Notes 15 : Branching processes

Math 733-734: Theory of Probability *Lecturer: Sebastien Roch*

References: [Wil91, Section 0], [Dur10, Section 5.3], [AN72, Section I.1-I.5].

1 Branching processes

1.1 Definitions

Recall:

DEF 15.1 *A* branching process *is an SP of the form:*

- *Let* $X(i, n)$, $i \geq 1$, $n \geq 1$, *be an array of iid* $\mathbb{Z}_+$*-valued RVs with finite mean* $m = \mathbb{E}[X(1, 1)] < +\infty$, *and inductively,*

$$Z_n = \sum_{1 \leq i \leq Z_{n-1}} X(i, n)$$

To avoid trivialities we assume $\mathbb{P}[X(1, 1) = i] < 1$ *for all* $i \geq 0$.

LEM 15.2 $M_n = m^{-n} Z_n$ *is a nonnegative MG.*

Proof: Use the following lemma (proved in homework):

LEM 15.3 *If* $Y_1 = Y_2$ *a.s. on* $B \in \mathcal{F}$ *then* $\mathbb{E}[Y_1 \,|\, \mathcal{F}] = \mathbb{E}[Y_2 \,|\, \mathcal{F}]$ *a.s. on* B.

Then, on $\{Z_{n-1} = k\}$,

$$\mathbb{E}[Z_n \,|\, \mathcal{F}_{n-1}] = \mathbb{E}[\sum_{1 \leq j \leq k} X(j, n) \,|\, \mathcal{F}_{n-1}] = mk = mZ_{n-1}.$$

This is true for all k. ■

COR 15.4 $M_n \to M_\infty < +\infty$ *a.s. and* $\mathbb{E}[M_\infty] \leq 1$.

The martingale convergence theorem in itself tells us little about the limit. Here we derive a more detailed picture of the limiting behavior—starting with extinction.

1.2 Extinction

Let $p_i = \mathbb{P}[X(1,1) = i]$ for all i and for $s \in [0,1]$

$$f(s) = p_0 + p_1 s + p_2 s^2 + \cdots = \sum_{i \geq 0} p_i s^i.$$

Similarly, $f_n(s) = \mathbb{E}[s^{Z_n}]$. One could hope to compute the generating function of the limit—but this is rarely possible. Instead, we derive some of its properties. In particular, note that

$$\begin{aligned} \pi &\equiv \mathbb{P}[Z_n = 0 \text{ for some } n \geq 0] \\ &= \lim_{n \to +\infty} \mathbb{P}[Z_n = 0] \\ &= \lim_{n \to +\infty} f_n(0), \end{aligned}$$

using the fact that 0 is an absorbing state and monotonicity. Moreover, by the Markov property, f_n as a natural recursive form:

$$\begin{aligned} f_n(s) &= \mathbb{E}[s^{Z_n}] \\ &= \mathbb{E}[\mathbb{E}[s^{Z_n} \,|\, \mathcal{F}_{n-1}]] \\ &= \mathbb{E}[f(s)^{Z_{n-1}}] \\ &= f_{n-1}(f(s)) = \cdots = f^{(n)}(s). \end{aligned}$$

So we need to study iterates of f. We will prove:

THM 15.5 (Extinction) *The probability of extinction π is given by the smallest fixed point of f in $[0,1]$:*

1. *If $m \leq 1$ then $\pi = 1$.*
2. *If $m > 1$ then $\pi < 1$.*

We first summarize some properties of f. To avoid uninteresting cases, we assume $p_0 + p_1 < 1$.

LEM 15.6 *The function f on $[0,1]$ satisfies:*

1. $f(0) = p_0$, $f(1) = 1$
2. *f is indefinitely differentiable on $[0,1)$*
3. *f is strictly convex and increasing*

4. $\lim_{s\uparrow 1} f'(s) = m < +\infty$

Proof: See Baby Rudin for the relevant power series facts. 1. is clear by definition. The function f is a power series with radius of convergence $R \geq 1$. This implies 2. In particular,

$$f'(s) = \sum_{i\geq 1} i p_i s^{i-1} \geq 0,$$

and

$$f''(s) = \sum_{i\geq 2} i(i-1) p_i s^{i-2} > 0.$$

because we must have $p_i > 0$ for some $i > 1$ by assumption. This proves 3. Since $m < +\infty$, $f'(1)$ is well defined and f' is continuous on $[0, 1]$. ■

COR 15.7 (Fixed points) *We have:*

1. *If $m > 1$ then f has a unique fixed point $\pi_0 \in [0, 1)$*
2. *If $m \leq 1$ then $f(t) > t$ for $t \in [0, 1)$ (Let $\pi_0 = 1$ in that case.)*

Proof: Since $f'(1) = m > 1$, there is $\delta > 0$ s.t. $f(1 - \delta) < 1 - \delta$. On the other hand $f(0) \geq 0$ so by continuity of f there must be a fixed point in $[0, 1 - \delta)$. Moreover, by strict convexity, if r is a fixed point then $f(s) < s$ for $s \in (r, 1)$, proving uniqueness.

The second part follows by strict convexity and monotonicity. ■

COR 15.8 (Dynamics) *We have:*

1. *If $t \in [0, \pi_0)$, then $f^{(n)}(t) \uparrow \pi_0$*
2. *If $t \in (\pi_0, 1)$ then $f^{(n)}(t) \downarrow \pi_0$*

Proof: We only prove 1. The argument for 2. is similar. By monotonicity, for $t \in [0, \pi_0)$, we have $t < f(t) < f(\pi_0) = \pi_0$. Iterating

$$t < f^{(1)}(t) < \cdots < f^{(n)}(t) < f^{(n)}(\pi_0) = \pi_0.$$

So $f^{(n)}(t) \uparrow L \leq \pi_0$. By continuity of f we can take the limit inside of

$$f^{(n)}(t) = f(f^{(n-1)}(t)),$$

to get $L = f(L)$. So by definition of π_0 we must have $L = \pi_0$. ■

Theorem 15.5 follows.

1.3 Discussion

The previous theorem "essentially" settles the subcritical and critical cases. For the supercritical case, however, it remains to understand when $M_\infty = 0$. When $M_\infty \equiv 0$ for instance, our convergence theorem provides less precise information. Note that convergence of expectations would help exclude that case since that would imply $\mathbb{E}[M_\infty] = 1$. But this requires some conditions. For instance, note that when $m \leq 1$

$$1 = \mathbb{E}[M_n] \nrightarrow \mathbb{E}[M_\infty] = 0.$$

In other words, the Martingale Convergence Theorem does not hold in L^1 under the same conditions.

More generally, one could conjecture that $M_\infty = 0$ exactly when we have extinction. We will see conditions under which this is true next time.

2 Martingales in $\mathcal{L}^2$

2.1 Preliminaries

DEF 15.9 *For $1 \leq p < +\infty$, we say that $X \in \mathcal{L}^p$ if*

$$\|X\|_p = \mathbb{E}[|X^p|]^{1/p} < +\infty.$$

By Jensen's inequality, for $1 \leq p \leq r < +\infty$ we have $\|X\|_p \leq \|X\|_r$ if $X \in \mathcal{L}^r$.

Proof: For $n \geq 0$, let

$$X_n = (|X| \wedge n)^p.$$

Take $c(x) = x^{r/p}$ on $(0, +\infty)$ which is convex. Then

$$(\mathbb{E}[X_n])^{r/p} \leq \mathbb{E}[(X_n)^{r/p}] = \mathbb{E}[(|X| \wedge n)^r] \leq \mathbb{E}[|X|^r].$$

Take $n \to \infty$ and use (MON). ■

DEF 15.10 *We say that X_n converges to X_∞ in $\mathcal{L}^p$ if $\|X_n - X_\infty\|_p \to 0$. By the previous result, convergence on $\mathcal{L}^r$ implies convergence in $\mathcal{L}^p$ for $r \geq p \geq 1$. (Moreover, by Chebyshev's inequality, convergence in $\mathcal{L}^p$ implies convergence in probability.)*

LEM 15.11 *Assume $X_n, X_\infty \in \mathcal{L}^1$. Then*

$$\|X_n - X_\infty\|_1 \to 0,$$

implies

$$\mathbb{E}[X_n] \to \mathbb{E}[X_\infty].$$

Proof: Note that

$$|\mathbb{E}[X_n] - \mathbb{E}[X_\infty]| \leq \mathbb{E}|X_n - X_\infty| \to 0.$$

■

DEF 15.12 *We say that $\{X_n\}_n$ is* bounded in $\mathcal{L}^p$ *if*

$$\sup_n \|X_n\|_p < +\infty.$$

2.2 $\mathcal{L}^2$ convergence

THM 15.13 *Let $\{M_n\}$ be a MG with $M_n \in \mathcal{L}^2$. Then $\{M_n\}$ is bounded in $\mathcal{L}^2$ if and only if*

$$\sum_{k \geq 1} \mathbb{E}[(M_k - M_{k-1})^2] < +\infty.$$

When this is the case, M_n converges a.s. and in $\mathcal{L}^2$. (In particular, it converges in $\mathcal{L}^1$.)

Proof:

LEM 15.14 (Orthogonality of increments) *Let $\{M_n\}$ be a MG with $M_n \in \mathcal{L}^2$. Let $s \leq t \leq u \leq v$. Then,*

$$\langle M_t - M_s, M_v - M_u \rangle = 0.$$

Proof: Use $M_u = \mathbb{E}[M_v \,|\, \mathcal{F}_u]$, $M_t - M_s \in \mathcal{F}_u$ and apply the $\mathcal{L}^2$ characterization of conditional expectations. ■

That implies

$$\mathbb{E}[M_n^2] = \mathbb{E}[M_0^2] + \sum_{1 \leq i \leq n} \mathbb{E}[(M_i - M_{i-1})^2],$$

proving the first claim.

By monotonicity of norms, M is bounded in $\mathcal{L}^2$ implies $\{M_n\}$ is bounded in $\mathcal{L}^1$ which, in turn, implies M converges a.s. Then using (FATOU) in

$$\mathbb{E}[(M_{n+k} - M_n)^2] = \sum_{n+1 \leq i \leq n+k} \mathbb{E}[(M_i - M_{i-1})^2],$$

gives

$$\mathbb{E}[(M_\infty - M_n)^2] \leq \sum_{n+1 \leq i} \mathbb{E}[(M_i - M_{i-1})^2].$$

The RHS goes to 0 which proves the second claim.

■

3 Back to branching processes

THM 15.15 *Let Z be a branching process with $Z_0 = 1$, $m = \mathbb{E}[X(1,1)] > 1$ and $\sigma^2 = \mathrm{Var}[X(1,1)] < +\infty$. Then, $M_n = m^{-n} Z_n$ converges in L^2, and in particular, $\mathbb{E}[M_\infty] = 1$.*

Proof: We bound $\mathbb{E}[M_n^2]$ by computing it explicitly by induction. From the orthogonality of increments

$$\mathbb{E}[M_n^2] = \mathbb{E}[M_{n-1}^2] + \mathbb{E}[(M_n - M_{n-1})^2].$$

On $\{Z_{n-1} = k\}$

$$\begin{aligned}
\mathbb{E}[(M_n - M_{n-1})^2 \,|\, \mathcal{F}_{n-1}] &= m^{-2n}\mathbb{E}[(Z_n - mZ_{n-1})^2 \,|\, \mathcal{F}_{n-1}] \\
&= m^{-2n}\mathbb{E}[(\sum_{i=1}^{k} X(i,n) - mk)^2 \,|\, \mathcal{F}_{n-1}] \\
&= m^{-2n} k\sigma^2 \\
&= m^{-2n} Z_{n-1}\sigma^2.
\end{aligned}$$

Hence

$$\mathbb{E}[M_n^2] = \mathbb{E}[M_{n-1}^2] + m^{-n-1}\sigma^2.$$

Since $\mathbb{E}[M_0^2] = 1$,

$$\mathbb{E}[M_n^2] = 1 + \sigma^2 \sum_{i=2}^{n+1} m^{-i},$$

which is uniformly bounded when $m > 1$. So M_n converges in L^2. Finally by (FATOU)

$$\mathbb{E}|M_\infty| \leq \sup \|M_n\|_1 \leq \sup \|M_n\|_2 < +\infty$$

and

$$|\mathbb{E}[M_n] - \mathbb{E}[M_\infty]| \leq \|M_n - M_\infty\|_1 \leq \|M_n - M_\infty\|_2,$$

implies the convergence of expectations. ■

In a homework problem, we will show that under the assumptions of the previous theorem

$$\{M_\infty = 0\} = \{Z_n = 0, \text{ for some } n\},$$

and

$$\mathbb{P}[M_\infty = 0] = \pi,$$

the probability of extinction.

EX 15.16 (Geometric Offspring) *Assume*

$$0 < p < 1,\ q = 1 - p,\ p_i = pq^i,\ \forall i \geq 0,\ m = \frac{q}{p}.$$

Then

$$f(s) = \frac{p}{1 - sq},\ \pi = \min\{\frac{p}{q}, 1\}.$$

- $\boxed{\text{Case } m \neq 1.}$ *If* G *is a* 2×2 *matrix, denote*

  $$G(s) = \frac{G_{11}s + G_{12}}{G_{21}s + G_{22}}.$$

  Then $G(H(s)) = (GH)(s)$. *By diagonalization,*

  $$\begin{pmatrix} 0 & p \\ -q & 1 \end{pmatrix}^n = (q-p)^{-1} \begin{pmatrix} 1 & p \\ 1 & q \end{pmatrix} \begin{pmatrix} p^n & 0 \\ 0 & q^n \end{pmatrix} \begin{pmatrix} q & -p \\ -1 & 1 \end{pmatrix}$$

  (the columns of the first matrix on the RHS are the right eigenvectors) leading to

  $$f_n(s) = \frac{pm^n(1-s) + qs - p}{qm^n(1-s) + qs - p}.$$

  In particular, when $m < 1$ *we have* $\pi = \lim f_n(0) = 1$. *On the other hand, if* $m > 1$, *we have by (DOM) for* $\lambda \geq 0$

  $$\begin{aligned} \mathbb{E}[\exp(-\lambda M_\infty)] &= \lim_n f_n(\exp(-\lambda/m^n)) \\ &= \frac{p\lambda + q - p}{q\lambda + q - p} \\ &= \pi + (1-\pi)\frac{(1-\pi)}{\lambda + (1-\pi)}. \end{aligned}$$

  The first term corresponds to a point mass at 0 *and the second term corresponds to an exponential with mean* $1/(1-\pi)$.

- $\boxed{\text{Case } m = 1.}$ *By induction*

  $$f_n(s) = \frac{n - (n-1)s}{n + 1 - ns},$$

  so that

  $$\mathbb{P}[Z_n > 0] = 1 - f_n(0) = \frac{1}{n+1},$$

  and

  $$\mathbb{E}[e^{-\lambda Z_n/n} \mid Z_n > 0] = \frac{f_n(e^{-\lambda/n}) - f_n(0)}{1 - f_n(0)} \to \frac{1}{1+\lambda},$$

  which is the Laplace transform of an eponential mean 1. *This is consistent with* $\mathbb{E}[Z_n] = 1$.

References

[AN72] Krishna B. Athreya and Peter E. Ney. *Branching processes*. Springer-Verlag, New York, 1972.

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010.

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.