Large deviations for mixtures

J. D. Biggins*

September 24, 2003

Abstract

Suppose the distributions $\{\mu^n\}$ on Θ obey a large deviation principle (LDP) and that for each $\theta \in \Theta$ the distributions $\{P^n_\theta\}$ on X also obey an LDP. The main purpose of this paper is to give conditions which allow an LDP for the mixtures $\{P^n\}$, given by $P^n(A) = \int P^n_\theta(A)d\mu^n(\theta)$, to be deduced. The treatment follows that of Dinwoodie and Zabell (1992) who, motivated by exchangeability, consider the case where $\{\mu^n\}$ does not vary with n.

1 Introduction

Let μ^n be a (mixing) probability measure on the Borel σ-algebra of a topological space Θ, concentrated on (the measurable set) Θ_n. For each $\theta \in \Theta_n$, let P^n_θ be a probability measure on the Borel σ-algebra of the topological space X, such that, for every measurable $A \subset X$, the map $\theta \to P^n_\theta(A)$ is measurable on Θ_n. For definiteness, let P^n_θ be given by some fixed probability measure on X when $\theta \notin \Theta_n$. (The complication of allowing Θ_n to depend on n occurs in the example which led to the search for a general theorem on mixtures.) Based on these, the joint distribution, $\widetilde{P}^n$, and the marginal distribution, P^n, obtained by mixing over θ, have the usual definitions:

$$d\widetilde{P}^n(\theta, x) = dP^n_\theta(x)d\mu^n(\theta) \quad \text{and} \quad dP^n(x) = \int_\Theta dP^n_\theta(x)d\mu^n(\theta) = \int_{\Theta_n} dP^n_\theta(x)d\mu^n(\theta).$$

Throughout Θ and X are assumed to be Hausdorff (distinct points can be separated by disjoint open sets) and Θ is assumed to be first countable (for each θ there is a countable collection of neighbourhoods such that every neighbourhood of θ contains one of this collection), which implies that convergence in Θ can be described by using sequences.

The sequence of probability measures (P^n) (on the Borel σ-algebra of a topological space X) obeys a large deviation principle (LDP) if there is a lower semicontinuous non-negative function λ (a rate function) such that for every closed F and open G

$$\limsup \frac{\log P^n(F)}{n} \le -\inf_{y \in F} \lambda(x) \quad \text{and} \quad \liminf \frac{\log P^n(G)}{n} \ge -\inf_{y \in G} \lambda(x)$$

The rate function λ is called 'good' (or proper) if for every finite β the set $\{x : \lambda(x) \le \beta\}$ is compact. The sequence satisfies a weak LDP if the upper bound holds for compact, rather than closed, F. Furthermore, the sequence of probability measures (P^n) is said to

*The University of Sheffield, U.K. (j.biggins@sheffield.ac.uk)

be exponentially tight if for every $\alpha > 0$ there is a set O_α whose complement is compact with

$$\limsup \frac{\log P^n(O_\alpha)}{n} < -\alpha. \tag{1}$$

The main idea is to combine large deviation results for P^n_θ and μ^n to give large deviation results for P^n. The treatment draws heavily on that in Section 2 of Dinwoodie and Zabell (1992), who consider the case where μ^n does not depend on n.

Four assumptions will be in force in all of the Theorems, but not necessarily in the supporting Lemmas. These are now described, for the third of these a little more notation is needed. Let $\widetilde{\Theta}$ be the limit set of sequences with the nth member from Θ_n; thus

$$\widetilde{\Theta} = \{\theta \in \Theta : \exists\, \theta(n) \in \Theta_n, \theta(n) \to \theta\}.$$

It is easy to check that $\widetilde{\Theta}$ is closed; see Lemma 6 in the next section. Most applications will have $\widetilde{\Theta} = \Theta$.

A 1 *(μ^n) satisfies an LDP with rate ψ.*

This is called a mixing LDP. Exponential tightness is also needed.

A 2 *(μ^n) is exponentially tight.*

A 3 *$\widetilde{\Theta}$ is non-empty and whenever $\theta(n) \in \Theta_n$ and $\theta(n) \to \theta \in \widetilde{\Theta}$, $\{P^n_{\theta(n)}\}$ satisfies an LDP with rate λ_θ. Also, when $\theta \notin \widetilde{\Theta}$, let $\lambda_\theta(x) = \infty$ for all x.*

Following Dinwoodie and Zabell (1992), this is called exponential continuity. Since λ_θ is a rate function it is lower semicontinuous on X for each θ. The third assumption is in similar vein.

A 4 *$\lambda_\theta(x)$ is jointly lower semi-continuous in $(\theta, x) \in \Theta \times X$.*

Lemma 3.1(i) in Dinwoodie and Zabell (1992) gives some general conditions for A4 to hold.

The main results are now stated. Some other results, for example on an LPD for the joint distributions $(\widetilde{P}^n)$, which are judged to be of less interest, are stated in the course of the proofs of the main results. Recall that a topological space is regular if for every open U containing x there is an open O also containing x with its closure contained in U.

Theorem 1 *Suppose A1-4 hold. If Θ is regular and ψ is good then $\{P^n\}$ satisfies an LDP with rate function*

$$\lambda(x) = \inf\{\lambda_\theta(x) + \psi(\theta) : \theta \in \Theta\}. \tag{2}$$

If, in addition, X is regular and, for each $\theta \in \widetilde{\Theta}$, λ_θ is good then λ is good. When Θ is compact and ψ takes only the value 0 the requirements that Θ is regular and (μ^n) is exponentially tight are not needed.

A family of sequences (of probability measures) is uniformly exponentially tight if, in (1), for every $\alpha > 0$ the same O_α can be used for every sequence.

Theorem 2 *Suppose A1-4 hold. If* Θ *and* X *are both regular and for each* $\theta \in \widetilde{\Theta}$*, the sequences* $\{P^n_{\theta(n)} : \theta(n) \in \Theta_n,\ \theta(n) \to \theta\}$ *are uniformly exponentially tight then* $\{P^n\}$ *satisfies an LDP with the good rate function* λ *defined at (2).*

Theorem 3 *Suppose A1-4 hold. If* Θ *and* X *are both regular,* X *is locally compact,* (μ^n) *is exponentially tight and for each* $\theta \in \widetilde{\Theta}$*,* λ_θ *is good then* $\{P^n\}$ *satisfies an LDP with the good rate function* λ *defined at (2).*

The next result, which is contained in exercises 1.2.19 and 4.1.10 in Dembo and Zeitouni (1993), notes that often the rate function being good implies exponential tightness. It shows that the hypothesis A2 in Theorems 1 and 3 is superfluous when Θ is locally compact or Polish.

Lemma 4 *If* ψ *is good and* Θ *is either locally compact or Polish then* (μ^n) *is exponentially tight.*

For easier references in the statement and proofs in the rest of the paper, three additional assumptions are also labelled.

A 5 ψ *is good.*

A 6 *For each* $\theta \in \widetilde{\Theta}$*,* $\lambda_\theta(x)$ *is good.*

A 7 *For each* $\theta \in \widetilde{\Theta}$*, the sequences* $\{P^n_{\theta(n)} : \theta(n) \in \Theta_n,\ \theta(n) \to \theta\}$ *are uniformly exponentially tight.*

Lemma 3.2 in Dunwoodie and Zabell (1992) gives some general conditions under which A7 holds when, for each θ, P^n_θ is the distribution of the average of independent identically distributed variables.

Dunwoodie and Zadell (1992) used their general results to consider large deviations for exchangeable sequences in rather general spaces; this motivation led naturally to the assumption that μ^n was independent of n. The motivating example for developing the results described here is much less topologically sophisticated. It is a problem arising in the study of random graphs. The classical random graph is very well understood, but fails to match up to the graphs occuring in many applications. Recently, Cannings and Penman (2003) suggested a model with more flexibility; see also Penman (1998). Suppose a graph is to have n vertices. Then, to produce random graphs with a correlation structure beteen edge occurences, Cannings and Penman (2003) proposed that each vertex is independently assigned one of a number of colours, and the probability that an edge arises depends on the colours of its two vertices. The problem posed is to find an LDP for the number of edges, as n becomes large. This falls exactly into the framework proposed. To ellucidate, consider the graph with n vertices. Let the proportions of these vertices of the various possible colours be θ; then μ^n is the distribution θ. Given n and θ the number of edges is obtained as the sum of independent (but not identically distributed) random varibales; this specifies P^n_θ. Note that for finite n the possible values of θ are confined to those with $n\theta$ containing integers; this defines Θ_n here. The details of this application are discussed in Biggins and Penman (2003).

The next section contains the proof of Theorems 1, the following one contains proofs of Theorems 2 and 3. A brief final section mentions some possible directions for further work.

2 Proof of Theorem 1

A function f on X is called lower semicontinuous at x if for each $c < f(x)$ there is an open U containing x such that $f(y) > c$ for every $y \in U$.

Lemma 5 *If X is regular and f is lower semicontinuous on X then for every x and $c < f(x)$ there is a closed set C_x with x in its interior and $f(y) > c$ for all $y \in C_x$.*

Proof. Fix x and $c < f(x)$. By the definition of lower semicontinuity, there is an open set U containing x with $f(y) > c$ for $y \in U$. Applying regularity, there is an open set V_x containing x with its closure inside O_x. Take C_x to be the closure of V_x. □

Lemma 6 *$\widetilde{\Theta}$ is closed.*

Proof. Suppose that $\theta_n^{(k)} \to \theta^{(k)} \to \theta$ with $\theta_n^{(k)} \in \Theta_n$. Take U_i from a countable open neighbourhood base of θ. For some $k(i) > k(i-1)$, $\theta^k \in U_i$ for $k \geq k(i)$. Now U_i is also an open neighbourhood of $\theta^{k(i)}$ and so there is an $n^{k(i)}$ with $\theta_n^{(k(i))} \in U_i$ for all $n \geq n^{k(i)}$. Now the sequence $\vartheta_n = \theta_n^{(k(i))} \in \Theta_n$ for $n^{k(i)} \leq n < n^{k(i+1)}$ converges to θ. □

Lemma 7 *Suppose the mixing LDP (i.e. A1) and exponential continuity (i.e. A3) hold. Let $(\theta, x) \in G^* \subset \Theta \times X$, where G^* is open. Then*

$$\liminf \frac{\log \widetilde{P}^n(G^*)}{n} \geq -(\lambda_\theta(x) + \psi(\theta)).$$

In particular, for G open in X,

$$\liminf \frac{\log P^n(G)}{n} \geq -\inf\{\lambda_\theta(x) + \psi(\theta) : \theta \in \Theta, x \in G\}.$$

Proof. The result is true when $\lambda_\theta(x) = \infty$. Hence attention can focus on $\lambda_\theta(x) < \infty$. There are open sets $O \subset \Theta$ and $U \subset X$ containing θ and x respectively with $O \times U \subset G^*$. Then

$$\widetilde{P}^n(O \times U) = \int_O P_\vartheta^n(U) d\mu^n(\vartheta).$$

By assumption A3, for any $\theta(n) \in \Theta_n$ with $\theta(n) \to \theta$

$$\liminf \frac{\log P_{\theta(n)}^n(U)}{n} \geq -\lambda_\theta(x)$$

and so, just as in the proof of Theorem 2.1 of Dinwoodie and Zabell (1992), for every $\epsilon > 0$, there exists an open set $O_\theta \subset O$ containing θ and an integer N_θ such that for $n \geq N_\theta$ and every $\gamma \in O_\theta \cap \Theta_n$

$$P_\gamma^n(U) > \exp(-n[\lambda_\theta(x) + \epsilon]).$$

To demonstrate this, suppose it fails. Then there are $\theta(i) \to \theta$ and $n(i) > n(i-1)$ such that

$$P_{\theta(i)}^{n(i)}(U) \leq \exp(-n[\lambda_\theta(x) + \epsilon]),$$

and then

$$\liminf \frac{\log P_{\theta(i)}^{n(i)}(U)}{n} \leq -\lambda_\theta(x) - \epsilon,$$

which contradicts the lower bound in the LDP in A3.

Thus, for $n \geq N_\theta$,

$$\begin{aligned}\widetilde{P}^n(G^*) \geq \widetilde{P}^n(O_\theta \times U) &= \int_{O_\theta} P^n_\vartheta(U) d\mu^n(\vartheta) \\ &= \int_{O_\theta \cap \Theta_n} P^n_\vartheta(U) d\mu^n(\vartheta) \\ &\geq \exp(-n[\lambda_\theta(x) + \epsilon])\mu^n(O_\theta \cap \Theta_n) \\ &= \exp(-n[\lambda_\theta(x) + \epsilon])\mu^n(O_\theta)\end{aligned}$$

and so, using the mixing LDP, A1,

$$\liminf \frac{\log \widetilde{P}^n(G^*)}{n} \geq -\lambda_\theta(x) - \epsilon + \liminf \frac{\log \mu^n(O_\theta)}{n} \geq -\lambda_\theta(x) - \epsilon - \psi(\theta).$$

The last part comes from taking $G^* = \Theta \times G$, for then $\widetilde{P}^n(G^*) = P^n(G)$. □

Lemma 8 *Suppose the mixing LDP (i.e. A1) and exponential continuity (i.e. A3) hold. Suppose too that (μ^n) is exponentially tight (i.e. A2) and Θ is regular. Let $F \subset X$ be closed. Then*

$$\limsup \frac{\log P^n(F)}{n} \leq -\inf\{\lambda_\theta(x) + \psi(\theta) : (\theta, x) \in \Theta \times F\}.$$

Proof. The idea of the proof follows that of Theorem 2.2 in Dinwoodie and Zabell (1992).

Fix F. Let c and d be such that

$$c < d = \inf\{\lambda_\theta(x) + \psi(\theta) : (x, \theta) \in F \times \Theta\}.$$

Using exponential tightness (i.e. A2), let O be such that

$$\limsup \frac{\log \mu^n(O)}{n} < -c$$

and let S be the (compact) complement of O. Then

$$P^n(F) = \int_\Theta dP^n_\theta(F) d\mu^n(\theta) \leq \int_S dP^n_\theta(F) d\mu^n(\theta) + \mu^n(O).$$

Let $\Lambda(\theta) = \inf\{\lambda_\theta(x) : x \in F\}$. Let $\epsilon > 0$ with $c < 1/\epsilon$. Now let

$$\Lambda^\epsilon(\theta) = \min\{\Lambda(\theta) - \epsilon, 1/\epsilon\} \quad \text{and} \quad \psi^\epsilon(\theta) = \min\{\psi(\theta) - \epsilon, 1/\epsilon\}.$$

For $\theta \in S$, by exponential continuity (i.e. A3), as in the proof of Theorem 2.2 in Dinwoodie and Zabell (1992), there is an open set U_θ containing θ and an integer N_θ such that for $n \geq N_\theta$ and every $\gamma \in U_\theta \cap \Theta_n$

$$P^n_\gamma(F) \leq \exp(-n\Lambda^\epsilon(\theta)).$$

To demonstrate this, suppose it fails. Then there are $\theta(i) \to \theta$ and $n(i) > n(i-1)$ such that

$$P^n_{\theta(i)}(F) > \exp(-n\Lambda^\epsilon(\theta)),$$

and then

$$\limsup \frac{\log P^n_{\theta(i)}(F)}{n} \geq -\Lambda^\epsilon(\theta)$$

which contradicts the upper bound in the LDP in A3.

Furthermore, using the lower semicontinuity of ψ, by taking U_θ to be smaller if necessary,

$$\psi(\vartheta) > \psi^\epsilon(\theta) \text{ for } \vartheta \in U_\theta,$$

and, using regularity of Θ, there is an open set V_θ with closure $\overline{V}_\theta$ such that $\theta \in V_\theta$ and $\overline{V}_\theta \subset U_\theta$.

Now $(V_\theta : \theta \in S)$ is an open covering of S. Since S is compact a finite subcover $(V_{\theta(i)})_{1\leq i\leq k}$ exists. Then, for sufficiently large n,

$$\begin{aligned}
P^n(F) &\leq \mu^n(O) + \sum_{i=1}^k \int_{V_{\theta(i)}} P^n_\vartheta(F) d\mu^n(\vartheta) \\
&= \mu^n(O) + \sum_{i=1}^k \int_{V_{\theta(i)}\cap\Theta_n} P^n_\vartheta(F) d\mu^n(\vartheta) \\
&\leq \mu^n(O) + \sum_{i=1}^k \exp(-n\Lambda^\epsilon(\theta(i)))\mu^n\left(\overline{V}_{\theta(i)}\right) \\
&\leq \mu^n(O) + \sum_{i=1}^k \exp(-n\Lambda^\epsilon(\theta(i)))\exp(-n\psi^\epsilon(\theta(i))).
\end{aligned}$$

Hence, since $c < 1/\epsilon$,

$$\begin{aligned}
\limsup \frac{\log P^n(F)}{n} &\leq -\min\left\{\min_{1\leq i\leq k}\{\Lambda^\epsilon(\theta(i)) + \psi^\epsilon(\theta(i))\}, c\right\} \\
&\leq -\min\left\{\min_{1\leq i\leq k}\{\Lambda(\theta(i)) + \psi(\theta(i)) - 2\epsilon\}, c\right\} \\
&\leq -\min\{d - 2\epsilon, c\}.
\end{aligned}$$

Since $c < d$ and $\epsilon > 0$ are arbitrary, the result follows. □

Lemma 9 *In Lemma 8, if Θ is compact and ψ takes only the value 0 then the hypotheses that (μ^n) is exponentially tight and Θ is regular are not needed.*

Proof. When Θ is compact A2 holds automatically. When ψ takes only the value 0 there is no need to introduce V_θ; it suffices to take a finite subcover from $(U_\theta : \theta \in \Theta)$. This is what is done in Dinwoodie and Zabell (1992). □

These lemmas provide most of what is needed to give the LPD for P^n. The final ingredient is the check that the putative rate function is indeed lower semicontinuous.

Lemma 10 *If $\lambda_\theta(x)$ is jointly lower semi-continuous (i.e. A4) and ψ is good (i.e. A5) then*

$$\lambda(x) = \inf\{\lambda_\theta(x) + \psi(\theta) : \theta \in \Theta\}$$

is lower semicontinuous.

Proof. It must be shown that for every x and $c < \lambda(x)$ there is a neighbourhood U of x with $\lambda(y) > c$ for every $y \in U$.

Fix x, $c < \lambda(x)$ and $\epsilon > 0$. Let $C = \{\theta : \psi(\theta) \leq c\}$, which is compact because ψ is good. Now let

$$\overline{\lambda}(y) = \inf\{\lambda_\theta(y) + \psi(\theta) : \theta \in C\}.$$

Then, since λ_θ is non-negative, it is easy to see that

$$\overline{\lambda}(y) \geq \lambda(y) \geq \min\{\overline{\lambda}(y), c\}$$

which implies that $\{y : \lambda(y) \leq c\}$ and $\{y : \overline{\lambda}(y) \leq c\}$ are the same, and that λ and $\overline{\lambda}$ agree on this set.

Let

$$\lambda^\epsilon_\theta(x) = \min\{\lambda_\theta(x) - \epsilon, 1/\epsilon\} \text{ and } \psi^\epsilon(\theta) = \min\{\psi(\theta) - \epsilon, 1/\epsilon\}.$$

For each θ, because $\lambda_\theta(y)$ is jointly lower semicontinuous, and ψ is lower semicontinuous there are open sets $O_\theta \subset \Theta$ and $U_\theta \subset X$ containing θ and x respectively such that throughout $O_\theta \times U_\theta$

$$\lambda_\vartheta(y) > \lambda^\epsilon_\theta(x) \text{ for } (\vartheta, y) \in O_\theta \times U_\theta$$

and

$$\psi(\vartheta) > \psi^\epsilon(\theta) \text{ for } \vartheta \in O_\theta.$$

The $\{O_\theta : \theta \in C\}$ cover C, and so there is a finite subcover, $(O_{\theta(i)})_{1\leq i\leq k}$. Let $U = \cap_i U_{\theta(i)}$, which is open and contains x. Then for $y \in U$

$$\begin{aligned}
\overline{\lambda}(y) &\geq \min_i\{\inf\{\lambda_\theta(y) + \psi(\theta) : \theta \in O_{\theta(i)}\}\} \\
&\geq \min_i\{\lambda^\epsilon_{\theta(i)}(x) + \psi^\epsilon(\theta(i))\} \\
&\geq \min\{\lambda(x) - 2\epsilon, (2\epsilon)^{-1}\} > c,
\end{aligned}$$

provided ϵ is small enough. Then

$$x \in U \subset \{y : \overline{\lambda}(y) > c\} = \{y : \lambda(y) > c\}$$

proving the result. □

Lemma 11 *Suppose A1-4 hold. In addition suppose ψ is good (i.e. A5), λ_θ is good for $\theta \in \widetilde{\Theta}$ (i.e. A6) and both Θ and X are regular. Then the rate function*

$$\lambda(x) = \inf\{\lambda_\theta(x) + \psi(\theta) : \theta \in \Theta\}.$$

is good.

Proof. Much of this proof follows that of Lemma 2.1 in Dinwoodie and Zabell (1992). However, a couple of points need additional argument. Take $\epsilon > 0$. Take α with $0 \leq \alpha < \alpha + 2\epsilon < \infty$. Let K be $\{\theta : \psi(\theta) \leq \alpha + 2\epsilon\}$ which is compact because ψ is good. Now let

$$\overline{\lambda}(y) = \inf\{\lambda_\theta(y) + \psi(\theta) : \theta \in K\};$$

then, as explained in the proof of Lemma 10,

$$\{y : \overline{\lambda}(y) \leq \alpha\} = \{y : \lambda(y) \leq \alpha\}.$$

Denote this set by L^α and suppose α was selected so that L^α is not compact. Then there exists a net $\{(\theta(i), x(i)) : i \in I\} \subset K \times X$ such that $\{x(i)\} \subset L^\alpha$, $\{x(i)\}$ has no convergent subnet and

$$\lambda_{\theta(i)}(x(i)) + \psi(\theta(i)) \leq \alpha + \epsilon$$

for all i. Note that this implies that $\theta(i) \in \widetilde{\Theta}$. Since K is compact and first countable there is a subsequence $(\theta(i_k), x(i_k))$ such that $\theta(i_k) \to \theta$, where $\theta \in \widetilde{\Theta}$ since $\widetilde{\Theta}$ is closed. Furthermore, because ψ is lower semicontinuous, $\liminf \psi(\theta(i)) \geq \psi(\theta)$ and $\psi(\theta) \leq \alpha + \epsilon$.

Take $\beta = \alpha + 3\epsilon$. The level set of λ_θ given by $L_\theta^\beta = \{x : \lambda_\theta(x) \leq \beta - \psi(\theta)\}$ is compact because λ_θ is good. Hence, for large enough k_0, $C_0 = \{x(i_k); k \geq k_0\}$ must be in the complement of L_θ^β. Then, following the argument in Dinwoodie and Zabell (1992), there is an open set U containing C_0 with closure C in the complement of L_θ^β.

Now take $\vartheta(i,n) \in \Theta_n$ with $\vartheta(i,n) \to \theta(i)$. By the LDP lower bound in A3,

$$\liminf \frac{\log P^n_{\vartheta(n,i)}(U)}{n} \geq -\inf\{\lambda_{\theta(i)}(x) : x \in U\} \geq \psi(\theta(i)) - \alpha - \epsilon.$$

Hence, selecting suitable subsequences, there is an increasing sequence $n(k)$ and $\vartheta(k) \in \Theta_{n(k)}$ such that $\vartheta(k) \to \theta$ and

$$\frac{\log P^{n(k)}_{\vartheta(k)}(U)}{n(k)} \geq \psi(\theta) - \alpha - 2\epsilon.$$

By the LDP upper bound in A3

$$\limsup \frac{\log P^{n(k)}_{\vartheta(k)}(C)}{n(k)} \leq \psi(\theta) - \beta = \psi(\theta) - \alpha - 3\epsilon,$$

since C is in the complement of L_θ^β. Since $U \subset C$ this contradicts the previous inequality. Therefore L^α must be compact. □

Proof of Theorem 1. This follows directly from the previous results. The last part of Lemma 7 gives the lower bound for open sets, Lemma 8 gives the upper bound for closed sets and Lemma 10 confirms that λ is lower semicontinuous. Lemma 11 shows that λ is good under the stated conditions and Lemma 9 gives the simplification contained in the final assertion. □

3 Proof of Theorems 2 and 3

Proposition 1 *Suppose A1, A3 and A4 hold. Suppose too that both Θ and X are regular. Then $\{\widetilde{P}^n\}$ satisfies a weak LDP with rate function $\lambda_\theta(x) + \psi(\theta)$. When Θ is locally compact the LDP in A1 can be replaced by a weak LDP. Similarly, when X is locally compact a weak LDP is enough in A3.*

Proof. First, the required result for open sets is contained in Lemma 7. Second, by A1 and A4, $\lambda_\theta(x) + \psi(\theta)$ is lower semicontinuous. It remains to show that the required bound holds for compact sets.

Fix $F \subset \Theta \times X$, compact. By lower semicontinuity and regularity, for each (θ, x) there are open sets $O \subset \Theta$ and $U \subset X$ containing θ and x respectively, with closures $\overline{O}$ and $\overline{U}$, such that

$$\lambda_\vartheta(y) > \lambda_\theta^\epsilon(x) = \min\{\lambda_\theta(x) - \epsilon, 1/\epsilon\} \text{ for } (\vartheta, y) \in \overline{O} \times \overline{U}$$

and

$$\psi(\vartheta) > \psi^\epsilon(\theta) \text{ for } \vartheta \in \overline{O}.$$

By taking O to be smaller, if necessary, there is an integer N such that for $n \geq N$ and $\gamma \in O \cap \Theta_n$

$$P_\gamma^n(\overline{U}) \leq \exp(-n\lambda^\epsilon(x,\theta)) \quad \text{and} \quad \mu^n(\overline{O}) \leq \exp(-n\psi^\epsilon(\theta)).$$

Thus, for $n \geq N$,

$$\begin{aligned} \widetilde{P}^n(O \times U) &\leq \int_{\overline{O} \cap \Theta_n} P_\vartheta^n(\overline{U}) d\mu^n(\vartheta) \\ &\leq \exp(-n\lambda_\theta^\epsilon(x))\mu^n(\overline{O}) \\ &\leq \exp(-n\lambda_\theta^\epsilon(x)) \exp(-n\psi^\epsilon(\theta)). \end{aligned}$$

Hence

$$\limsup_n \frac{\log \widetilde{P}^n(U \times O)}{n} \leq -\lambda_\theta^\epsilon(x) - \psi^\epsilon(\theta).$$

As (x, θ) varies over F the corresponding sets $U \times O$ cover F. Taking a finite subcover, using it to get an upper bound on $\widetilde{P}(F)$ and then letting ϵ go to zero completes the proof. In the locally compact cases, O and U can be taken so that $\overline{U}$ and $\overline{O}$ are compact and so a weak LDP is enough to bound the corresponding terms. □

Lemma 12 *If A2 and A7 hold then $\{\widetilde{P}^n\}$ is exponentially tight.*

Proof. Fix α. Using A2, let O be such that

$$\limsup \frac{\log \mu^n\{O\}}{n} < -\alpha$$

and let S be the (compact) complement of O. For $\theta \in \widetilde{\Theta}$, let $U_\theta \subset X$ be a set with compact complement such that for any $\theta(n) \in \Theta_n$ with $\theta(n) \to \theta$

$$\limsup \frac{\log P_{\theta(n)}^n\{U_\theta\}}{n} < -\alpha.$$

The existence of U_θ is guaranteed by A7. For $\theta \notin \widetilde{\Theta}$, let $U_\theta = X$. Then there is an open set V_θ, containing θ, and an integer N_θ such that for $n \geq N_\theta$ and $\gamma \in V_\theta \cap \Theta_n$

$$P_\gamma^n(U_\theta) < \exp(-n\alpha).$$

Otherwise a suitable subsequence contradicts A7.

The collection $\{V_\theta : \theta \in S\}$ cover S. Take a finite cover $(V_{\theta(i)} : 1 \leq i \leq k)$ of S; then let U be the set $\cap_i U_{\theta(i)}$ and K be its complement, which, as the union of a finite number of compact sets is itself compact. Then $(O \times X) \cup (S \times U)$ has the complement $S \times K$, which is compact, and, for n large enough

$$\begin{aligned} \widetilde{P}^n \left((O \times X) \cup (S \times U) \right) &\leq \mu^n(O) + \sum_{i=1}^{k} \int_{V_{\theta(i)} \cap \Theta_n} P_\vartheta(U_{\theta(i)}) d\mu^n(\vartheta) \\ &< (k+1) \exp(-n\alpha). \end{aligned}$$

Hence

$$\limsup \frac{\log \widetilde{P}^n (O \times X) \cup (S \times U))}{n} \leq -\alpha,$$

which suffices, since α was arbitrary. □

Theorem 13 *Suppose A1-4 hold. Suppose too that A7 holds and both Θ and X are regular. Then $\{\widetilde{P}^n\}$ satisfies an LDP with the good rate function $\lambda_\theta(x) + \psi(\theta)$.*

Proof. This follows immediately from Lemma 1.2.18 in Dembo and Zeitouni (1993) together with Theorem 1 and Lemma 12. □

Proof of Theorem 2. This is an application of the Theorem 13 and the contraction principle (given in Theorem 4.2.1 of Dembo and Zeitouni (1993)) applied to the projection from $\Theta \times X$ to X. □

Lemma 14 *If exponential continuity (i.e. A3) holds, λ_θ is good for $\theta \in \widetilde{\Theta}$ (i.e. A6) and X is locally compact then A7 holds.*

Proof. Locally compact means that for every $x \in X$ there is U_x open and C_x compact with $x \in U_x \subset C_x$. Fix $\theta \in \widetilde{\Theta}$. Take $\beta < \alpha$. Since λ_θ is good,

$$K = \{x : \lambda_\theta(x) \leq \alpha\}$$

is compact. Let $\{U_{x(i)} : i = 1, 2, \ldots, k\}$ be a finite subcover of $\{U_x : x \in K\}$. Now let O be the complement of the compact set $\cup_i C_{x(i)}$, and let F be the complement of the open set $\cup_i U_{x(i)}$.

Consider $\{P^n_{\theta(n)}\}$ where $\theta(n) \in \Theta_n$, and $\theta(n) \to \theta$. Then, by A3,

$$\limsup \frac{\log P^n_{\theta(n)}(O)}{n} \leq \limsup \frac{\log P^n_{\theta(n)}(F)}{n} \leq -\inf_{y \in F} \lambda_\theta(y) \leq -\alpha < -\beta.$$

Since the set O is independent of the particular sequence $(\theta(n))$ the result is proved. □

Proof of Theorem 3. This follows immediately from Theorem 2 and Lemma 14. □

4 Possible extensions and refinements

This is a brief note of things that have not been attempted but seem to have some interest.

Clearly it would be desirable to have conditions, other than local compactness of X, which force the rather awkward assumption A7 to hold.

This note aims to generalize Theorem 2.3 in Dinwoodie and Zabell (1992). In that Theorem, the mixing LDP, A1, and the associated exponential tightness, A2, hold automatically, while exponential continuity, A3, and joint lower semicontinuity of $\lambda_\theta(x)$, A4, are taken as hypotheses. In a further study, Dinwoodie and Zabell (1993), they give results that relax these assumptions and also their assumption that Θ is compact, which is, in a sense, analogous to A2 here.

Finally, the approach to large deviations described in Puhalskii (2001) could be explored. Theorem 1.8.9 and Lemma 1.8.12 there are relevant. Roughly translated into the language here, they give conditions on $\psi(\theta)$ and $\lambda_\theta(x)$ which make $\lambda_\theta(x) + \psi(\theta)$ a rate function on $\Theta \times X$.

Acknowledgements

I am grateful to David Penman who brought Dinwoodie and Zabell (1992) to my attention, worked with me on the motivating problem, reported in Biggins and Penman (2003), and commented on this manuscript.

References

[1] Biggins, J.D., Penman, D. B. Large deviations for randomly coloured random graphs. (2003). Submitted to *Combinatorics Probability and Computing.*
Preprint available at `http://www.shef.ac.uk/~st1jdb/ldircrg.html`

[2] Cannings, C., Penman, D. B. Models of random graphs and their applications *Handbook of Statistics 21. Stochastic Processes: Modeling and Simulation.* Eds: D.N. Shanbhag, and C.R. Rao. Elsevier (2003) 51–91.

[3] Dembo, A., Zeitouni, O. *Large deviation techniques*, Jones and Bartlett, Boston (1993).

[4] Dinwoodie, I. H., Zabell, S. L. Large deviations for exchangeable random vectors. Ann. Prob. **20** (1992) 1147–1166.

[5] Dinwoodie, I. H., Zabell, S. L. Large deviations for sequences of mixtures. In *Statistics and Probability: A Raghu Raj Bahadur Festschrift*, eds: J.K. Ghosh, S.K. Mitra, K.R. Parthasarathy and B.L.S. Prakasa Rao. Wiley Eastern, New Delhi, (1993) 171–190.

[6] Penman, D. B. Random graphs with correlation structure. Ph.D thesis, University of Sheffield, 1998.

[7] Puhalskii, A. *Large Deviations and Idempotent Probability.* Chapman & Hall/CRC, Boca Raton, 2001.

J.D. Biggins
Dept of Probability and Statistics,
Hicks Building,
The University of Sheffield,
Sheffield, S3 7RH,
U.K.