Large Deviations Problem and Shapes of Young Diagrams

V.M.Blinovsky
email: blinov@postman.ru

Abstract

We consider the local process level large deviations principle for the shape of random Young diagrams with different distribution This is the survey of the results which contained in three theorems, obtaining by the author

To the problem of establishing the process level large deviations principle devoted a lot of recent works. Arising from the minimum action principle in physics and from the large deviations problems in probability theory the process level large deviations theory nowdays spread its influence to other branches of mathematics such as number theory, theory of representations. The detailed description of the large deviations theory can be found in monographs [1], [2], [3]. The main idea of the process level large deviations theory (LDP) in discrete space is as follows: for every natural n we have the finite number of curves κ_n which for every n belong to the same metric space L and we have some the distribution on the Borel σ – algebra on L. Then we find the probability

$$P_n(\kappa_n \in B_n(\epsilon, y)),$$

where $B_n(\epsilon, y) = \{z \in L : \rho(z, y) < \epsilon\}$ is the open ball of radius ϵ with the center at $y \in L$. Then the local LDP (LLDP) is the statement that there exists functional $I : L \to [0, \infty]$ and sequence $A_n \to \infty$ such that $I \not\equiv 0, \infty$ and

$$I(y) = -\lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{\ln P_n(\kappa_n \in B_n(\epsilon, y))}{A_n}. \tag{1}$$

Roughly to say the last relation means that

$$P_n(\kappa_n \in B_n(\epsilon, y)) \sim e^{-A_n I(y)}.$$

In some cases it is possible to prove the global LDP (or simply LDP). It establish the existence of the functional $I : L \to [0, \infty]$ and the sequence

$A_n \to \infty$ such that $I \not\equiv 0, \infty$ and for the every Borel set $B \subset L$ the following relations are valid

$$\begin{aligned} -\inf_{y\in B^0} I(y) &\leq \liminf_{n\to\infty} \frac{\ln P_n(\kappa_n \in B)}{A_n} \leq \limsup_{n\to\infty} \frac{\ln P_n(\kappa_n \in B)}{A_n} \\ &\leq -\inf_{y\in \bar{B}} I(y), \end{aligned} \tag{2}$$

where B^0, $\bar{B}$ is the open kernel and closure of the set B correspondingly. Usually (2) follows from (1) and the exponential compactness of the distribution P_n on L. However in many cases P_n does not satisfy the exponential compactness, but (1) is still valid. Moreover we should like to mention the interesting phenomena that the relation (1) has very weak dependence on choice the metric on L. It means that (1) is true whith the same functional for different metrics ρ on L. And usually the exponential compactness is valid for given distribution P_n only for one choosing of metrics on L.

Usually it is necessary to normalize (to scale) the linear sizes of the curves γ_n in order to make the proper definition of the set $\{\kappa_n : \rho(y, \kappa_n) < \epsilon\}$, where κ_n is the scaled shape. In many cases the scailing factor follows from the formulation of the problem and some intuitive considerations about the behaviour of the asymptotics of the linear sizes of the curves κ_n. We will show that in some cases it is necessary to scale the 'time' axis and the values of the shape γ_n in some cases by the same factor and in some cases by the different factor. Moreover there exist examples (and we will introduce one of them) when to obtain the proper solution it is necessary not to divide the linear sizes by some factor $\varphi(n)$ but to make some nonlinear transformation as taking $\log_n$.

We should like to introduce here three examples of formulation of LLDP. We will consider three different distributions on the set of Young diagrams of weight n and we will consider L^1 spaces. In L^1 metrics the exponential compactness is not valid and as we mentioned above the global LDP is not valid in this case.

Let's make some definitions. We remind that the Young diagram of weight n is the diagram which we suppose is drawing in the positive quater of the plane, such, that its shape is the piecewise constant nonincreasing functions with integer lengths and heights of steps and area of this diagram is equal to n. Without loosing of generality we suppose that the shapes of Young diagrams are continuous from the right.

First consider the case, when the heights of the steps of the shapes take values from the set $\mathcal{A} \subset \{1, 2, \ldots\}$ and the lengths of the steps take values from the set $\mathcal{B} \subset \{1, 2, \ldots\}$. In other words we consider the case when the values of heights and lengths of the steps belong to some (in general different) subsets of natural numbers. Then for given n we consider the uniform distribution on the set of such diagrams of weight n. The proper scailing in this case is to divide the linear sizes of the diagrams by $\sqrt{n}$. Then the area of the diagrams become unit. We call the diagrams with the corresponding restrictions the $(\mathcal{A}, \mathcal{B})_n$ – diagrams. Let's define the set $C \subset L^1([0, \infty))$ of functions f such that for an arbitrary $f \in C$ the following conditions are true: $||f|| = 1$, for every $f \in C$ there exists $\hat{f} = f$ $a.s.$, such that $\hat{f} \geq 0$ and $\hat{f}$ is nonincreasing continuous from the right function.

Next we define the functional $L_1(z)$ and the value L_2 as follows

$$\begin{aligned} L_1(z) &= \sup_{\ell \leq \min\{1,z\}} \left[\inf_{h \in R^1} \left(hz + \ell \ln \left(\sum_{i \in \mathcal{A}} e^{-ih} \right) \right) \right. \\ &+ \left. \inf_{h \in R^1} \left(h + \ell \ln \left(\sum_{i \in \mathcal{B}} e^{-ih} \right) \right) \right] \\ L_2 &= 2 \int_0^\infty h_2(x) dx, \end{aligned} \tag{3}$$

where $h_2(x)$ satisfies the equality

$$\sum_{i \in \mathcal{A}} e^{-iCx} \sum_{i \in \mathcal{B}} e^{-ih_2} = 1$$

and in turn the constant C satisfies the relation

$$\int_0^\infty h_2(x) dx = C.$$

By the simple analytic considerations one can rewrite formula (3) as follows:

$$L_1(z) = zh^1(z) + h^2(z),$$

where h^1, h^2 satisfy the equality

$$\sum_{i \in \mathcal{A}} e^{-h_1 i} \sum_{i \in \mathcal{B}} e^{-h_2 i} = 1$$

and $h^1(z)$ satisfies the relation

$$z = -\frac{dh^2(h^1)}{dh^1}.$$

We also consider one additional property of the functions $f \in C$. For an arbitary $0 \le x_1 < x_2 < \infty$ the following relation is valid

$$L_1\left(\frac{\hat{f}(x_1) - \hat{f}(x_2)}{x_2 - x_1}\right) > -\infty. \tag{4}$$

From the definition of the set C and (4) it follows that if $|\mathcal{A}| < \infty,\ |\mathcal{B}| < \infty$ then $\hat{f}$ is continuous and

$$\min_{i\in\mathcal{A},\ j\in\mathcal{B}} \frac{i}{j} \le \hat{f}' \le \max_{i\in\mathcal{A},\ j\in\mathcal{B}} \frac{i}{j}\ a.s.$$

The following theorem is valid.

Theorem 1 *For the sequence κ_n of the scaled shapes of the Young diagrams the LLDP is valid with the rate function N satisfying the relation*

$$N(y) = \begin{cases} L_2 - \int_0^\infty L_1(-\hat{y}'(x))dx, & y \in C, \\ \infty, & y \notin C. \end{cases} \tag{5}$$

Note 1. Values L_2 and $\int_0^\infty L_1(-\hat{y}'(x))dx$ are symmetric under changing $\mathcal{A} \leftrightarrow \mathcal{B}$ (it means only that we change the axis $OX \leftrightarrow OY$).

Note 2. In the case of absence of any restrictions i.e. $\mathcal{A} = \mathcal{B} = \{1, 2, \ldots\}$ formula (5) reduces to the following relation

$$N(y) = \begin{cases} \pi\sqrt{2/3} - \int_0^\infty H\left(\frac{-\hat{y}'}{1-\hat{y}'}\right)dx, & y \in C, \\ \infty, & y \notin C, \end{cases} \tag{6}$$

where $H(\xi) = \xi \ln \xi - (1 - \xi)\ln(1 - \xi)$ is the binary entropy. This result was published in [4] [5], see also [9], where it was obtained using some other method. The case when (6) is valid has transparent combinatorial structure. For given continuous function $y(x)$ we consider the partition of the interval $[a, b]$ into consecutive subintervals $[a_i, a_{i+1}]$ the maximal length of which tends to zero. Consider the spline

$$\ell(x) = y(a_i) + \frac{x - a_i}{a_{i+1} - a_i} y(a_{i+1}),\ x \in [a_i, a_{i+1}],$$

then $\ell(x) \to y(x)$ for every $x \in [a,b]$. and the number of scaled shapes κ_n such, that for all i

$$|\kappa_n(a_i) - y(a_i)| < \epsilon \tag{7}$$

is approximately equal to

$$\binom{\sqrt{n}(y(a_i) - y(b_i))}{\sqrt{n}(y(a_i) - y(a_{i+1}) + a_{i+1} - a_i) - 1} 2^{O(\epsilon)+o(\sqrt{n})}.$$

Then the rough logarithmic asymptotics of the number of the whole number of shapes of diagrams for which for all i the relation (7) is valid approximate the value

$$\frac{1}{\sqrt{n}} \ln \prod_i \binom{\sqrt{n}(y(a_i) - y(b_i))}{\sqrt{n}(y(a_i) - y(a_{i+1}) + a_{i+1} - a_i) - 1} \sim \int_a^b H\left(\frac{-y'}{1-y'}\right) dx + O(\epsilon) + o(1).$$

Then we put $a \to 0$; $b \to \infty$ and obtain the rough logarithmic asymptotics of the number of shapes κ_n which belong to the $\epsilon-$ neighborhood of the curve y. To prove the Theorem 1 it is necessary instead of the simple combinatorial arguments make the Cramer – type considerations about the probability of large deviations of the sum of independent random variables taking values in $\mathcal{B}$ and the numbers of these variables in the sum taking values from $1, 2, \dots N$ also randomly in such a way that if $\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_1}$ is such sequence, then $i_{j+1} - i_j \in \mathcal{A}$. The case when $\mathcal{A} = \{1, 2, \dots\}$ was considered in [7].

It is interesting using functional $N(y)$ and variational mathods find the unique curve $\hat{y}$ for which $N(y) = 0$. In the case of absence of the restrictions ($\mathcal{A} = \mathcal{B} = \{1, 2, \dots\}$) such a curve was known long time ago. It satisfies the relation

$$e^{-y\pi/\sqrt{6}} + e^{-x\pi/\sqrt{6}} = 1$$

and was obtained by using as we think less transparent method in the work [10]. In the general case of an arbitrary $\mathcal{A}$, $\mathcal{B}$ it is not possible to write the equation for extremal of $N(y)$ in such a compact form. It can be written in the parametrical form and we omit here detailes.

In the above example we introduce the case when the functional $N(y)$ actually depends on $\hat{y}'$ and scailing of the diagrams is uniform (we divide the linear sizes of the diagrams by factor $\sqrt{n}$). In the next example we apply the different scailing in OX and OY directions. Such different scailing used before. More essential is that the functional $N(y)$ in the following example depends only on y but not on $\hat{y}'$. This example is quite natural for those who knows from physics that in general the action functional $N(y)$ depends at the same time on y and $\hat{y}'$. It is so because the most evaluation in the probability is delivered by the only one shape near the curve y and this shape 'try' to be constant on the as large as possible intervals.

Now we introduce the distribution (so called Bell distribution) on the set of diagrams Y_n:

$$\mu(Y_n) = \frac{n!}{B_n \prod_{k=1}^{n} (k!)^{i_k} i_k!},$$

where i_k is the number of columns of length i_k in the diagram Y_n and

$$B_n = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}$$

is the Bell number. This distribution is generated by the mapping $\varphi : \{(k_1, k_2, \ldots, k_n)\} \to Y_n$ of the set of partitions of the set of n elements to the set of Young diagrams and the uniform distribution on the set of partitions $\{(k_1, k_2. \ldots, k_n)\}$. This distribution is rather sensitive on the choosing of the partition $(k_1, k_2, \ldots, k_n)$. This leads to the unusial for such kind of problems dependens of the rate function $N(y)$. Considering the same as before L^1 – metrics we establish the LLDP in this case. To achieve the proper results it is necessary to make the following scaling: we divide the linear sizes of the Young diagrams of weight n in the OX direction by $n/\ln n$ and in the OY direction by $\ln n$. Once more the scaled Young diagrams have unit area. As before we consider the special class of functions $C_1 \subset L^1([0, \infty))$: $y \in C_1$ iff there exists $\hat{y} = y$ $a.s.$ which is nonegative nonincreasing sunction and $||y|| = 1$. then the following theorem is valid.

Theorem 2 *For the sequence of probabilistic spaces, described above the LLDP is valid with $A_n = n$ and the rate function $N(y)$ satisfying the relation*

$$N(y) = \begin{cases} \int_0^\infty (1 + y \ln \hat{y} - y) dx, & y \in C_1, \\ \infty, & y \notin C_1. \end{cases}$$

In this problem functional $N(y)$ depends on y, but not on $\hat{y}'$. It is so because the real evaluation to the probability $P(\rho(y, \kappa_n) < \epsilon)$ is from the most 'heavy' shape whose behaviour is to be piecewise constant with as long as possible steps.

Note, that the unique curve $\hat{y}$, such that $N(y) = 0$ (up to the choosing values in the point of discontinuity) is the step function

$$\hat{y}(x) = \begin{cases} 1, & x \in [0,1), \\ 0, & x \notin [0,1). \end{cases}$$

Next problem we introduce here is the problem of establishing LLDP for the random shape of Young diagram with distribution

$$\mu_\theta(Y_n) = \frac{n!}{\theta(\theta+1)\ldots(\theta+n-1)\prod_{k=1}^n i_k! \left(\frac{k}{\theta}\right)^{i_k}}. \tag{8}$$

When $\theta = 1$ this distribution generated by the uniform distribution on the symmetric group S_n and mapping $\varphi : S_n \to (k_1, k_2, \ldots, k_n)$. We establish the proper LLDP in the case of slightly different distribution

$$\mu'_{\theta,n}(Y) = \prod_{k=1}^n \frac{e^{-1/k}\left(\frac{\theta}{k}\right)^{i_k}}{i_k!}. \tag{9}$$

In this case it is necessary to choose the special changing of scale. It is not right way here to divide the linear sizes of the diagrams by the fixed factor, but it is necessary to apply the logarithmic scale. Let's make some preliminary notes. Consider the sequence of Poisson random variables $\chi_{n^{t_1}}, \chi_{n^{t_1}+1}, \ldots, \chi_{n^{t_2}}$ with $E\chi_i = \frac{\theta}{i}$, $\theta > 0$. Then paying into account the Cramer – type considerations (see[11]) we can state the following large deviations lemma for the sum $\chi_{t_1 t_2} = \sum_{i=n^{t_1}}^{n_2^t} \chi_i / \ln n$

Lemma 1 *The following relation is valid:*

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{P(|\chi_{t_1 t_2} - a| < \epsilon)}{\ln n} = -(t_2 - t_1)\left(b \ln \frac{b}{\theta} + \theta - b\right), \tag{10}$$

where $b = a/(t_2 - t_1)$ and $a \geq 0$.

Omitting the detailes of the proof of this lemma we note that the expression in the rhs of relation (10) follows from the consideration that the moment generating function of the sum $\chi_{t_1 t_2}$ satisfies the following asymptotic relation

$$\ln E e^{h\chi_{t_1 t_2}} = (e^h - 1) \sum_{i=n^{t_1}}^{n^{t_2}} \frac{\theta}{i} \overset{n \to \infty}{\sim} (e^h - 1)(t_2 - t_1)\theta \ln n. \tag{11}$$

The rhs of (10) is the Legendre transform of the rhs of (11) (divided by $\ln n$).

Next as before we define the 'proper' set $C_2 \subset L^1([0,1])$ in the following way: $y \in C_2$ iff there exists $\hat{y} \geq 0$ such that $\hat{y} = y$ $a.s.$ and $\hat{y}$ is nonincreasing absolutely continuous bounded function. We will choose the following scailing: we make the transformation $x \rightarrow \log_n x$ in the OX direction and linear size in the OY direction we divide by $(\ln n)/\theta$. Also let's attract your attention to the following property:

$$\int_0^1 y(t)dn^t \sim \theta \frac{n}{\ln n}, \ n \rightarrow \infty. \tag{12}$$

The relation (12) is equivalent to the condition that the initial (nonscaled) Young diagram Y_n has the area equal to $n\theta$. It follows from the relation

$$\int_0^n \tilde{y}(x)dx = 1 \rightarrow \frac{1}{\theta}\int_0^1 y(t)dn^t \sim \frac{n}{\ln n}, \tag{13}$$

where $\tilde{y}$ is the function y after back scaling. The last relation is local. It means that it depends on behavour of function y only near the point $t = 1$, i.e.

$$\frac{1}{\theta}\int_{1-\alpha}^1 y(t)dn^t \sim \frac{n}{\ln n}.$$

From here it follows that $\hat{y}(1) = 0$ and if there exists the derivation $\hat{y}'(1)$, then $\hat{y}(t) = (\hat{y}'(1) + \beta)(t-1)$, where $\beta \overset{t \rightarrow 0}{\rightarrow} 0$ and from the relation

$$\frac{1}{\theta}\int_0^1 (\hat{y}'(1) + \beta)(t-1)n^t \ln n dt \sim \frac{n}{\ln n}$$

and the condition that the value β can be chosen an arbitrary small it follows that $\hat{y}'(1) = -\theta$. If there exists $\hat{y}'(1)$, then the conditions $y(1) = 0, \hat{y}'(1) = -\theta$ are equivalent to (13).

Next if there exists at least one point $x_o \in (0,1)$ such that $\hat{y}(x_o - 0) - \hat{y}(x_o + 0) = \gamma > 0$ (we set $\hat{y}(0) = \hat{y}(0+0)$), then

$$N(y) = \infty.$$

Indeed in this case for some $a_i, b_i, \ 0 < a_i < x_o < b_i < 1$

$$|\kappa_n(a_i) - \kappa_n(b_i)| > \gamma - 2\epsilon$$

and from (10) we have

$$
\begin{aligned}
&\lim_{\epsilon\to 0}\lim_{\to\infty}\frac{\ln P(\kappa_n:\kappa_n(a_i)=\chi,\ |\kappa_n(b_i)-\chi+\gamma|<\epsilon)}{\ln n}\\
&= -\tau\left(b\ln\frac{b}{\theta}+\theta-b\right)\stackrel{\tau\to 0}{\to}-\infty,\\
b &= \frac{\gamma}{\tau},\ \tau=b_i-a_i.
\end{aligned}
$$

Because the last expression is the upper bound for the value $-N(y)$ we have $N(y)=\infty$. Hence it is necessary to consider the function y such that $\hat{y}$ is continuous. Moreover it can be shown that if $\hat{y}$ is not absolutely continuous then also $N(y)=\infty$ (we omit detailes of the proof). At last if $\hat{y}=\infty$, then $N(y)=\infty$. So we propose for the functions $y\in C_2$ the condition $\hat{y}<\infty$.

Next theorem establish the LLDP for random Young diagram with the distribution (9).

Theorem 3 *There exists the following limit*

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\frac{P_{\theta,n}(\kappa_n\in B(y,\epsilon))}{\ln n}=-N(y),$$

where

$$N(y)=\begin{cases}\int_0^1\left(-\hat{y}'\ln\left(-\frac{\hat{y}'}{\theta}\right)+\theta+\hat{y}'\right)dt, & y\in C_2,\\ \infty, & y\notin C_2.\end{cases}$$

Probability $P_{\theta,n(1\pm\delta)}$ generated by the distribution (8) and string $n(1\pm\delta)$ means that actually we consider Young diagrams of weight not only n but the weights from the interval $[n(1-\delta),n(1+\delta)]$.

Note. We can rewrite

$$\int_0^1\left(-\hat{y}'\ln\left(-\frac{\hat{y}'}{\theta}\right)+\theta+\hat{y}'\right)dt=-\theta\left[h(\rho)+\frac{\hat{y}(0)}{\theta}-1\right],$$

where

$$h(\rho)=-\int_0^1\rho\ln\rho dt$$

and $\rho=-\hat{y}'/\theta$.

References

[1] J. Deuschel and D. Strook, Large Deviations. Academic Press, 1989.

[2] A. Dembo and O. Zeitouni, Large Deviations Techniques and Applications. Jones and Barlett Publ., Boston/London, 1992.

[3] V. Blinovsky and R. Dobrushin, Process Level Large Deviations for a Class of Piecewise Homogeneous Random Walks. The Dynkin Festschrift, Markov Processes and their Applications, Progress in Probability, Boston, Birkhouse, vol. 34, 1994, 1–60.

[4] V.Blinovsky, LDP for Shape of Random Young Diagram, in Proc. of IEEE Int. Symp. on Inf. Theory, 16 –21 Aug., Boston, 1998, p.416.

[5] V. Blinovsky, LDP of Random Young Diagram. Problemi Peredachi Informatsii, vol 34, N1, 1999, 61-74.

[6] V.Blinovsky, General Approach to LDP for Random Young Diagram. In Research Communications of the Conference in the memory of Paul Erdos, Budapest, July 4–11, 1999, 37–41.

[7] V.Blinovsky, Local LDP for the Random Young Diagram With Restrictions', in Proc. of the 1999 IEEE Information Theory and Communication Workshop, Kruger Park, South Africa, p.115

[8] V. Blinovsky, Large Deviations For the Shape of Random Young Diagram With Restrictions, in book: 'Numbers, Information and Complexity', Kluwer Acad. Publishers, 2000, pp/473–488.

[9] A. Dembo, A.Vershik, O. Zeitouni, Large Deviation Principle for Integer Partitions, manuscript.

[10] A.Vershik Statistical Sum Connecting with Young Diagrams, Zapiski Nauchnogo Ceminara LOMI, 164, Nauka, 1987, 20–29 (in Russian)

[11] H. Cramer, Sur un nouveau theoreme – limite de la theorie des probabilities, Actualites Scientifiques et Industrielles 736, 1938, 5-23. Colloque consacre a la theorie des probabilities, vol. 3, Hermann, Paris.