Superexponential decay for the GEM process

O. Zeitouni *
Department of Electrical Engineering
Technion—Israel Institute of Technology
Haifa 32000, Israel

November 18, 1996

AMS Subject classification: Primary 60F10. Secondary 60G09, 60K35.
Keywords: Large deviations. GEM process.

Abstract We show that the GEM process has strong ordering properties: the probability that the k-th largest element in the GEM sequence is beyond the first ck elements ($c > 1$) decays super-exponentially in k.

Let $\{U_i\}_{i=1}^{\infty}$ denote a sequence of $[0,1]$ valued i.i.d. random variables, with common law μ possessing a density $p_\theta(x) = \theta x^{\theta-1}$. Here, $\theta > 0$ is a fixed known parameter, and throughout we use $\overline{U}_i = 1 - U_i$.

Define the random sequence (GEM process) $A_1 = U_1$ and

$$A_i = U_i \prod_{j=1}^{i-1} \overline{U}_j \,, i \geq 2 \,.$$

For references and background on the GEM process and its properties, see [2]. Note that stochastically, A_i dominates A_{i+1}, but of course it is still possible that $A_i < A_{i+1}$. Our goal here is to estimate how unlikely is really this reverse inequality. More precisely, let $\{X_i\}$ denote the reordered sequence of $\{A_i\}$. That is, for each i there is a $j = j(i)$ such that $X_i = A_j$ and $X_{i+1} < X_i$. For $c > 1$, define the event

$$\Omega_{k,c} = \{X_k \text{ is not among } A_i \,, i < ck\} \,,$$

and let $P_{\theta,c,k} = \text{Prob}(\Omega_{k,c})$. Our goal is to prove the

Theorem 1.

$$\lim_{k\to\infty} \frac{\log P_{\theta,c,k}}{k \log k} = -\theta(c-1) \,.$$

Proof: We begin by quickly demonstrating a lower bound (which, incidentally, captures the correct order of magnitude but does not exhibit necessarily the most likely event).

*This work was partially supported by a US-Israel BSF grant.

Fix $\alpha > 0$ independent of k, and denote by $\Omega'_{k,c}$ the event

$$\Omega'_{k,c} = \{U_{ck} > \frac{1}{2}, U_j < \frac{\alpha}{(c-1)k}\,,\ j = k, k+1, \ldots, ck-1\}\,.$$

Because, in the event $\Omega'_{k,c}$,

$$A_{ck} \geq \frac{1}{2}\left(1 - \frac{\alpha}{(c-1)k}\right)^{(c-1)k} \prod_{i=1}^{k-1} \overline{U}_i \geq \frac{1}{2} e^{-2\alpha} \prod_{i=1}^{k-1} \overline{U}_i\,,$$

while

$$A_j \leq \frac{2\alpha}{(c-1)k} \prod_{i=1}^{k-1} \overline{U}_i\,,\ \ j = k, \ldots, ck-1\,,$$

it holds that for all k large enough, $A_{ck} \geq A_j\,,\ j = k, \ldots, ck-1$. Hence, for such k, $\Omega'_{k,c} \subset \Omega_{k,c}$. Thus, since for some constant $c_{\alpha,c}$ independent of k which may change from line to line,

$$\text{Prob}(U_1 < \frac{\alpha}{(c-1)k}) > c_{\alpha,c} k^{-\theta}\,,$$

it holds that

$$P_{\theta,c,k} \geq P(U_{ck} > \frac{1}{2}) c_{\alpha,c}^k k^{-\theta(c-1)k}\,,$$

which is more than enough to imply the required lower bound.

We next turn to establish the (harder) complementary upper bound. Note first that

$$\begin{aligned} P_{\theta,c,k} &= \text{Prob}(\exists j \geq ck\,,\ \ A_j \geq X_k) \\ &\leq \sum_{j=ck}^{\infty} \text{Prob}\left(A_j \geq X_k\right) \\ &\leq \sum_{j=ck}^{\infty} \text{Prob}\left(\text{for some } I \in \mathcal{I}_{j,k}, A_j \geq A_i\ \forall i \in I\right), \end{aligned} \tag{1}$$

where in the last inequality,

$$\mathcal{I}_{j,k} = \{\text{all subsets of length } j-k \text{ of } \{1, \ldots, j-1\}\}\,.$$

Note that the cardinality of $\mathcal{I}_{j,k}$ is $\left(\begin{array}{c} j \\ k \end{array}\right)$, while, from the definition of A_i and the i.i.d. assumption,

$$\max_{I \in \mathcal{I}_{j,k}} \text{Prob}\left(A_j \geq A_i\ \forall i \in I\right) \leq \text{Prob}\left(A_j \geq A_i, i = k, \ldots, j-1\right).$$

It thus follows from (1) that

$$\begin{aligned} P_{\theta,c,k} &\leq \sum_{j=ck}^{\infty} \left(\begin{array}{c} j \\ k \end{array}\right) \text{Prob}\left(A_j \geq A_i\,,\ i = k, \ldots, j-1\right) \\ &\leq \sum_{j=ck}^{\infty} \left(\begin{array}{c} j \\ k \end{array}\right) \text{Prob}\left(A_{j-k} \geq A_i\,,\ i = 1, \ldots, j-k-1\right) \\ &\leq \sum_{j=ck}^{\infty} \left(\begin{array}{c} j \\ k \end{array}\right) \text{Prob}\left(U_{j-k} \prod_{\ell=i+1}^{j-k-1} \overline{U}_\ell \geq \frac{U_i}{\overline{U}_i}\,,\ i = 1, \ldots, j-k-2\right) \triangleq \sum_{j=ck}^{\infty} \left(\begin{array}{c} j \\ k \end{array}\right) P_{j,k} \end{aligned} \tag{2}$$

Since for $j \geq ck$ there exists a $c_{\alpha,c}$ independent of j, k such that $\begin{pmatrix} j \\ k \end{pmatrix} \leq e^{c_{\alpha,c} k \log(j/k)}$, the proof is completed by the following lemma:

Lemma 1. *There exists a constant $c_{\theta,c}$, independent of k, j, such that for $j > ck$,*

$$P_{j,k} \leq c_{\theta,c} e^{-\theta(j-k)\log k} . \tag{3}$$

Proof of Lemma 1: Throughout this proof, we use c_α to denote constants, whose values may change from line to line, which are independent of k, j but may depend on θ, c. Let $n = j - k$. Then

$$P_{j,k} \leq \text{Prob}\,(\forall\ 2 \leq \ell \leq n\,,\ \sum_{j=1}^{\ell-1} \log \overline{U}_j \geq \log V_\ell) \triangleq P_n\,,$$

where $V_\ell = U_\ell / \overline{U}_\ell$.

For simplicity in notations, we assume below that both $\log n$ and $n/\log n$ are integers, the general case posing no new difficulties. Define

$$A_i = \frac{\sum_{j=i\log n}^{(i+1)\log n - 2} \log \overline{U}_j}{\log n}\,, \quad Z_i = \log V_{(i+1)\log n - 1}\,,$$

and let $\mathbf{x} \in \mathbb{R}^{n/\log n}$ have components x_i. Further, let

$$\mathcal{A}_n = \{\mathbf{x} \in \mathbb{R}^{n/\log n} : 0 > x_i > -n(\log n + 1), x_i = -jn^{-2}, \text{some integer } j\}\,.$$

Note that the cardinality of $\mathcal{A}_n$ is bounded by $(n^2(n+1)\log n)^{n/\log n} \leq e^{c_\alpha n}$. Then,

$$\begin{aligned} P_n &\leq \text{Prob}\,(\sum_{i=1}^{j} A_i \geq Z_j/\log n, j = 1, \ldots, n/\log n) \\ &\leq \frac{n}{\log n}\text{Prob}\,(Z_1 < -n\log^2 n/2) \\ &\quad + \sum_{\mathbf{x}\in\mathcal{A}_n} \text{Prob}\,(A_i \in [x_i, x_i + n^{-2}], \sum_{\ell=1}^{i} x_\ell + in^{-2} \geq \frac{Z_i}{\log n}, i = 1, \ldots, n/\log n)\,. \end{aligned}$$

Since $\text{Prob}\,(Z_1 < -n\log^2 n/2) \leq e^{-c_\alpha n \log^2 n}$, the bound on the cardinality of $\mathcal{A}_n$ and the independence of the $\{A_i\}$ and $\{Z_i\}$ reveals that for some C_n, C'_n with

$$\log(C_n)/n\log n \to -\infty\,, \log(C'_n)/n\log n \to 0\,, \tag{4}$$

$$P_n \leq C_n + C'_n \max_{\mathbf{x}\in\mathcal{A}_n} \prod_{i=1}^{n/\log n} \text{Prob}\,(x_i + n^{-2} \geq A_i \geq x_i) \prod_{i=1}^{n/\log n} \text{Prob}\,(Z_i \leq \log n \sum_{j=1}^{i} x_j + \frac{1}{n})\,. \tag{5}$$

Define next

$$\Lambda_\theta(\lambda) = \log\left(\int (1-x)^\lambda p_\theta(x)dx\right)\,,$$

and its Fenchel-Legendre transform

$$\Lambda^*_\theta(x) = \sup_{\lambda\in\mathbb{R}} (\lambda x - \Lambda(\lambda))\,.$$

Finally, let $\overline{\Lambda_\theta^*(x)} = \min_{y\in[x,x+n^{-2}]} \Lambda_\theta^*(y)$. By Cramér's theorem (see, e.g., [1, pg. 27]),

$$\text{Prob}\,(x_i + n^{-2} \geq A_i \geq x_i) \leq 2e^{-\log n \overline{\Lambda_\theta^*(x_i)}} .$$

On the other hand,

$$\text{Prob}\,(Z_i \leq \log n \sum_{j=1}^{i} x_j + n^{-1}) \leq c_\alpha e^{\theta \log n \sum_{j=1}^{i} x_j} .$$

Combining the above, and still using C_n, C_n' to denote (possibly different) constants still satisfying (4), one obtains

$$\begin{aligned} P_n &\leq C_n + C_n' \max_{\mathbf{x}\in A_n} \exp\left(-\log n (\sum_{i=1}^{n/\log n} \overline{\Lambda_\theta^*}(x_i) - \theta \sum_{i=1}^{n/\log n} \sum_{j=1}^{i} x_j) \right) \\ &\leq C_n + C_n' \max_{\phi\in\mathcal{A}} \exp -n \left(\int_0^1 (\overline{\Lambda_\theta^*}(\dot\phi_s) - \theta \frac{n}{\log n}\phi_s) ds \right) \\ &= C_n + C_n' \max_{\phi\in\mathcal{A}} \exp -n \left(\int_0^1 (\overline{\Lambda_\theta^*}(\frac{\log n \dot\phi_s}{n}) - \theta \phi_s) ds \right) \\ &\leq C_n + C_n' \max_{\phi\in\mathcal{A}} \exp -n \left(\int_0^1 (\Lambda_\theta^*(\frac{\log n \dot\phi_s}{n}) - \theta \phi_s) ds \right) \triangleq C_n + C_n' \max_{\phi\in\mathcal{A}} \exp -n I_n(\phi) , \end{aligned} \tag{6}$$

where

$$\mathcal{A} = \{\phi \text{ absolutely continuous, nonincreasing}\,, \;\; \phi_0 = 0\} ,$$

the second inequality is obtained by noting that polygonal decreasing functions (at steps of size $\log n/n$) form a subset of $\mathcal{A}$, and the last one by the continuity of Λ_θ^* away from 0 and a change in the value of C_n'.

Let next $\eta \in (0,1)$ be arbitrary. Using the convexity of Λ_θ^*, one notes that for $\phi \in \mathcal{A}$,

$$I_n(\phi) \geq \eta \Lambda_\theta^*(\frac{\phi_\eta \log n}{n\eta}) + (1-\eta)\Lambda_\theta^*(\frac{(\phi_1 - \phi_\eta)\log n}{(1-\eta)n}) - \theta \int_0^1 \phi_s ds .$$

Fixing η and ϕ_η, recalling that $\Lambda_\theta^* \geq 0$ and that ϕ is non-increasing,

$$\min_{\phi\in\mathcal{A}} I_n(\phi) \geq \min_{\phi_\eta < 0} \left(\eta\Lambda_\theta^*(\frac{\phi_\eta \log n}{n\eta}) - \theta(1-\eta)\phi_\eta \right) .$$

In Lemma 2 below, we collect some properties of $\Lambda_\theta^*(\cdot)$. In particular, it holds that $\Lambda_\theta^*(x) \geq -\theta \log x(1+o(1))$ for x small. A direct optimization over ϕ_η reveals then that there exit negative constants $c_1(\eta), c_2(\eta)$ independent of n such that

$$\min_{\phi_\eta<0} \left(\eta\Lambda_\theta^*(\frac{\phi_\eta \log n}{\eta n}) - \theta(1-\eta)\phi_\eta \right) = \min_{c_2 \log n < \phi_\eta < c_1} \left(\eta\Lambda_\theta^*(\frac{\phi_\eta \log n}{\eta n}) - \theta(1-\eta)\phi_\eta \right) \geq \theta\eta \log n(1+o(1)) .$$

Taking now $\eta \to 1$ yields

$$\min_{\phi\in\mathcal{A}} I_n(\phi) \geq \theta \log n (1+o(1)) .$$

Substituting back in (6), this concludes the proof of the Lemma 1 and hence of Theorem 1. □

The following lemma was used in the course of the proof of Lemma 1.

Lemma 2. *Λ_θ^* is strictly convex, $\Lambda_\theta^*(x)=\infty$ for $x\geq 0$, $\lim_{x\to-\infty}\Lambda_\theta^*(x)=\infty$, and $\Lambda_\theta^*(y)=0$ if and only if $y=\int\log(1-x)p_\theta(x)dx$. Finally, $\Lambda_\theta^*(x)\geq-\theta\log(x)(1+o(1))$ for x small.*

Proof of Lemma 2: The first part of the lemma is a trivial consequence of the fact that $\Lambda_\theta(\lambda)<\infty$ for all λ with $|\lambda|<\lambda_0(\theta)$ (c.f. [1, pg. 28]). To see the second part, note first that for $\theta=1$ and $x<0$, U_1 is uniformly distributed and a straight forward computation reveals that $\Lambda_\theta(\lambda)=-\log(\lambda+1)$ for $\lambda>-1$ and $\Lambda_\theta^*(x)=-1-x-\log(-x)$. We use below c_θ to denote various constants, whose value may change from line to line but which are independent of λ. To see the claim for $0<\theta<1$, simply note that for $\lambda>1$,

$$\begin{aligned}\int_0^1 y^\lambda(1-y)^{\theta-1}dy &= \int_0^{1-\lambda^{-1}} y^\lambda(1-y)^{\theta-1}dy+\int_{1-\lambda^{-1}}^1 y^\lambda(1-y)^{\theta-1}dy\\ &\leq (1-\lambda^{-1})^\lambda\lambda^{1-\theta}\left(\theta\frac{1-\lambda^{-1}}{1+\lambda}+\lambda^{-1}\right)\leq c_\theta\lambda^{-\theta}\,,\end{aligned}$$

whereas for $\theta>1$ and $\lambda>0$,

$$\begin{aligned}\int_0^1 y^\lambda(1-y)^{\theta-1}dy &\leq \sum_{k=0}^\infty\int_{1-(k+1)\lambda^{-1}}^{1-k\lambda^{-1}} y^\lambda(1-y)^{\theta-1}dy\\ &\leq \sum_{k=0}^\infty \lambda^{-\theta}k^{\theta-1}e^{-k}\leq c_\theta\lambda^{-\theta}\,.\end{aligned}$$

Hence, for any $\theta>0$ and $\lambda>1$,

$$\Lambda_\theta(\lambda)\leq c_\theta-\theta\log\lambda\,.$$

It follows that, with the choice $\lambda=-x^{-1}$,

$$\Lambda_\theta^*(x)\geq -1-c_\theta-\theta\log(-x)\,,$$

as claimed. □

Remark: In fact, the exact form of p_θ was never used. In order to get Theorem 1, all that is needed is that the common law μ of the $(0,1)$ valued i.i.d. random variables U_i possesses a density near $0,1$ such that, for some positive constants θ,α_θ,

$$\lim_{x\to\infty}\frac{1}{x}\log\text{Prob}\left(\log(U_i/\overline{U}_i)<-x\right)=-\alpha_\theta\,, \tag{7}$$

$$\lim_{x\to 0^-}\frac{\Lambda^*(x)}{\log(-x)}=-\theta\,. \tag{8}$$

Here, $\Lambda^*(x)$ is the Fenchel-Legendre transform of

$$\Lambda(\theta)=\log\left(\int_0^1(1-x)^\lambda\mu(dx)\right)\,.$$

Acknowledgement: I thank Richard Arratia for asking me the question answered in this note, and for several useful discussions.

References

[1] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Jones and Bartlett, Boston, 1993.

[2] J. Pitman and M. Yor, *The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator*, to appear, Ann. Probab., 1996.