Negative Moments for Branching Process (An application of small value probability)

Wenbo V. Li

University of Delaware

http://www.math.udel.edu/~wli
wli@math.udel.edu

The 8th Workshop at BNU and Wuyishan, July 16-21, 2012

Small Value Probability

Small value (deviation) probability studies the asymptotic rate of approaching zero for rare events that positive random variables take smaller values. To be more precise, let V_n be a sequence of *non-negative* random variables and suppose that some or all of the probabilities

$$\mathbb{P}\left(V_n \leq \varepsilon_n\right), \quad \mathbb{P}\left(V_n \leq C\right), \quad \mathbb{P}\left(V_n \leq (1-\delta)\mathbb{E}\, V_n\right)$$

tend to zero as $n \to \infty$, for $\varepsilon_n \to 0$, some constant $C > 0$ and $0 < \delta \leq 1$. Of course, they are all special cases of $\mathbb{P}\left(V_n \leq h_n\right) \to 0$ for some function $h_n \geq 0$, but examples and applications given later show the benefits of the separated formulations.

•What is often an important and interesting problem is the determination of just how "rare" the event $\{V_n \leq h_n\}$ is, that is, the study of the *small value (deviation) probabilities* of V_n associated with the sequence h_n.

•If $\varepsilon_n = \varepsilon$ and $V_n = \|X\|$, the norm of a random element X on a separable Banach space, then we are in the setting of small ball/deviation probabilities.

Small Value Theory

A theory of small value phenomenon is being developed and centered on:

- systematically studies of the existing techniques and applications

- applications of the existing methods to a variety of fields

- new techniques and problems motivated by current interests of advancing knowledge such as smooth Gaussian fields, random matrices/polynomials.

♢ W.V. Li, Ten lectures on Small Value Probabilities: Theory and Applications, NSF/CBMS Regional Research Conference in the Mathematical Sciences, University of Alabama in Huntsville, June 04-08, 2012.
Lecture notes are available at http://www.math.uah.edu/~cbms/

Small Value Theory

A theory of small value phenomenon is being developed and centered on:

• systematically studies of the existing techniques and applications

• applications of the existing methods to a variety of fields

• new techniques and problems motivated by current interests of advancing knowledge such as smooth Gaussian fields, random matrices/polynomials.

◊ W.V. Li, Ten lectures on Small Value Probabilities: Theory and Applications, NSF/CBMS Regional Research Conference in the Mathematical Sciences, University of Alabama in Huntsville, June 04-08, 2012.
Lecture notes are available at http://www.math.uah.edu/~cbms/
◊ Mikhail Lifshits maintains an excellent updated bibliography (over 300 papers only on estimates) at
http://www.proba.jussieu.fr/pageperso/smalldev/

Deviations: Large vs Small

• Both are estimates of rare events and depend on one's point of view in certain problems. E.g $\mathbb{P}(V \leq \varepsilon) = \mathbb{P}(V^{-1} \geq \varepsilon^{-1})$.

• Large deviations deal with a class of sets rather than special sets. And results for special sets may not hold in general.

• Similar techniques can be used, such as exponential Chebychev's inequality, change of measure argument, isoperimetric inequalities, concentration of measure, chaining, etc.

• Second order behavior of certain large deviation estimates depends on small deviation type estimates.

• General theory for small deviations has been developed for Gaussian processes and measures.

Some technical difficulties between small and large values

- Let X and Y be two positive r.v's (not necessarily ind.). Then

$$\begin{aligned} \mathbb{P}(X+Y>t) &\geq \max(\mathbb{P}(X>t), \mathbb{P}(Y>t)) \\ \mathbb{P}(X+Y>t) &\leq \mathbb{P}(X>\delta t) + \mathbb{P}(Y>(1-\delta)t) \end{aligned}$$

but

$$?? \leq \mathbb{P}(X+Y \leq \varepsilon) \leq \min(\mathbb{P}(X \leq \varepsilon), \mathbb{P}(Y \leq \varepsilon))$$

Some technical difficulties between small and large values

- Let X and Y be two positive r.v's (not necessarily ind.). Then

$$\begin{aligned} \mathbb{P}(X+Y>t) &\geq \max(\mathbb{P}(X>t), \mathbb{P}(Y>t)) \\ \mathbb{P}(X+Y>t) &\leq \mathbb{P}(X>\delta t)+\mathbb{P}(Y>(1-\delta)t) \end{aligned}$$

but

$$?? \leq \mathbb{P}(X+Y \leq \varepsilon) \leq \min(\mathbb{P}(X \leq \varepsilon), \mathbb{P}(Y \leq \varepsilon))$$

- Moment estimates $a_n \leq \mathbb{E}\, X^n \leq b_n$ can be used for

$$\mathbb{E}\, e^{\lambda X} = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \mathbb{E}\, X^n$$

but $\mathbb{E}\, \exp\{-\lambda X\}$ is harder to estimate.

Some technical difficulties between small and large values

- Let X and Y be two positive r.v's (not necessarily ind.). Then

$$\begin{aligned}\mathbb{P}(X+Y>t) &\geq \max(\mathbb{P}(X>t),\mathbb{P}(Y>t))\\ \mathbb{P}(X+Y>t) &\leq \mathbb{P}(X>\delta t)+\mathbb{P}(Y>(1-\delta)t)\end{aligned}$$

but

$$?? \leq \mathbb{P}(X+Y\leq\varepsilon)\leq\min(\mathbb{P}(X\leq\varepsilon),\mathbb{P}(Y\leq\varepsilon))$$

- Moment estimates $a_n \leq \mathbb{E}\,X^n \leq b_n$ can be used for

$$\mathbb{E}\,e^{\lambda X}=\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}\mathbb{E}\,X^n$$

but $\mathbb{E}\exp\{-\lambda X\}$ is harder to estimate.

- Tauberian theorem: Let V be a positive random variable. Then for $\alpha>0$ and slowly varying function L,

$$\mathbb{E}\,e^{-\lambda V}\sim C/(\lambda^{\alpha}L(\lambda)) \quad \text{as} \quad \lambda\to\infty.$$

if and only if

$$\mathbb{P}(V\leq t)\sim\frac{C}{\Gamma(1+\alpha)}t^{\alpha}/L(1/t) \quad \text{as} \quad t\to 0.$$

EX: Related formulations for BM

•For one-dim Brownian motion $B(t)$ under the sup-norm, we have by scaling

$$\log \mathbb{P}\left(\sup_{0\le t\le 1}|B(t)|\le\varepsilon\right) = \log \mathbb{P}\left(\sup_{0\le t\le T}|B(t)|\le 1\right) = \log \mathbb{P}\left(\tau_2 \ge T\right)$$

$$\sim -\frac{\pi^2}{8}\cdot T \sim -\frac{\pi^2}{8}\frac{1}{\varepsilon^2}$$

as $\varepsilon \to 0$ and $T = \varepsilon^{-2} \to \infty$. Here $\tau_2 = \inf\{s : |B(s)| \ge 1\}$ is the first two-sided exit (or passage) time.

•Lower tail and one sided exit time:

$$\mathbb{P}\left(\sup_{0\le t\le 1} B(t)\le\varepsilon\right) = \mathbb{P}\left(\sup_{0\le t\le T} B(t)\le 1\right) = \mathbb{P}\left(\tau_1 > T\right)$$

$$= \mathbb{P}\left(|B(T)|\le 1\right) \sim (2/\pi)^{1/2}\, T^{-1/2} = (2/\pi)^{1/2}\varepsilon$$

where $\tau_1 = \inf\{s : B(s) = 1\}$ is the one-sided exit time.

Negative Moments: Smoothness via Malliavin Matrix

Consider $F = (F^1, \cdots, F^m) : \Omega \to \mathbb{R}^m$ with $F^i \in D^{1,2}$. Then Malliavin Matrix of F is

$$\gamma_F = (\gamma_F^{ij}), \quad \gamma_F^{ij} = \langle DF^i, DF^j \rangle$$

Thm: (Malliavin) If (1) $F^i \in D^\infty$ and (2) $\mathbb{E}\,|\det \gamma_F|^{-p} < \infty$ for any $p > 0$, then F has a C^∞ density.

•The condition (2) is called non-degeneracy for F.

•All these have been extended into theory of SDE and SPDE. It is curial to check the non-degeneracy condition which is small value probability. In fact, the negative moments estimates

$$\mathbb{E}\, V^{-p} < \infty \quad \text{for any/all} \quad p > 0$$

is equivalent to the upper small value estimates

$$\mathbb{P}(V \le \varepsilon) \le C_p \varepsilon^p \quad \text{for any/all} \quad p > 0, \quad \text{as} \quad \varepsilon \to 0.$$

•Mueller and Nualart (2008): Regularity of the density for the stochastic heat equation.

•Nualart (2010, book): Malliavin Calculus and its Applications.

SVP for the Martingale Limit of a Galton-Watson Tree

Consider the Galton-Watson branching process $(Z_n)_{n\geq 0}$ with offspring distribution $(p_k)_{k\geq 0}$ starting with $Z_0 = 1$. In any subsequent generation individuals independently produce a random number of offspring according to $\mathbb{P}(X = k) = p_k$. Suppose $m = \mathbb{E}\, X > 1$ and $\mathbb{E}\, X \log X < \infty$. Then by Kesten-Stigum theorem, the martingale limit (a.s and in L^1)

$$W = \lim_{n\to\infty} \frac{Z_n}{m^n}$$

exists and is nontrivial almost surely with $\mathbb{E}\, W = 1$. WOLG, assume $p_0 = 0$ and $p_k < 1$ for all $k \geq 1$. Then in the case $p_1 > 0$, there exist constants $0 < c < C < \infty$ such that for all $0 < \varepsilon < 1$

$$c\varepsilon^{\tau} \leq \mathbb{P}(W \leq \varepsilon) \leq C\varepsilon^{\tau}, \quad \tau = -\log p_1 / \log m$$

and in the case $p_1 = 0$, there exist constants $0 < c < C < \infty$ such that for all $0 < \varepsilon < 1$

$$c\varepsilon^{-\alpha/(1-\alpha)} \leq -\log \mathbb{P}(W \leq \varepsilon) \leq C\varepsilon^{-\alpha/(1-\alpha)}.$$

with $\nu = \min\{k \geq 2 : p_k \neq 0\}$ and $\alpha = \log \nu / \log m < 1$.

•These results are due to Dubuc (1971a,b) in the $p_1 > 0$ case, and up to a Tauberian theorem also in the $p_1 = 0$ case, see Bingham (1988). The proofs are relying on nontrivial complex analysis and are therefore difficult to generalize, for example to processes with immigration and/or dependent offsprings.

•Examples, near-constancy phenomena and various refinements, see Harris (1948), Karlin and McGregor (1968 a,b), Dubuc (1982), Barlow and Perkins (1987), Goldstein (1987) and Kusuoka (1987), Bingham (1988), Biggins and Bingham (1991), Biggins and Bingham (1993), Biggins and Nadarajah (1994), Hambly (1995), Fleischman and Wachtel (2007, 2009).

•A probabilistic argument is given in Mörters and Ortgiese (2008).

•These results are due to Dubuc (1971a,b) in the $p_1 > 0$ case, and up to a Tauberian theorem also in the $p_1 = 0$ case, see Bingham (1988). The proofs are relying on nontrivial complex analysis and are therefore difficult to generalize, for example to processes with immigration and/or dependent offsprings.

•Examples, near-constancy phenomena and various refinements, see Harris (1948), Karlin and McGregor (1968 a,b), Dubuc (1982), Barlow and Perkins (1987), Goldstein (1987) and Kusuoka (1987), Bingham (1988), Biggins and Bingham (1991), Biggins and Bingham (1993), Biggins and Nadarajah (1994), Hambly (1995), Fleischman and Wachtel (2007, 2009).

•A probabilistic argument is given in Mörters and Ortgiese (2008).

•Fleischmann and Wachtel (2009): For $p_1 > 0$ and $x_n = o(m^n)$,

$$\mathbb{P}(Z_n \le x_n) \sim \mathbb{P}(W \le x_n m^{-n})$$

•Probabilistic arguments are under development for related problems (up to constants).

•These results are due to Dubuc (1971a,b) in the $p_1 > 0$ case, and up to a Tauberian theorem also in the $p_1 = 0$ case, see Bingham (1988). The proofs are relying on nontrivial complex analysis and are therefore difficult to generalize, for example to processes with immigration and/or dependent offsprings.
•Examples, near-constancy phenomena and various refinements, see Harris (1948), Karlin and McGregor (1968 a,b), Dubuc (1982), Barlow and Perkins (1987), Goldstein (1987) and Kusuoka (1987), Bingham (1988), Biggins and Bingham (1991), Biggins and Bingham (1993), Biggins and Nadarajah (1994), Hambly (1995), Fleischman and Wachtel (2007, 2009).
•A probabilistic argument is given in Mörters and Ortgiese (2008).
•Fleischmann and Wachtel (2009): For $p_1 > 0$ and $x_n = o(m^n)$,

$$\mathbb{P}(Z_n \leq x_n) \sim \mathbb{P}(W \leq x_n m^{-n})$$

•Probabilistic arguments are under development for related problems (up to constants).
•SVP for supercritical branching processes with Immigration is studied in Chu, Li and Ren (2012).

Negative Moments of Z_n vs Large Deviations of Z_{n+1}/Z_n

Negative moments of Z_n are closely related to large deviation of Z_{n+1}/Z_n which is the well-known Lotka-Nagaev estimator of the offspring mean.

Negative Moments of Z_n vs Large Deviations of Z_{n+1}/Z_n

Negative moments of Z_n are closely related to large deviation of Z_{n+1}/Z_n which is the well-known Lotka-Nagaev estimator of the offspring mean.

$$
\begin{aligned}
& \mathbb{P}(|Z_{n+1}Z_n^{-1} - m| \geq a) \\
= & \sum_{k\geq 1} \mathbb{P}\left(\frac{1}{Z_n} |\sum_{j=1}^{Z_n} (X_{n,j} - m)| \geq a \,\middle|\, Z_n = k \right) \cdot \mathbb{P}(Z_n = k) \\
= & \sum_{k\geq 1} \mathbb{P}\left(|\sum_{j=1}^{k} (X_{n,j} - m)| \geq ka \right) \cdot \mathbb{P}(Z_n = k)
\end{aligned}
$$

Negative Moments of Z_n vs Large Deviations of Z_{n+1}/Z_n

Negative moments of Z_n are closely related to large deviation of Z_{n+1}/Z_n which is the well-known Lotka-Nagaev estimator of the offspring mean.

$$
\begin{aligned}
& \mathbb{P}(|Z_{n+1}Z_n^{-1} - m| \geq a) \\
= & \sum_{k\geq 1} \mathbb{P}\left(\frac{1}{Z_n}|\sum_{j=1}^{Z_n}(X_{n,j} - m)| \geq a \,\Big| Z_n = k\right) \cdot \mathbb{P}(Z_n = k) \\
= & \sum_{k\geq 1} \mathbb{P}\left(|\sum_{j=1}^{k}(X_{n,j} - m)| \geq ka\right) \cdot \mathbb{P}(Z_n = k) \\
\simeq & \sum_{k\geq 1} k\mathbb{P}\left(|X - m| \geq ka\right) \cdot \mathbb{P}(Z_n = k) \quad \text{by Heyde's thm}
\end{aligned}
$$

Negative Moments of Z_n vs Large Deviations of Z_{n+1}/Z_n

Negative moments of Z_n are closely related to large deviation of Z_{n+1}/Z_n which is the well-known Lotka-Nagaev estimator of the offspring mean.

$$
\begin{aligned}
& \mathbb{P}(|Z_{n+1}Z_n^{-1} - m| \geq a) \\
= & \sum_{k\geq 1} \mathbb{P}\left(\frac{1}{Z_n}|\sum_{j=1}^{Z_n}(X_{n,j} - m)| \geq a \,\Big|Z_n = k\right) \cdot \mathbb{P}(Z_n = k) \\
= & \sum_{k\geq 1} \mathbb{P}\left(|\sum_{j=1}^{k}(X_{n,j} - m)| \geq ka\right) \cdot \mathbb{P}(Z_n = k) \\
\simeq & \sum_{k\geq 1} k\mathbb{P}\left(|X - m| \geq ka\right) \cdot \mathbb{P}(Z_n = k) \quad \text{by Heyde's thm} \\
\simeq & \sum_{k\geq 1} k^{-\beta} h(k) \cdot \mathbb{P}(Z_n = k) = \mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right)
\end{aligned}
$$

where h is a slowly varying function and $\beta > 0$.

•Ney and Vidyashankar (2003): Assume $1 < m < \infty$ and $\mathbb{E}\, X \log X < \infty$. Then

$$\lim_{n\to\infty} A_n(\beta)\mathbb{E}\, Z_n^{-\beta} = C(\beta), \quad 0 < C(\beta) < \infty,$$

where

$$A_n(\beta) = \begin{cases} p_1^{-n} & \text{if} \quad p_1 m^\beta > 1, \\ n^{-1} p_1^{-n} & \text{if} \quad p_1 m^\beta = 1, \\ m^{\beta n} & \text{if} \quad p_1 m^\beta < 1, \end{cases}$$

•Ney and Vidyashankar (2003): Assume $1 < m < \infty$ and $\mathbb{E}\, X \log X < \infty$. Then

$$\lim_{n\to\infty} A_n(\beta)\mathbb{E}\, Z_n^{-\beta} = C(\beta), \quad 0 < C(\beta) < \infty,$$

where

$$A_n(\beta) = \begin{cases} p_1^{-n} & \text{if} \quad p_1 m^\beta > 1, \\ n^{-1} p_1^{-n} & \text{if} \quad p_1 m^\beta = 1, \\ m^{\beta n} & \text{if} \quad p_1 m^\beta < 1, \end{cases}$$

Idea of Pf:

$$\mathbb{E}\, Z_n^{-\beta} = \frac{1}{\Gamma(\beta)} \int_0^\infty \mathbb{E}\, e^{-tZ_n} t^{\beta-1} dt = \frac{1}{\Gamma(\beta)} \int_0^\infty f_n(e^{-t}) t^{\beta-1} dt$$

where $f_n(\cdot)$ is the n-th iterate of $f(t) = \mathbb{E}\, t^X$.

•Kuelbs, Li and Vidyashankar (2012+). Assume $1 < m < \infty$, $\mathbb{E}\, X \log X < \infty$ and $h(x)$ is a positive slowly varying function with certain monotonicity.
(1). If $p_1 m^\beta < 1$, then

$$\lim_{n\to\infty} \frac{m^{\beta n}}{h(m^n)} \mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) = C(\beta).$$

(2). If $p_1 m^\beta > 1$, then

$$\lim_{n\to\infty} p_1^{-n} \mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) = \sum_{k=1}^{\infty} \frac{h(k)}{k^\beta} q_k$$

where $q_k = \lim_{n\to\infty} p_1^k \mathbb{P}(Z_n = k)$ and $0 < \sum_{k=1}^{\infty} k^{-\beta} h(k) q_k < \infty$.
(3). If $p_1 m^\beta = 1$, then

$$\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) \approx p_1^n \sum_{k \le m^n} \frac{h(k)}{k}$$

•Applications to limit theorems for multiple generations $(R_n, R_{n-1}, \cdots R_{n-r(n)})$ with $R_n = Z_{n+1}/Z_n$ are given, see also Kuelbs and Vidyashankar (2008).

Idea of Proof: $p_1 m^\beta < 1$

Upper bound: Holder and $Z_n/m^n \to^d W$.

$$\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) \le \left(\mathbb{E}\left(Z_n^{-\beta(1+\delta)}\right)\right)^{1/(1+\delta)} \cdot \left(\mathbb{E}\left(h(Z_n)^{(1+\delta)/\delta}\right)\right)^{\delta/(1+\delta)}$$

for $\delta > 0$ small enough such that $p_1 m^{\beta(1+\delta)} < 1$.
Lower bound: Truncation and Holder's inequality. For $h(x)$ decreasing,

$$\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) \ \ge \ \mathbb{E}\left(Z_n^{-\beta} h(Z_n) \mathbb{I}(Z_n \ge t_n)\right)$$

Idea of Proof: $p_1 m^\beta < 1$

Upper bound: Holder and $Z_n/m^n \to^d W$.

$$\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) \le \left(\mathbb{E}\left({Z_n}^{-\beta(1+\delta)}\right)\right)^{1/(1+\delta)} \cdot \left(\mathbb{E}\left(h(Z_n)^{(1+\delta)/\delta}\right)\right)^{\delta/(1+\delta)}$$

for $\delta > 0$ small enough such that $p_1 m^{\beta(1+\delta)} < 1$.
Lower bound: Truncation and Holder's inequality. For $h(x)$ decreasing,

$$\begin{aligned}\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) &\ge \mathbb{E}\left(Z_n^{-\beta} h(Z_n)\mathbb{I}(Z_n \ge t_n)\right)\\ &\ge h(t_n)\cdot \mathbb{E}\left(Z_n^{-\beta}\mathbb{I}(Z_n \ge t_n)\right)\end{aligned}$$

Idea of Proof: $p_1 m^\beta < 1$

Upper bound: Holder and $Z_n/m^n \to^d W$.

$$\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) \le \left(\mathbb{E}\left(Z_n^{-\beta(1+\delta)}\right)\right)^{1/(1+\delta)} \cdot \left(\mathbb{E}\left(h(Z_n)^{(1+\delta)/\delta}\right)\right)^{\delta/(1+\delta)}$$

for $\delta > 0$ small enough such that $p_1 m^{\beta(1+\delta)} < 1$.
Lower bound: Truncation and Holder's inequality. For $h(x)$ decreasing,

$$\begin{aligned}
\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) &\ge \mathbb{E}\left(Z_n^{-\beta} h(Z_n)\mathbb{I}(Z_n \ge t_n)\right) \\
&\ge h(t_n) \cdot \mathbb{E}\left(Z_n^{-\beta}\mathbb{I}(Z_n \ge t_n)\right) \\
&= h(t_n) \cdot \left(\mathbb{E}\, Z_n^{-\beta} - \mathbb{E}\left(Z_n^{-\beta}\mathbb{I}(Z_n \ge t_n)\right)\right)
\end{aligned}$$

where $t_n \approx m^n$. The last term can be controled just like the upper bound estimates.

Idea of Proof: $p_1 m^\beta = 1$

Note that for $k \le m^n$,

$$\mathbb{P}(Z_n \le k) \approx (km^{-n})^{-(\log p_1)/\log m} = p_1^n k^\beta$$

Thus by using summation by parts twice,

$$\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) \approx \sum_{k \le m^n} \frac{h(k)}{k^\beta} \cdot \mathbb{P}(Z_n = k)$$

Idea of Proof: $p_1 m^\beta = 1$

Note that for $k \le m^n$,

$$\mathbb{P}(Z_n \le k) \approx (km^{-n})^{-(\log p_1)/\log m} = p_1^n k^\beta$$

Thus by using summation by parts twice,

$$\begin{aligned} \mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) &\approx \sum_{k \le m^n} \frac{h(k)}{k^\beta} \cdot \mathbb{P}(Z_n = k) \\ &\approx \sum_{k \le m^n} \left(\frac{h(k)}{k^\beta} - \frac{h(k+1)}{(k+1)^\beta} \right) \cdot \mathbb{P}(Z_n \le k) \end{aligned}$$

Idea of Proof: $p_1 m^\beta = 1$

Note that for $k \leq m^n$,

$$\mathbb{P}(Z_n \leq k) \approx (km^{-n})^{-(\log p_1)/\log m} = p_1^n k^\beta$$

Thus by using summation by parts twice,

$$\begin{aligned}
\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) &\approx \sum_{k \leq m^n} \frac{h(k)}{k^\beta} \cdot \mathbb{P}(Z_n = k) \\
&\approx \sum_{k \leq m^n} \left(\frac{h(k)}{k^\beta} - \frac{h(k+1)}{(k+1)^\beta}\right) \cdot \mathbb{P}(Z_n \leq k) \\
&\approx \sum_{k \leq m^n} \left(\frac{h(k)}{k^\beta} - \frac{h(k+1)}{(k+1)^\beta}\right) \cdot p_1^n k^\beta
\end{aligned}$$

Idea of Proof: $p_1 m^\beta = 1$

Note that for $k \le m^n$,

$$\mathbb{P}(Z_n \le k) \approx (km^{-n})^{-(\log p_1)/\log m} = p_1^n k^\beta$$

Thus by using summation by parts twice,

$$\begin{aligned}
\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) &\approx \sum_{k \le m^n} \frac{h(k)}{k^\beta} \cdot \mathbb{P}(Z_n = k) \\
&\approx \sum_{k \le m^n} \left(\frac{h(k)}{k^\beta} - \frac{h(k+1)}{(k+1)^\beta}\right) \cdot \mathbb{P}(Z_n \le k) \\
&\approx \sum_{k \le m^n} \left(\frac{h(k)}{k^\beta} - \frac{h(k+1)}{(k+1)^\beta}\right) \cdot p_1^n k^\beta \\
&\approx p_1^n \sum_{k \le m^n} \left((k+1)^\beta - k^\beta\right) \cdot \frac{h(k)}{k^\beta}
\end{aligned}$$

Idea of Proof: $p_1 m^\beta = 1$

Note that for $k \le m^n$,

$$\mathbb{P}(Z_n \le k) \approx (km^{-n})^{-(\log p_1)/\log m} = p_1^n k^\beta$$

Thus by using summation by parts twice,

$$\begin{aligned}
\mathbb{E}\left(Z_n^{-\beta} h(Z_n)\right) &\approx \sum_{k \le m^n} \frac{h(k)}{k^\beta} \cdot \mathbb{P}(Z_n = k) \\
&\approx \sum_{k \le m^n} \left(\frac{h(k)}{k^\beta} - \frac{h(k+1)}{(k+1)^\beta}\right) \cdot \mathbb{P}(Z_n \le k) \\
&\approx \sum_{k \le m^n} \left(\frac{h(k)}{k^\beta} - \frac{h(k+1)}{(k+1)^\beta}\right) \cdot p_1^n k^\beta \\
&\approx p_1^n \sum_{k \le m^n} \left((k+1)^\beta - k^\beta\right) \cdot \frac{h(k)}{k^\beta} \\
&\approx p_1^n \sum_{k \le m^n} \frac{h(k)}{k}
\end{aligned}$$