n\big(B(\nu,r)\big) \le e^{-n\langle\varphi,\nu\rangle + n\epsilon_\varphi(r)}\mathbb{E}^{\tilde{\mu}_n}\left[e^{n\langle\varphi,\nu\rangle}\right] = \exp\Big(n\big(-\langle\varphi,\nu\rangle + \tilde{\Lambda}_\mu(\varphi) + \epsilon_\varphi(r)\big)\Big),$$

and so

$$\varlimsup_{n\to\infty}\frac{1}{n}\log\tilde{\mu}_n\big(B(\nu,r)\big) \le -\langle\varphi,\nu\rangle + \tilde{\Lambda}(\varphi) + \epsilon_\varphi(r).$$

Now take the limit as $r \searrow 0$ and then the supremum over $\varphi \in C_{\mathrm{b}}(E;\mathbb{R})$. $\square$

LEMMA 3. *For $\nu \in \mathbf{M}_1(\Sigma)$, define*

$$H(\nu|\mu) = \begin{cases} \int_\Sigma f\log f\,d\mu & \text{if } \nu \ll \mu \text{ and } f = \frac{d\nu}{d\mu} \\ \infty & \text{otherwise.}\end{cases}$$

Then $\tilde{\Lambda}^*_\mu(\nu) = H(\nu|\mu)$. In particular, $\nu \rightsquigarrow H(\nu|\mu)$ is l.s.c. and convex. In fact, if $H(\nu_1|\mu)\vee H(\nu_2|\mu) < \infty$ and $\theta \in (0,1)$, then $H\big((1-\theta)\nu_1 + \theta\nu_2\big) < (1-\theta)H(\nu_1|\mu) + \theta H(\nu_2|\mu)$.

PROOF: The final assertion follows immediately from the strict convexity of $x \in [0,\infty) \longmapsto x\log x \in \mathbb{R}$.

If $\nu \ll \mu$ and $\nu_\theta \equiv \theta\mu + (1-\theta)\nu$ for $\theta \in [0,1]$, then $H(\nu|\mu) = \lim_{\theta\searrow 0} H(\nu_\theta|\mu)$. To see this, set $f = \frac{d\nu}{d\mu}$ and $f_\theta = \theta + (1-\theta)f$. Since $x \in [0,\infty) \longmapsto x\log x$ is convex, Jensen's inequality says that

$$H(\nu_\theta|\mu) = \int f_\theta \log f_\theta \, d\mu \le (1-\theta)\int f\log f\, d\mu = (1-\theta)H(\nu|\mu).$$

At the same time, since $x \in [0,\infty) \longmapsto \log x$ is non-decreasing and concave, $\log f_\theta$ dominates both $\log\theta$ and $(1-\theta)\log f$; and therefore

$$H(\nu_\theta|\mu) = \theta\int \log f_\theta\, d\mu + (1-\theta)\int f\log f_\theta\, d\mu \ge \theta\log\theta + (1-\theta)^2 H(\nu|\mu).$$

After combining these two, one clearly gets the asserted convergence.

I next show that if $\nu \ll \mu$, then $\tilde\Lambda^*_\mu(\nu) \le H(\nu|\mu)$. In view of the preceding and the obvious fact that $\nu \in \mathbf{M}_1(\Sigma) \longmapsto \tilde\Lambda^*_\mu(\nu)$ is lower semi-continuous, I may and will assume that $f = \frac{d\nu}{d\mu} \ge \theta$ for some $\theta \in (0,1)$. In particular, by Jensen's inequality,

$$\exp\left[\int \varphi\, d\nu - H(\nu|\mu)\right] = \exp\left[\int (\varphi - \log f)\, d\nu\right] \le \int \frac{\exp[\varphi]}{f}\, d\nu = \int \exp[\varphi]\, d\mu;$$

from which it is clear that $\tilde\Lambda^*_\mu(\nu) \le H(\nu|\mu)$.

As a consequence of the preceding, all that remains is to show that if $\tilde\Lambda^*_\mu(\nu) < \infty$, then $d\nu = f\,d\mu$ and

$$(*) \qquad \tilde\Lambda^*_\mu(\nu) \ge \int f\log f\, d\mu.$$

Given ν with $\tilde\Lambda^*_\mu(\nu) < \infty$, one has

$$(4) \qquad \int \varphi\, d\nu - \log\left(\int \exp[\varphi]\, d\mu\right) \le \tilde\Lambda^*_\mu(\nu) < \infty$$

for every bounded continuous φ. Since the class of φ's for which (4) holds is closed under bounded point-wise convergence, (4) continues to be true for every bounded $\mathcal{B}_E$-measurable φ. In particular, one can now show that $\nu \ll \mu$. Indeed, suppose that $\Gamma \in \mathcal{B}_E$ with $\mu(\Gamma) = 0$. Then, by (4) with $\varphi = r\,\mathbf{1}_\Gamma$, $r\nu(\Gamma) \le \tilde\Lambda^*_\mu(\nu)$, $r > 0$; and therefore $\nu(\Gamma) = 0$. Knowing that $\nu \ll \mu$, set $f = \frac{d\nu}{d\mu}$. If f is uniformly positive and uniformly bounded, then (*) is an immediate consequence of (4) with $\varphi = \log f$. If f is uniformly positive but not necessarily uniformly bounded, set $f_n = f \wedge n$ and use (4) together with Fatou's Lemma to justify

$$\int f\log f\, d\mu = \int \log f\, d\nu \le \varliminf_{n\to\infty} \int \log f_n\, d\nu \le \tilde\Lambda^*_\mu(\nu) + \varlimsup_{n\to\infty} \log\left(\int f\wedge n\, d\mu\right) = \tilde\Lambda^*_\mu(\nu).$$

Finally, to treat the general case, define ν_θ and $f_\theta = \theta + (1-\theta)f$ for $\theta \in [0,1]$ as in the first paragraph of this proof. By the preceding, $\int f_\theta \log f_\theta\, d\mu \le \tilde\Lambda^*_\mu(\nu_\theta)$ as long as $\theta \in (0,1)$. Moreover, since $\theta \in [0,1] \longmapsto \tilde\Lambda^*_\mu(\nu_\theta)$ is bounded, lower semi-continuous, and convex on $[0,1]$, it is continuous there. In conjunction with the result obtained in the first paragraph, this now completes the proof. $\square$

As a consequence of Lemma 3 and (4), one knows that

$$(5) \qquad \langle \varphi, \nu\rangle \le H(\nu|\mu) + \log E^\mu\left[e^\varphi\right]$$

for any $\mathcal{B}_E$-measurable $\varphi : E \longrightarrow \mathbb{R}$ which is bounded below.

Theorem 6 (**Sanov**). *The map $\nu \in \mathbf{M}_1(E) \longmapsto H(\nu|\mu) \in [0,\infty]$ is a good rate function and*

$$-\inf_{\nu\in A^o} H(\nu|\mu) \le \varliminf_{n\to\infty} \frac{1}{n}\log\tilde{\mu}_n(A) \le \varlimsup_{n\to\infty} \frac{1}{n}\log\tilde{\mu}_n(A) \le -\inf_{\nu\in\overline{A}} H(\nu|\mu)$$

for all $A \in \mathcal{B}_{\mathbf{M}_1(E)}$.

Proof: In view of Lemmas 1, 2, and 3, it suffices to prove that if G is open in $\mathbf{M}_1(E)$ and $\nu \in G$ with $H(\nu|\mu) < \infty$ then $\varliminf_{n\to\infty} \frac{1}{n}\log\tilde{\mu}_n(G) \ge H(\nu|\mu)$. To this end, suppose that $\nu \in G$ with $H(\nu|\mu) < \infty$, and let $f = \frac{d\nu}{d\mu}$. For $n \ge 1$, set $F_n(x) = \prod_{m=1}^n f(x_m)$ for $x \in E^n$ and $A_n = \{x \in E^n : L_n(x) \in G \text{ and } F_n(x) > 0\}$. Then, because $t\log t \ge -\frac{1}{e}$, Jensen's inequality implies

$$\begin{aligned}
\log\bigl(\tilde{\mu}_n(G)\bigr) &\ge \log\left(\int_{A_n} \frac{1}{F_n(\sigma)}\,\nu^n(d\sigma)\right)\\
&\ge \log\bigl(\nu^n(A_n)\bigr) - \frac{1}{\nu^n(A_n)}\int_{A_n} \log\bigl(F_n(\sigma)\bigr)\,\nu^n(d\sigma)\\
&\ge \log\bigl(\nu^n(A_n)\bigr) - \frac{1}{e\nu^n(A_n)} - \frac{1}{\nu^n(A_n)}\int_{\Sigma^n} \log\bigl(F_n(\sigma)\bigr)\,\nu^n(d\sigma)\\
&= \log\bigl(\nu^n(A_n)\bigr) - \frac{1}{e\nu^n(A_n)} - n\frac{H(\nu|\mu)}{\nu^n(A_n)}
\end{aligned}$$

as long as $\nu^n(A_n) > 0$. Finally, by the Strong Law of Large Numbers, $\nu^n(A_n) \longrightarrow 1$ as $n \to \infty$. □

Cramér vs. Sanov

Let E be a separable, real Banach space, and assume that $\mu \in \mathbf{M}_1(E)$ satisfies $\mathbb{E}^\mu\bigl[e^{\alpha\|x\|_E}\bigr] < \infty$ for all $\alpha \ge 0$. Next, set $\Lambda_\mu(x^*) = \log\mathbb{E}^\mu\bigl[e^{\langle x,x^*\rangle}\bigr]$ for $x^* \in E^*$, and define

$$\Lambda_\mu^*(x) = \sup\bigl\{\langle x,x^*\rangle - \Lambda_\mu(x^*) : x^* \in E^*\bigr\} \quad \text{for } x \in E.$$

Finally, let $\mu_n \in \mathbf{M}_1(E)$ denote the distribution of

$$x \in E^n \longmapsto \overline{S}_n \equiv \frac{1}{n}\sum_{m=1}^n x_m \in E \quad \text{under } \mu^n.$$

The goal here is to show that Λ_μ^* is good, that $\{\mu_n : n \ge 1\}$ satisfies the full large deviations principle with respect to Λ_μ^*, and that

$$\Lambda_\mu^*(x) = J_\mu(x) \equiv \inf\left\{H(\nu|\mu) : \int \|x\|\,d\nu < \infty \text{ and } \int y\,\nu(dy) = x\right\}. \tag{7}$$

Let $\mathcal{I}$ be the set $\nu \in \mathbf{M}_1(E)$ such that $\mathbb{E}^\nu[\|x\|] < \infty$. Then $\mathcal{I} \in \mathfrak{F}_\sigma\bigl(\mathbf{M}_1(E)\bigr)$ and $\tilde{\mu}_n(\mathcal{I}) = 1$ for all $n \ge 1$.

Lemma 8. *There is a l.s.c. $V : E \longrightarrow [0,\infty)$ such that $V(x) \ge \|x\|_E$, $\lim_{\|x\|_E\to\infty} \frac{V(x)}{\|x\|_E} = \infty$, and $\mathbb{E}^\mu\bigl[e^V\bigr] < \infty$. In particular, if $0 \le R_\ell \nearrow \infty$ is chosen so that $V(x) \ge \ell\|x\|_E$ for $\|x\|_E \ge R_\ell$, then for each $M > 0$ there is an $C_M \in (0,\infty)$ such that $\varlimsup_{n\to\infty} \frac{1}{n}\log\tilde{\mu}_n\bigl(\mathbf{M}_1(E)\setminus\mathcal{F}_M\bigr) \le -M$ where*

$$\mathcal{F}_M \equiv \left\{\nu : \int_{\|x\|_E<R_\ell} \|x\|_E\,d\nu \le \frac{C_M}{\ell} \text{ for all } \ell \in \mathbb{Z}^+\right\}.$$

PROOF: For each $\ell \in \mathbb{Z}^+$, choose $R_\ell \in (0,\infty)$ so that $\int_{\|x\|_E \ge R_\ell} e^{\ell \|x\|_E}\,\mu(dx) \le 2^{-\ell}$. Without loss in generality, assume that $R_\ell \nearrow \infty$. Set

$$V(x) = \|x\|_E \left(1 + \sum_{\ell=1}^{\infty} \mathbf{1}_{(R_\ell,\infty)}(\|x\|_E)\right).$$

Then V is l.s.c., $V(x) \ge \|x\|_E$ for all $x \in E$, $V(x) \le \ell\|x\|_E$ when $\|x\|_E \le R_\ell$, and $V(x) \ge \ell \|x\|_E$ when $\|x\|_E \ge R_\ell$. Hence,

$$\mathbb{E}^\mu\big[e^V\big] \le e + \sum_{\ell=1}^{\infty} \int_{R_\ell < \|x\|_E \le R_{\ell+1}} e^{(\ell+1)\|x\|_E}\,\mu(dx) \le 2e,$$

and so

$$\tilde\mu_n\big(\langle V,\nu\rangle \ge C\big) \le e^{-nC}\mathbb{E}^{\tilde\mu_n}\big[e^{n\langle V,\nu\rangle}\big] = \exp\big(-nC + n\log \mathbb{E}^\mu\big[e^V\big]\big) \le e^{-n(C-\log 2e)}.$$

Therefore, if $C_M = M + \log 2e$, then $\overline{\lim}_{n\to\infty} \frac{1}{n}\log\tilde\mu_n\big(\langle V,\nu\rangle\big) \le -M$. Since $\langle V,\nu\rangle \le C_M \implies \nu \in \mathcal{F}_M$, this completes the proof. $\square$

Continuing with the notation in Lemma 8, note that each of the sets $\mathcal{F}_M$ is closed in $\mathbf{M}_1(E)$ and contained in $\mathcal{I}$. Next, define $\Psi : \mathcal{I} \longrightarrow E$ so that $\Psi(\nu) = \int x\,\nu(dx)$. Then $\Psi \restriction \mathcal{F}_M$ is bounded and continuous for each M. Indeed, the boundedness is obvious. To prove continuity, choose $\eta \in C\big(\mathbb{R};[0,1]\big)$ so that $\eta(t) = 1$ if $t \le 0$ and $\eta(t) = 0$ if $t \ge 1$. Then, for each $\ell \in \mathbb{Z}^+$, the function $\Psi_\ell : \mathbf{M}_1(E) \longrightarrow E$ given by $\Psi_\ell(\nu) = \int \eta(\|x\|_E - R_\ell)x\,\nu(dx)$ is bounded and continuous. Furthermore, as $\ell \to \infty$, $\Psi_\ell \longrightarrow \Psi$ uniformly on $\mathcal{F}_M$.

LEMMA 9. *For each $M \in (0,\infty)$ there is a $K_M \subset\subset E$ such that $\overline{\lim}_{n\to\infty} \frac{1}{n}\log\mu_n(E\setminus K_M) \le -M$.*

PROOF: Choose $\mathcal{K}_M \subset\subset \mathbf{M}_1(E)$ as in Lemma 1, and set $K_M = \Psi\big(\mathcal{K}_M \cap \mathcal{F}_M\big)$. Then, because $\mathcal{K}_M \cap \mathcal{F}_M \subset\subset \mathbf{M}_1(E)$ and $\Psi \restriction \mathcal{F}_M$ is continuous, $K_M \subset\subset E$. In addition, by Lemmas 1 and 8,

$$\overline{\lim_{n\to\infty}} \frac{1}{n}\log\mu_n(E\setminus K_M) = \overline{\lim_{n\to\infty}} \frac{1}{n}\log\tilde\mu_n\big(\mathbf{M}_1(E)\setminus(\mathcal{K}_M\cap\mathcal{F}_M)\big) \le -M. \quad \square$$

LEMMA 10. *The function J_μ is a good rate function which is convex. Moreover, if $x \in E$ and $J_\mu(x) < \infty$, there exists a unique $\nu \in \mathcal{I}$ such that $\Psi(\nu) = x$ and $H(\nu|\mu) = J_\mu(x)$.*

PROOF: Suppose that $J_\mu(x) < \infty$. Then I can find $\{\nu_k : k \ge 1\} \subseteq \mathcal{I}$ such that $\Psi(\nu_k) = x$ and $H(\nu_k|\mu) \le J_\mu(x) + \frac{1}{k}$. In particular, $\{\nu_k : k \ge 1\}$ is relatively compact and, by (5) with φ equal to the function V in Lemma 8, $\sup_{k\ge1}\langle V,\nu_k\rangle \le J_\mu(x) + 1 + \log\mathbb{E}^\mu\big[e^V\big]$, which means that $\{\nu_k : k \ge 1\} \subseteq \mathcal{F}_M$ for some $M < \infty$. Thus, without loss in generality, I may assume that $\nu_k \implies \nu$ for some $\nu \in \mathcal{F}_M$. Since $\Psi \restriction \mathcal{F}_M$ is continuous and $H(\,\cdot\,|\mu)$ is l.s.c., this implies that $\Psi(\nu) = x$ and that $H(\nu|\mu) \le \Lambda^*_\mu(x)$, which means that $H(\nu|\mu) = \Lambda^*_\mu(x)$. Further, if ν_1,ν_2 were distinct elements of $\mathcal{I}$ satisfying $\Psi(\nu_1) = x = \Psi(\nu_2)$ and $H(\nu_1|\mu) = J_\mu(x) = H(\nu_2|\mu)$, then one would have that $\frac{\nu_1+\nu_2}{2} \in \mathcal{I}$, $\Psi\left(\frac{\nu_1+\nu_2}{2}\right) = x$, and $H\left(\frac{\nu_1+\nu_2}{2}\right) < J_\mu(x)$, which is impossible.

To prove that $\{J_\mu \le L\} \subset\subset E$, suppose $\{x_k : k \ge 1\} \subseteq E$ with $J_\mu(x_k) \le L$. For each k, choose $\nu_k \in \mathcal{I}$ so that $\Psi(\nu_k) = x_k$ and $H(\nu_k|\mu) = J_\mu(x_k)$. Then, because $H(\,\cdot\,|\mu)$ is good, $\{\nu_k : k \ge 1\}$ is relatively compact. In addition, just as above, $\{\nu_k : k \ge 1\} \subseteq \mathcal{F}_M$ for some $M < \infty$. Finally, choose a subsequence $\{\nu_{k_m} : m \ge 1\}$ so that $\nu_{k_m} \Longrightarrow \nu$. Then $\nu \in \mathcal{F}_M$ and, because $\Psi \restriction \mathcal{F}_M$ is continuous, $x_{k_m} = \Psi(\nu_{k_m}) \longrightarrow x = \Psi(\nu)$. Because $H(\nu|\mu) \le \underline{\lim}_{m\to\infty} H(\nu_{k_m}|\mu) \le L$, $J_\mu(x) \le L$.

To prove that J_μ is convex, suppose that $x_1, x_2 \in E$ with $J_\mu(x_1) \vee J_\mu(x_2) < \infty$, and choose $\nu_1,\nu_2 \in \mathcal{I}$ so that $\Psi(\nu_i) = x_i$ and $J_\mu(x_i) = H(\nu_i|\mu)$ for $i \in \{1,2\}$. Then $\Psi\big((1-\theta)\nu_1 + \theta\nu_2\big) = (1-\theta)x_1 + \theta x_2$ and

$$J_\mu\big((1-\theta)x_1 + \theta x_2\big) \le H\big((1-\theta)\nu_1 + \theta\nu_2\big) \le (1-\theta)J_\mu(x_1) + \theta J_\mu(x_2). \quad \square$$

LEMMA 11. *For any closed $F \subseteq E$, $\overline{\lim}_{n\to\infty} \frac{1}{n} \log \mu_n(F) \le -\inf_F J_\mu$.*

PROOF: Refer to the notation in Lemma 8 and the paragraph following the lemma.

For any $M > 0$,

$$\mu_n(F) = \tilde{\mu}_n\big(\{\nu \in \mathcal{I} : \Psi(\nu) \in F\}\big) \le \tilde{\mu}_n\big(\{\nu \in \mathcal{F}_M : \Psi(\nu) \in F\}\big) + \tilde{\mu}_n\big(\mathbf{M}_1(E) \setminus \mathcal{F}_M\big),$$

and so, by Sanov's Theorem and Lemma 8,

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \le -\Big[\inf\big\{H(\nu|\mu) : \nu \in \mathcal{F}_M \ \&\ \Psi(\nu) \in F\big\} \wedge M\Big] \le -\Big[\inf_F J_\mu \wedge M\Big]. \quad \square$$

LEMMA 12. *For each open $G \subseteq E$, $\underline{\lim}_{n\to\infty} \frac{1}{n} \log \mu_n(G) \ge -\inf_G J_\mu$.*

PROOF: Again refer to Lemma 8.

Let $\nu_0 \in \mathcal{I}$ with $H(\nu_0|\mu) < \infty$ and $x_0 = \Psi(\nu_0) \in G$ be given, and choose $r > 0$ so that $B_E(x_0, 2r) \subseteq G$.

By (5), $C \equiv \langle V, \nu_0\rangle \le H(\nu_0|\mu) + \log \mathbb{E}^\mu\big[e^V\big] < \infty$. Hence $\nu_0 \in \mathcal{F}_M$ for any M with $C_M \ge C$. Choose $M > H(\nu_0|\mu) + 2$ so that $C_M \ge C$ and $\ell \in \mathbb{Z}^+$ so that $\frac{C_M}{\ell} < r$. Then $\Psi(\nu) \in B_E(x_0, 2r)$ if $\nu \in \mathcal{F}_M$ and $\Psi_\ell(\nu) \in B_E(x_0, r)$, and so

$$\begin{aligned}\mu_n(G) \ge \mu_n\big(B_E(x_0, 2r)\big) &\ge \tilde{\mu}_n\big(\{\nu \in \mathcal{F}_M : \Psi_\ell(\nu) \in B(x_0, r)\}\big)\\ &\ge \tilde{\mu}_n\big(\Psi_\ell(\nu) \in B_E(x_0, r)\big) - \tilde{\mu}_n\big(\mathbf{M}_1(E) \setminus \mathcal{F}_M\big).\end{aligned}$$

Finally, since $\Psi_\ell(\nu_0) \in B_E(x_0, r)$, Sanov's Theorem and Lemma 8 say that, for each $0 < \delta < 1$,

$$\tilde{\mu}_n\big(\Psi_\ell(\nu) \in B_E(x_0, r)\big) \ge e^{-n(H(\nu_0|\mu)+\delta)} \text{ and } \tilde{\mu}\big(\mathbf{M}_1(E) \setminus \mathcal{F}_M\big) \le e^{-n(M-\delta)}$$

for all sufficiently large n's. Hence, $\underline{\lim}_{n\to\infty} \frac{1}{n} \log \tilde{\mu}_n(G) \ge -H(\nu_0|\mu) - \delta$. $\square$

By combining Lemmas 11 and 12, one sees that $\{\mu_n : n \ge 1\}$ satisfies the full large deviations principle with the good rate function J_μ. As a consequence of this and Varadhan's Theorem,

$$\Lambda_\mu(x^*) = \sup_{x\in E}\big(\langle x, x^*\rangle - J_\mu(x)\big).$$

Since J_μ is l.s.c. and convex, this means that (7) holds.

THEOREM 13 (**Cramér**). *The rate function Λ^*_μ is good, and $\{\mu_n : n \ge 1\}$ satisfies the full large deviations principle with respect to it. In addition, (7) holds.*

Gibbs Measures

Given $x^* \in E^*$, define the **Gibbs measure** $\gamma_{x^*} \in \mathbf{M}_1(E)$ by

$$\gamma_{x^*}(dy) = \frac{1}{M_\mu(x^*)} e^{\langle y, x^*\rangle}\, \mu(dy) \quad \text{where } M_\mu(x^*) = \int e^{\langle y, x^*\rangle}\, \mu(dy), \tag{14}$$

LEMMA 15. *For each $x^* \in E^*$, $\gamma_{x^*} \in \mathcal{I}$ and*

$$H(\gamma_{x^*}|\mu) = \big\langle \Psi(\gamma_{x^*}), x^*\big\rangle - \Lambda_\mu(x^*) = \Lambda^*_\mu\big(\Psi(\gamma_{x^*})\big). \tag{16}$$

In particular, $H(\nu|\mu) > H(\gamma_{x^*}|\mu)$ for any $\nu \in \mathcal{I} \setminus \{\gamma_{x^*}\}$ with $\Psi(\nu) = \Psi(\gamma_{x^*})$.

Proof: The first equality in (16) is obvious. To prove the second, suppose that $\nu \in \mathcal{I}$ with $\Psi(\nu) = x \equiv \Psi(\gamma_{x^*})$ and $H(\nu|\mu) < \infty$. Set $f = \frac{d\nu}{d\gamma_{x^*}}$ and $g = \frac{d\gamma_{x^*}}{d\mu}$, and note that

$$H(\nu|\mu) = \int \log f \, d\nu + \int \log g \, d\nu = H(\nu|\gamma_{x^*}) + \int \langle y, x^* \rangle \, \nu(dy) - \Lambda_\mu(x^*) \geq \langle x, x^* \rangle - \Lambda_\nu(x^*) = H(\gamma_{x^*}|\mu).$$

Hence, by (7), $H(\gamma_{x^*}|\mu) = \Lambda_\mu^*(x)$, and equality in the preceding hold only if $\nu = \gamma_{x^*}$. □

Theorem 17. *Given $x \in E$, there exists a $\nu \in \mathcal{I}$ such that $\Psi(\nu) = x$ and $H(\nu|\mu) < \infty$ if and only if $\Lambda_\mu^*(x) < \infty$. Moreover, if $\Lambda_\mu^*(x) < \infty$, then there exists a unique $\nu \in \mathcal{I}$ such that $\Psi(\nu) = x$ and $H(\nu|\mu) = \Lambda_\mu^*(x)$. Finally, $x^* \in E^*$ satisfies $\Psi(\gamma_{x^*}) = x$ if and only if $\langle x, x^* \rangle - \Lambda_\mu(x^*) = \Lambda_\mu^*(x)$, in which case $H(\gamma_{x^*}|\mu) = \Lambda_\mu^*(x)$.*

Proof: The only assertion yet to be proved is that $\langle x, x^* \rangle - \Lambda_\mu(x^*) = \Lambda_\mu^*(x) \implies \Psi(\gamma_{x^*}) = x$. But $\langle x, x^* \rangle - \Lambda_\mu(x^*) = \Lambda_\mu^*(x)$ implies that

$$y^* \in E^* \longmapsto F(y^*) = \langle x, y^* \rangle - \log \int e^{\langle y, y^* \rangle} \, d\mu(dy)$$

achieves a maximum at x^*, and therefore, by the first derivative test,

$$x - \Psi(\gamma_{x^*}) = x - \int y \, \gamma_{x^*}(dy) = DF(x^*) = 0. \quad \square$$

Corollary 18. *If $(H, E, \mathcal{W})$ is an abstract Wiener space and $\nu \in \mathbf{M}_1(E)$, then $H(\nu|\mathcal{W}) < \infty$ implies that $\int \|x\|_E^2 \, \nu(dx) < \infty$ and that $\Psi(\nu) \in H$. Furthermore, for any $x \in E$ and $x^* \in E$, $x = h_{x^*}$ if and only if $x = \Psi(\gamma_{x^*})$. Thus, if $x \in E$, then*

$$y^* \in E^* \longmapsto \int e^{\langle y - x, y^* \rangle} \, \mathcal{W}(dy) \in (0, \infty)$$

achieves a minimum if and only if $x = h_{x^*}$ for some $x^* \in E^*$.

Proof: By Fernique's Theorem, $A = \mathbb{E}^{\mathcal{W}}\big[e^{\alpha \|x\|_E^2}\big] < \infty$ is for some $\alpha > 0$. Thus, if $H(\nu|\mathcal{W}) < \infty$, then, by (5), $\alpha \int \|x\|_E^2 \, d\nu \leq H(\nu|\mu) + \log A < \infty$. Furthermore, by (7), $\Lambda_{\mathcal{W}}^*\big(\Psi(\nu)\big) < \infty$, and therefore $\Psi(\nu) \in H$. The second assertion is simply that observation that $\gamma_{x^*} = T_{h_{x^*}} \mathcal{W}$ and therefore that $\Psi(\gamma_{x^*}) = h_{x^*}$. Given the earlier ones, the final assertion is an easy application of the last part of Theorem 17. □