

106349 - Advanced probability

Nick Crawford

November 6, 2018

Abstract

1 Introduction. Summary of course through an example. Branching process

We have an individual that gives a birth to a random number of offsprings – random variable X . X define a distribution, i.e., $P : \mathbb{Z}^+ \rightarrow [0, 1]$, i.e., $P(X = k) \in [0, 1]$, and $\sum_{k=0}^{\infty} P(X = k) = 1$.

Definition 1.1. $f_X(\theta) = \sum_{k=0}^{\infty} \theta^k P(X = k)$ – moment-generating function.

The series is absolutely convergent for $\theta \in [-1, 1]$ since k sums to 1. For $\theta \in (-1, 1)$, f_X is analytic, thus we can differentiate it term-by-term:

$$f'_X(\theta) = \sum_{k \geq 1} \theta^{k-1} P(X = k)$$

Since, f_X is analytic, knowing it means knowing $P(X = k)$ and vice versa.

Note that $f_X(0) = P(X = 0)$ and $f_X(1) = 1$. Also

$$f'_X(1) = \sum_{k \geq 0} k P(X = k) = \mathbb{E}X = \mu$$

$$\lim_{\theta \rightarrow 1} \frac{f_X(1) - f_X(\theta)}{1 - \theta} = \lim_{\theta \rightarrow 1} \frac{1 - f_X(\theta)}{1 - \theta}$$

Note also that f_X is convex, since second derivative is positive.

Size of n^{th} generation Let $(X_r^{(n)})_{n,r=1}^{\infty}$, where n is generation and r is offspring number (index) in n^{th} generation.

Assume $X_r^{(n)}$ are i.i.d. (independent, identically distributed) random variables.

Identically distributed means

$$P(X_n^r = k) = P(X = k)$$

Independence means

$$P(\forall i < J \ X_{r_i}^{n_i} = k) = \prod_{i=1}^J P(X_{r_i}^{n_i} = k)$$

Define $z_1 = X_1^1$. $z_2 = \sum_{r=1}^{z_1} X_r^2$ and so on:

$$z_{n+1} = \sum_{r=1}^{z_n} X_r^n$$

We want to study asymptotics of z_n .

Given U and V taking values in \mathbb{Z}^+ ,

$$\mathbb{E}[U|V = k] = \sum_{j=0}^{\infty} j P(U = j|V = k)$$

, where

$$P(U = j|V = k) = \frac{P(U = j, V = k)}{P(V = k)}$$

If U, V are independent, $P(U = j|V = k) = P(U = j)$ and thus $\mathbb{E}[U|V = k] = \mathbb{E}U$.

Definition 1.2. Define random variable $\mathbb{E}[U|V]$ such that

$$\mathbb{E}[U|V] = \mathbb{E}[U|V = k]$$

if $V = k$.

Definition 1.3 (Tower property).

$$\mathbb{E}[\mathbb{E}[U|V]] = \mathbb{E}U$$

Define

$$f_n = \sum_{k=0}^{\infty} \sum_{k=0}^{\infty} \theta^k P(z_n = k) = \mathbb{E}\theta^{z_n}$$

Theorem 1.1.

$$f_{n+1}(\theta) = f_n(f_X(\theta))$$

or

$$f_n(\theta) = \underbrace{f \circ f \circ \dots \circ f}_{n \text{ times}}(\theta)$$

Proof. Use tower property with $U^{z_{n+1}}$ and $V = \theta^{z_n}$. By tower property

$$\mathbb{E}[\theta^{z_{n+1}}] = \mathbb{E}[\mathbb{E}[\theta^{z_{n+1}}|\theta^{z_n}]]$$

$$\mathbb{E}[\mathbb{E}[\theta^{z_{n+1}}|\theta^{z_n}]] = \sum_{k=0}^{\infty} P(z_n = k) \mathbb{E}[\theta^{z_{n+1}}|\theta^{z_n} = k]$$

What is $\mathbb{E}[\theta^{z_{n+1}}|\theta^{z_n} = k]$?

$$\mathbb{E}[\theta^{z_{n+1}}|\theta^{z_n} = k] = \mathbb{E}[\theta^{\sum_{j=1}^k X_j^{n+1}}|\theta^{z_n} = k] \stackrel{\text{independence}}{=} \mathbb{E}[\theta^{\sum_{j=1}^{z_n} X_j^{n+1}}] \stackrel{\text{independence}}{=} \prod_{j=1}^k \mathbb{E}[\theta^{X_j^{n+1}}] \stackrel{\text{i.d.}}{=} (f_X(\theta))^k$$

Thus

$$\mathbb{E}[\mathbb{E}[\theta^{z_{n+1}}|\theta^{z_n}]] = \sum_{k=0}^{\infty} P(z_n = k) (f_X(\theta))^k = f_n(f(\theta))$$

Also we can say

$$\mathbb{E}[\theta^{z_{n+1}}|z_n] = (f_X(\theta))^{z_n}$$

□

Study of z_n What is $\pi_n = P(z_n = 0) = f_n(0) = f(\pi_{n-1})$, probability that population is extinguished. Since $z_{n-1} = 0 \Rightarrow z_n = 0$, i.e. π_n is non-decreasing.

Let $P(z_n = 0 \text{ for some } n) = \pi$.

We hope that $\{z_n = 0\}$ such that

$$\bigcup_n \{z_n = 0\} = \{z_n = 0 \text{ for some } n\}$$

i.e., $\pi = \lim_{n \rightarrow \infty} \pi_n$. We call π the extinction probability.

Theorem 1.2. If $\mu = \mathbb{E} > 1$ then π is a unique root of $\pi = f(\pi)$ and $\pi \in [0, 1]$. If $\mu \leq 1$, $\pi = 1$.

If we look at $f(\pi)$ and π , they intersect in 1, and they can intersect in two points since $f(x)$ is convex. There is second intersection iff $f'(1) = \mu > 1$.

Construction of X_n^r Construct set Ω , $f_{n,r} : \Omega \rightarrow \mathbb{Z}^+$ and \mathcal{F} a collection of subsets of Ω with $P : \mathcal{F} \rightarrow [0, 1]$.

Let $\Omega = \mathbb{Z}^+ \times \mathbb{Z}^+$, $\mathcal{F} = \{0, 1\}^\Omega$.

The problem is when we have infinitely number of variables.

Example Example of not well-behaved triple (Ω, \mathcal{F}, P) . $\Omega = \mathbb{N}$. Now $\mathcal{F} = \{C \subset \mathbb{N} : C \text{ has density}\}$. C has density means

$$\frac{|C \cap \mathbb{N}|}{n} \xrightarrow{n \rightarrow \infty} \rho(C)$$

However, for $C(m) = \{1, 2, \dots, m\}$, $\forall m$ $\rho(C_m)$, and

$$\rho\left(\bigcup C_m\right) = 1$$

Thus $(\mathbb{N}, \mathcal{F}, \rho)$ is not a good probability space, since it doesn't fulfill this $\pi_n \rightarrow \pi$ property. Note we can define other probabilities on naturals, for example

$$P(\{i\}) = 2^{-i}$$

Asymptotics of z_i Assuming $\pi \in (0, 1)$, what is behavior of z_n ?

Definition 1.4. z_n is a Markov chain if

$$P(z_{n+1} = j | z_i = k_i \quad \forall i \leq n) = P(z_{n+1} = j | z_n = k_n)$$

We can use to compute expectation:

$$\mathbb{E}[z_{n+1} | z_i = k_i \quad \forall i < n] = E[z_{n+1} | z_n = k_n]$$

Then, since $E\left[\sum_{i=1}^J X_i^n\right] = J\mu$

$$E[z_{n+1} | z_n] = \mu z_n$$

Let $M_n = \frac{z_n}{\mu^n}$ then $\mathbb{E}[M_n] = 1$. Also

$$\mathbb{E}[M_{n+1} | z_0, \dots, z_n] = M_n$$

This is a definition of martingale with respect to z_0, \dots, z_n .

Let (Ω, \mathcal{F}, P) we say S happens almost surely (a.s.) if

$$P(\{w \in \Omega : S \text{ is true for } w\}) = 1$$

Theorem 1.3 (Martingale convergence theorem). If M_n is a positive martingale then $\lim_{n \rightarrow \infty} M_n = M_\infty$ exists a.s. and

- $\mu \leq 1$. $M_\infty = 0$ a.s. That means $\mathbb{E}M_\infty = 0$ but $\mathbb{E}M_n = 1$, i.e.,

$$\mathbb{E}\left[\liminf_{n \rightarrow \infty} M_n\right] < \liminf_{n \rightarrow \infty} \mathbb{E}[M_n]$$

- $\mu > 1$. If $M_\infty > 0$ with positive probability then $z_n \sim \mu^n M_\infty$.

Lemma 1.1 (Fatou's lemma).

$$\mathbb{E}\left[\liminf_{n \rightarrow \infty} M_n\right] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[M_n]$$

Theorem 1.4.

$$\mathbb{E}[M_\infty] = 1 \iff \mu > 1 \quad \text{and} \quad \mathbb{E}[X \log(X)] < \infty$$

2 Overview of measure theory

Notation

- S is a set.
- \mathcal{A} is algebra of subsets of S

1. $S \in \mathcal{A}$

2.

$$E \in \mathcal{A} \Rightarrow E^C \in \mathcal{A}$$

, where $E^C = S \setminus E$

3.

$$E_1, E_2 \in \mathcal{A} \Rightarrow E_1 \cup E_2 \in \mathcal{A}$$

meaning

$$E_1, E_2 \in \mathcal{A} \Rightarrow E_1 \cap E_2 \in \mathcal{A}$$

- \mathcal{F} is a σ -algebra if lest item works for countable union.
- $E \Delta F = E \setminus F \cup F \setminus E$

Definition 2.1. A measurable space is a pair $\{S, \mathcal{F}\}$.

Proposition 2.1. If we have $(\mathcal{F}_i)_{i \in I}$, then $\bigcap_{i \in I} \mathcal{F}_i$ is also a σ -algebra.

Definition 2.2. Let C be a collection of subsets of S . $\sigma(C)$ is a smallest σ -algebra containing C (σ -algebra generated by C). It is easy to construct one

$$I = \{\mathcal{F} : \mathcal{F} \supset C\}$$

and then

$$\sigma(C) = \bigcap_{\mathcal{F} \in I} \mathcal{F}$$

Definition 2.3. Let $\{S, \mathcal{F}\}$ be a topological space. $\mathcal{B}(X)$ (Borel σ -algebra) is defined as σ -algebra generated by open sets. We denote $\mathcal{B} = \mathcal{B}(\mathbb{R})$.

Exercise

$$\pi(\mathbb{R}) = \{(-\infty, x], x \in \mathbb{R}\}$$

Show that $\sigma(\pi(\mathbb{R})) = \mathcal{B}$

Definition 2.4. Additive set function on a collection of sets \mathcal{F} is

$$\mu : \mathcal{F} \rightarrow [0, \infty)$$

$$\forall E, F \in \mathcal{F} \ E \cap F = \emptyset \quad \mu(E \cup F) = \mu(E) + \mu(F)$$

We say μ is σ -additive if same holds of countable infinite sets

$$\forall \{E_i\}_{i=1}^{\infty} \ E_i \cap E_j = \emptyset \quad \mu(E \cup F) = \sum_{i=1}^{\infty} \mu(E_i)$$

Definition 2.5. A triple (S, \mathcal{F}, μ) is a measure space if \mathcal{F} is a σ -algebra on S and μ is σ -additive on \mathcal{F} .

Definition 2.6. (S, \mathcal{F}, μ) is finite if $\mu(S) < \infty$

(S, \mathcal{F}, μ) is σ -finite if

$$\exists \{E_i, \mu(E_i) < \infty\}_{i=1}^{\infty} \quad S = \bigcup_{i=1}^{\infty} E_i$$

Definition 2.7. If $\mu(S) = 1$, (S, \mathcal{F}, μ) is probability space.

Definition 2.8. E is null if $\mu(E) = 0$.

Definition 2.9. ϕ is said to be true almost everywhere with respect of μ if

$$\mu(\{X : \phi(X) = \text{False}\}) = 0$$

2.1 Results from measure theory

Definition 2.10. A collection of sets \mathcal{D} is called a π -system if $E, F \in \mathcal{D} \Rightarrow E \cap F \in \mathcal{D}$

Theorem 2.2 (Uniqueness). Let \mathcal{D} be a π -system generating a σ -algebra \mathcal{F} . Let μ_1 and μ_2 be two finite measures on \mathcal{F} which agree on \mathcal{D} . Then $\mu_1 = \mu_2$.

Collary 2.1. $(S, \mathcal{F}, P_1), (S, \mathcal{F}, P_2)$ probability spaces, $P_1 = P_2$ on π -system \mathcal{D} , then $P_1 = P_2$.

Theorem 2.3 (Carathéodory's extension theorem). Let \mathcal{A} be an algebra of sets. $\mu_0 : \mathcal{A} \rightarrow \mathbb{R}^+$ σ -additive set function on \mathcal{A} . Then exists unique extension $\bar{\mu} : \sigma(\mathcal{A}) \rightarrow \mathbb{R}^+$ such that $\bar{\mu}|_{\mathcal{A}} = \mu_0$.

Homework Lebesgue on \mathbb{R} . $\mathcal{A} = \{\text{open set}\}$. If we have

$$O = \bigcup_{i=1}^{\infty} (a_i, b_i)$$

then

$$\mu_0(O) = \sum_{i=1}^{\infty} b_i - a_i$$

Check that μ_0 is well defined and σ -additive.

Lemma 2.1. (S, \mathcal{F}, μ) measure space. $A, B \in \mathcal{F}$, then

$$\mu(A \cup B) \leq \mu(A) + \mu(B)$$

$$\mu\left(\bigcup_{i=1}^{\infty} F_i\right) \leq \sum_{i=1}^{\infty} \mu(F_i)$$

If $\mu(S) < \infty$

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

From that we get inclusion-exclusion:

$$\mu\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mu(A_i) - \sum_{i \neq j} \mu(A_i \cap A_j) + \cdots + (-1)^{n-1} \mu\left(\bigcap_{i=1}^n A_i\right)$$

Exercise Proof the lemma

Lemma 2.2. If $F_n \subseteq F_{n+1}$ then

$$\mu\left(\bigcup_{i=1}^{\infty} F_i\right) = \lim_{n \rightarrow \infty} \mu(F_n)$$

If $\mu(S) < \infty$ and $F_n \supseteq F_{n+1}$ then

$$\mu\left(\bigcap_{i=1}^{\infty} F_i\right) = \lim_{n \rightarrow \infty} \mu(F_n)$$

Proof. Assume $\mu(S) < \infty$. Define $F_{\infty} = \bigcup_{i=1}^{\infty} F_i$. Let $G_n = F_n \setminus F_{n+1}$. Then

$$F_{\infty} = \bigcup_{i=1}^{\infty} G_i$$

Meaning

$$\mu(F_{\infty}) = \sum_{i=1}^{\infty} \mu(G_i)$$

$$\mu(F_n) = \sum_{k=1}^n \mu(G_k)$$

Since measure is finite, the tail of series tends to 0, thus

$$\mu(F_{\infty}) - \mu(F_n) = \sum_{k=n}^{\infty} \mu(G_k) \rightarrow 0$$

Then we can take complements and get the second statement. □

Exercise Proof unconditionally

3 Recasting measure theory as probability

Definition 3.1. A probability space is a (Ω, \mathcal{F}, P) is a measure space such that $P(\Omega) = 1$. We call $\omega \in \Omega$ an outcome. $E \in \mathcal{F}$ is an event. $P(E)$ is probability of the event.

Example Tossing finite or infinite sequence of coins.

Tossing 4 coins

$$\Omega = \{HHHH, HHHT, HHTH, \dots, TTTH, TTTT\}$$

$$\mathcal{F} = 2^\Omega$$

$$P(\omega \in \Omega) = \frac{1}{|\Omega|}$$

Tossing infinite number coins

$$\Omega = \{0, 1\}^{\mathbb{N}}$$

Ω has a natural topology which is called a product topology. It is coarsest topology such that $\pi_i : \Omega \rightarrow \{0, 1\}$ $\pi_i(\omega) = \omega_i$ is continuous.

Let $\mathcal{F} = \mathcal{B}(\Omega)$.

Smallest σ -algebra such that

$$\pi_i^{-1}(0) \subset \Omega \in \mathcal{F}$$

$$\pi_i^{-1}(1) \subset \Omega \in \mathcal{F}$$

$$\pi_i(\Omega, \mathcal{F}) \rightarrow \left(\{0, 1\}, \{0, 1\}^{\{0, 1\}} \right)$$

Natural π -system \mathcal{F}_n smallest σ -algebra making π_1, \dots, π_n measurable.

Note that

Proposition 3.1.

$$\bigcup_n \mathcal{F}_n \neq \mathcal{F}$$

Proof. Define $S_n(\omega) = \sum_{i=1}^n \omega_i$.

$$X_n = \frac{S_n(\omega)}{n}$$

Define

$$Y(\omega) = \limsup X_n(\omega)$$

$$E = \left\{ \omega : Y(\omega) \geq \frac{1}{3} \right\}$$

$$E \in \mathcal{F} \setminus \bigcup_n \mathcal{F}_n$$

□

What \mathcal{F}_n looks like? For example, \mathcal{F}_2 has 4 outcomes, deciding only first two tosses.

Note If we take $(\Omega_4, \mathcal{F}^{(4)}, P_4)$, restricting to $(\Omega_3, \mathcal{F}^{(3)}, P_3)$

$$P_4(\{(0, 0, 0, \omega_4)\}) = P_3(\{(0, 0, 0)\})$$

Thus we want P_{fair} defined on Ω to fulfill same property:

$$P_{fair}(E) = P_n(\tilde{E})$$

where $E \in \mathcal{F}_n$ and $\tilde{E} \in \mathcal{F}^{(n)}$.

Definition 3.2. $E \subset \mathcal{F}$ occurs almost surely (a.s.) if $P(E) = 1$.

Definition 3.3 (lim sup and lim inf). Let $\{E_n\}$ be a sequence of events.

$$\limsup E_n = \bigcap_m \bigcup_{n \geq m} E_n = \{E_n \text{ occurs infinitely often (i.o.)}\} = \{\omega \in \Omega : \forall m \exists n(\omega) > m \quad \omega \in E_n(\omega)\}$$

Alternatively, (Ω, \mathcal{F}) and $\{E_n\}$ there is a natural map

$$I : \Omega \rightarrow \{0, 1\}^{\mathbb{N}}$$

$$\omega \mapsto \{1_{E_n}(\omega)\}$$

where

$$1_E(\omega) = \begin{cases} 0 & \omega \notin E \\ 1 & \omega \in E \end{cases}$$

Now

$$\liminf E_n = \bigcup_m \bigcap_{n \geq m} E_n = \{E_n \text{ occurs eventually}\} = \{\omega \in \Omega : \exists m(\omega) \forall n \geq m(\omega) \quad \omega \in E_n(\omega)\}$$

Remark Since everything is countable, if $E_n \in \mathcal{F}$, then $\limsup E_n, \liminf E_n \in \mathcal{F}$

We can write

$$\left\{ \frac{S_n}{n} \rightarrow \frac{1}{2} \right\} = \left\{ \limsup \frac{S_n}{n} \leq \frac{1}{2} \right\} \cap \left\{ \liminf \frac{S_n}{n} \geq \frac{1}{2} \right\}$$

Choose $q \in \mathbb{Q}^+$ and take a look at

$$\left\{ \liminf \frac{S_n}{n} > q \right\} = \liminf E_n(q)$$

where $E_n = \{\omega : \frac{S_n}{n} > q\}$.

In addition

$$\left\{ \limsup \frac{S_n}{n} < q \right\} = \liminf F_n(q)$$

where $F_n = \{\omega : \frac{S_n}{n} < q\}$.

Therefore $\{\liminf \frac{S_n}{n} > q\} \in \mathcal{F}$.

Finally,

$$\left\{ \liminf \frac{S_n}{n} \geq \alpha \right\} = \bigcap_{q < \alpha} \left\{ \liminf \frac{S_n}{n} > q \right\}$$

Lemma 3.1 (Fatou's lemma).

$$P\left[\liminf_{n \rightarrow \infty} E_n\right] \leq \liminf_{n \rightarrow \infty} p(E_n)$$

Proof.

$$\liminf_{n \rightarrow \infty} E_n = \bigcup_m \bigcap_{n \geq m} E_n$$

Sets $F_m = \bigcap_{n \geq m} E_n$ are increasing and $F_n \subseteq E_n$, thus

$$P\left[\liminf_{n \rightarrow \infty} E_n\right] = \lim_{n \rightarrow \infty} P(F_n) \leq \liminf_{n \rightarrow \infty} P(E_n)$$

□

Lemma 3.2 (Fatou's lemma).

$$P\left[\limsup_{n \rightarrow \infty} E_n\right] \geq \limsup_{n \rightarrow \infty} p(E_n)$$

Proof. Note that $(\limsup E_n)^C = \liminf E_n^C$, thus this is straightforward from previous lemma.

□

Lemma 3.3 (First Borel-Cantelli lemma). Let $\{E_n\} \subseteq \mathcal{F}$ be a sequence of events s.t. $\sum_n P(E_n) < \infty$, then

$$P(E_n \text{ happens i.o.}) = 0$$

Proof.

$$P(E_n \text{ i.o.}) = P\left(\bigcap_m \bigcup_{n \geq m} E_n\right) \leq P\left(\bigcup_{n \geq m} E_n\right) \leq \sum_{n=m}^{\infty} P(E_n) \xrightarrow{m \rightarrow \infty} 0$$

Since $P(E_n \text{ i.o.})$ is independent on m , it got to be 0.

□

Example Fix $\epsilon > 0$. Look at $P\left(\left|\frac{S_n(\omega)}{n} - \frac{1}{2}\right| > \epsilon\right)$.

Claim

$$P\left(\left|\frac{S_n(\omega)}{n} - \frac{1}{2}\right| > \epsilon\right) \leq \frac{12}{\epsilon^4} \frac{1}{n^2}$$

By 3.3 $P\left(\left|\frac{S_n(\omega)}{n} - \frac{1}{2}\right| > \epsilon \text{ i.o.}\right) = 0$ thus

$$\left\{\frac{S_n}{n} \rightarrow \frac{1}{2}\right\} = \bigcap_{\epsilon > 0} \left\{\left|\frac{S_n(\omega)}{n} - \frac{1}{2}\right| < \epsilon \text{ eventually}\right\} = 1$$