Section Notes for PSTAT 213

Haosheng Zhou

Sept, 2023

Contents

This note contains extra exercises, examples and materials for PSTAT 213. The notes may be subject to typos, and you are welcome to email me at hzhou593@ucsb.edu for any possible advice.

Week 1

Example for Indicator

Lemma 1 (Example)**.** *The indicator I_A of event A is a random variable defined as*

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{else} \end{cases} \tag{1}$$

(a): Write PMF for $X = I_A$ and calculate $\mathbb{E}X, Var(X), G_X(s)$.

(b): For random variable $Y \geq 0$ and ϕ as any non-negative increasing function on $[0, +\infty)$, show that $\forall a > 0, \phi(a) \cdot \mathbb{P}(Y \geq a) \leq \mathbb{E}\phi(Y)$ so that $\forall \varepsilon > 0, \mathbb{P}(|Z| \geq \varepsilon) \leq \frac{\mathbb{E}Z^2}{\varepsilon^2}$ for any random variable Z.

(c): Assume Y is a random variable such that its MGF $M_Y(t) = \mathbb{E}e^{tY}$ is finite for all $t \in \mathbb{R}$, show that when $t \geq 0$, $\mathbb{P}(X \geq x) \leq e^{-tx} M_X(t)$ so that $\mathbb{P}(X \geq x) \leq \inf_{t \geq 0} e^{-tx} M_X(t)$.

Proof. (a): X has support $\{0, 1\}$ with $\mathbb{P}(X = 1) = \mathbb{P}(A), \mathbb{P}(X = 0) = \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ gives the PMF.

From the PMF, it's easy to calculate $\mathbb{E}X = \mathbb{P}(A), \mathbb{E}X^2 = \mathbb{P}(A)$ so $Var(X) = \mathbb{E}X^2 - \mathbb{E}^2 X = \mathbb{P}(A) - [\mathbb{P}(A)]^2$.

$$G_X(s) = \mathbb{E}s^X = 1 \cdot \mathbb{P}(X = 0) + s \cdot \mathbb{P}(X = 1) = 1 - \mathbb{P}(A) + s \cdot \mathbb{P}(A) \tag{2}$$

(b): This is the classical trick on indicator

$$\forall a > 0, \phi(a) \cdot \mathbb{P}(Y \geq a) = \mathbb{E}[\phi(a)\mathbb{I}_{Y \geq a}] \leq \mathbb{E}[\phi(Y)\mathbb{I}_{Y \geq a}] \leq \mathbb{E}\phi(Y) \tag{3}$$

since ϕ is increasing and indicator is non-negative and takes value no larger than 1.

Consider $\phi(x) = x^2$ non-negative and increasing on $[0, +\infty)$ plugging in $Y = |Z| \geq 0, a = \varepsilon$ to conclude the proof.

(c):

Since for $t > 0$, e^{tx} is non-negative and increasing in x, resulting in

$$\mathbb{P}(X \geq x) = \mathbb{P}\left(e^{tX} \geq e^{tx}\right) \leq e^{-tx}\mathbb{E}e^{tX} = e^{-tx} M_X(t) \tag{4}$$

applying the conclusion in (b) for $Y = e^{tX}, a = e^{tx}, \phi(x) = x$. When $t = 0$, check $e^{-tx} M_X(0) = 1$ so $\mathbb{P}(X \geq x) \leq 1$ naturally holds. This proves that the inequality holds for $\forall t \geq 0$. Taking inf on both sides w.r.t. t concludes the proof.

□

Remark. *Part (c) is a very important technique that will appear once again in 213BC to derive the Chernoff bound of concentration of measures. The basic idea is to* ***introduce some unspecified parameter t, build a bound for the probability and optimize the bound to get the tightest bound by specifying an appropriate value of t.***

Example for Branching Process

Lemma 2 (Example)**.** *There is an isolated island with the original stock of 100 family surnames, and the survival of family names is modelled by branching process, different surnames' survivals are independent. Each surname has extinction probability $\eta = \frac{9}{10}$.*

(a): After many generations how many surnames do you expect to be on the island?

(b): Do you expect the total population on the island to be increasing or decreasing?

Proof. (a): Each surname has η probability of disappearing independent of other surnames so the number of surname survived after a long enough time denoted X has binomial distribution $X \sim B(100, 1-\eta)$. It's clear that $\mathbb{E}X = 100(1-\eta) = 10$.

(b): Since $\eta > 0, \eta \neq 1$, the branching process $\{Z_n\}$ for each family surname is in the supercritical phase with offspring mean $\mu > 1$. It's clear that $\mathbb{E}Z_n = \mu^n \to +\infty$ $(n \to \infty)$ so the expected total population is increasing. □

Lemma 3 (Example)**.** *Branching process $\{Z_n\}$ originates from one individual, i.e. $Z_0 = 1$ has Poisson offspring distribution $Z_1 \sim P(\lambda)$ $(\lambda > 1)$. If it's known that a branching process conditional on extinction is still a branching process, i.e. let A stands for the event that $\{Z_n\}$ extinct, $\{E_n\} = \{Z_n\}|_A$ is still a branching process. Can you derive the offspring distribution for $\{E_n\}$?*

Proof. Since $E_0 = Z_0|_A = 1$, the offspring distribution for $\{E_n\}$ is just the distribution of E_1. Let's denote η as the extinction probability of $\{Z_n\}$, i.e. $\eta = \mathbb{P}(A)$ and $p_k = \mathbb{P}(Z_1 = k)$ as the offspring distribution PMF of $\{Z_n\}$.

$$\mathbb{P}(E_1 = k) = \mathbb{P}(Z_1 = k|A) = \frac{\mathbb{P}(A|Z_1 = k)\,\mathbb{P}(Z_1 = k)}{\mathbb{P}(A)} \tag{5}$$

using Bayes formula. Notice that conditional on $Z_1 = k$, extinction happens if and only if all k subtrees generated in generation 1 are extinct. Since all k subtrees are independent and follow the same offspring distribution, they have exactly the same probability of being extinct, resulting in

$$\mathbb{P}(A|Z_1 = k) = [\mathbb{P}(A|Z_1 = 1)]^k = [\mathbb{P}(A)]^k = \eta^k \tag{6}$$

where the second equation comes from the fact that if $Z_1 = 1$, restarting the branching process at generation 1 makes no difference to the extinction probability (this is actually the Markov property of branching process). At this point, we see that

$$\mathbb{P}(E_1 = k) = \eta^{k-1} p_k = \eta^{k-1}\frac{\lambda^k}{k!}e^{-\lambda} \tag{7}$$

Since $\lambda > 1$, the offspring mean is larger than 1, the extinction probability η is thus the fixed point of $G(s)$ with

$$G(s) = \mathbb{E}s^{Z_1} = \sum_{k=0}^{\infty} s^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{s\lambda - \lambda} \tag{8}$$

telling us

$$e^{\eta\lambda - \lambda} = \eta \tag{9}$$

turning it into $e^{-\lambda} = \eta e^{-\eta\lambda}$ and replace the $e^{-\lambda}$ term in the expression of $\mathbb{P}(E_1 = k)$ to get

$$\mathbb{P}(E_1 = k) = \frac{(\eta\lambda)^k}{k!} e^{-\eta\lambda}, E_1 \sim P(\eta\lambda) \tag{10}$$

the offspring distribution of $\{E_n\}$ is still Poisson but it's $P(\eta\lambda)$.

□

Remark. *Actually* ***any branching process conditional on extinction is still a branching process****. Unfortunately, there is no easy approach to prove this conclusion since it's a statement for the whole process but not for pointwise evaluation of the process. Proving this conclusion requires the correspondence between branching process and random walk which we might have the chance to introduce in the future.*

However, we can do heuristic calculations as above to calculate the offspring distribution of the new branching process. From what we have shown above, the new branching process $\{E_n\}$ has offspring distribution with PMF

$$\mathbb{P}(E_1 = k) = p'_k = \eta^{k-1} p_k \tag{11}$$

this is called ***the duality principle of branching process****. In particular, Poisson branching process conditional on extinction still provides a Poisson branching process.*

Lemma 4 (Example)**.** *A branching process $\{Z_n\}$ is given such that $Z_0 = 8$ with offspring distribution PMF $p_0 = 0.2, p_1 = 0.5, p_2 = 0.3$.*

(a): Derive its extinction probability η.

(b): Derive the probability that the process is extinct in generation 3 but survives in generation 1 and generation 2.

Proof. (a): Such branching process is actually the sum of 8 branching process $\left\{Z_n^{(1)}\right\}, ..., \left\{Z_n^{(8)}\right\}$ with the same offspring distribution but with $Z_0^{(1)} = ... = Z_0^{(8)} = 1$. Moreover, those 8 branching processes are independent (by the definition of branching process).

Denote $E_n^{(i)}$ as the event that $\left\{Z_n^{(i)}\right\}$ is extinct in generation n and $S_n^{(i)}$ as the event that $\left\{Z_n^{(i)}\right\}$ survives in generation n, $E^{(i)}$ as the event that $\left\{Z_n^{(i)}\right\}$ is extinct. It's clear that $\{Z_n\}$ is extinct if and only if $\left\{Z_n^{(1)}\right\}, ..., \left\{Z_n^{(8)}\right\}$ are all extinct.

$$\eta = \mathbb{P}\left(E^{(1)}, E^{(2)}, ..., E^{(8)}\right) = \left[\mathbb{P}\left(E^{(1)}\right)\right]^8 \tag{12}$$

since offspring mean $\mu = 0.5 + 2 \times 0.3 = 1.1 > 1$, $\left\{Z_n^{(i)}\right\}$ is in supercritical phase, $\mathbb{P}\left(E^{(1)}\right)$ is the fixed point of $G(s)$. Let's first derive generating function

$$G(s) = 0.2 + 0.5s + 0.3s^2 \tag{13}$$

and solve $G(s) = s$ to get the solution $\mathbb{P}\left(E^{(1)}\right) = \frac{2}{3}$. We get the answer

$$\eta = \left(\frac{2}{3}\right)^8 \tag{14}$$

(b): $\{Z_n\}$ is extinct in generation 3 iff all $\left\{Z_n^{(i)}\right\}$ are extinct in generation 3. $\{Z_n\}$ survives in generation 2 iff there exists some $\left\{Z_n^{(i)}\right\}$ survive in generation 2. Notice that $\{Z_n\}$ survives in generation 2 implies $\{Z_n\}$ survives in generation 1 so the probability we want to find is the probability that $\{Z_n\}$ is extinct in generation 3 and survives in generation 2.

$$\mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)} \cap \bigcup_{i=1}^8 S_2^{(i)}\right) = \mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)}\right) - \mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)} \cap \left[\bigcup_{i=1}^8 S_2^{(i)}\right]^c\right) \tag{15}$$

$$= \mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)}\right) - \mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)} \cap \bigcap_{i=1}^8 \left[S_2^{(i)}\right]^c\right) \tag{16}$$

$$= \mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)}\right) - \mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)} \cap \bigcap_{i=1}^8 E_2^{(i)}\right) \tag{17}$$

is what we want to calculate by noticing $\forall n, i$, $\left[S_n^{(i)}\right]^c = E_n^{(i)}$. Use the fact that extinction in generation 2 implies extinction in generation 3, this tells us

$$\mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)} \cap \bigcap_{i=1}^8 E_2^{(i)}\right) = \mathbb{P}\left(\bigcap_{i=1}^8 (E_3^{(i)} \cap E_2^{(i)})\right) = \mathbb{P}\left(\bigcap_{i=1}^8 E_2^{(i)}\right) \tag{18}$$

the structure of independence helps us again

$$\mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)} \cap \bigcup_{i=1}^8 S_2^{(i)}\right) = \mathbb{P}\left(\bigcap_{i=1}^8 E_3^{(i)}\right) - \mathbb{P}\left(\bigcap_{i=1}^8 E_2^{(i)}\right) = \left[\mathbb{P}\left(E_3^{(1)}\right)\right]^8 - \left[\mathbb{P}\left(E_2^{(1)}\right)\right]^8 \tag{19}$$

the final step is to calculate those two probabilities. Recall the property of generating function that $G_X(0) = \mathbb{P}(X = 0)$. Now $\mathbb{P}\left(E_2^{(1)}\right) = \mathbb{P}\left(Z_2^{(1)} = 0\right) = G_{Z_2^{(1)}}(0)$ and we have proved in class that $Z_2^{(1)}$ has generating function

$G(G(s))$. This tells us

$$\begin{cases} \mathbb{P}\left(E_2^{(1)}\right) = G(G(0)) = G(0.2) = 0.312 \\ \mathbb{P}\left(E_3^{(1)}\right) = G(G(G(0))) = G(0.312) = 0.3852 \end{cases} \tag{20}$$

so the probability we want to find is

$$0.3852^8 - 0.312^8 \tag{21}$$

□

Extra Materials: Total Progeny

For branching process $\{Z_n\}$ with $Z_0 = 1$, offspring distribution $\{p_k\}$ and generating function of offspring distribution $G(s)$, the **total progeny** is defined as

$$T = \sum_{n=0}^{\infty} Z_n \tag{22}$$

the overall number of individuals in the branching process. It's easy to see that if extinction probability $\eta = 1$, then $T < \infty$ *a.s.*, otherwise T has positive probability taking value ∞. Due to this fact, the generating function of the total progeny is defined as

$$G_T(s) = \mathbb{E}\left(s^T \cdot \mathbb{I}_{T<\infty}\right) \tag{23}$$

with the indicator added to make sure that $G_T(s)$ is well-defined. Deriving the generating function of T would provide us with a taste of how things work in branching process.

Theorem 1. ***(Generating Function of Total Progeny)***

$$\forall s \in [0,1), G_T(s) = s \cdot G(G_T(s)) \tag{24}$$

Proof. Tear apart the expectation w.r.t. the value of Z_1 to get

$$G_T(s) = \sum_{k=0}^{\infty} \mathbb{P}(Z_1 = k) \cdot \mathbb{E}(s^T \cdot \mathbb{I}_{T<\infty} | Z_1 = k) \tag{25}$$

now under the condition that $Z_1 = k$, $T = 1 + T_1 + ... + T_k$ where T_j denotes the total progeny of the descendants of the j-th person in generation 1

$$G_T(s) = \sum_{k=0}^{\infty} p_k \cdot s \cdot \mathbb{E}(s^{T_1} ... s^{T_k} \cdot \mathbb{I}_{T_1<\infty} ... \mathbb{I}_{T_k<\infty} | Z_1 = k) \tag{26}$$

notice that $T_1, ..., T_k, Z_1$ are independent and $T_1, ..., T_k$ are identically distributed, so

$$
\begin{aligned}
G_T(s) &= \sum_{k=0}^{\infty} p_k \cdot s \cdot \mathbb{E}(s^{T_1}...s^{T_k} \cdot \mathbb{I}_{T_1<\infty}...\mathbb{I}_{T_k<\infty}) && (27)\\
&= s \cdot \sum_{k=0}^{\infty} p_k \cdot \mathbb{E}(s^{T_1} \cdot \mathbb{I}_{T_1<\infty})...\mathbb{E}(s^{T_k} \cdot \mathbb{I}_{T_k<\infty}) && (28)\\
&= s \cdot \sum_{k=0}^{\infty} p_k \cdot [G_{T_1}(s)]^k && (29)\\
&= s \cdot G(G_{T_1}(s)) && (30)
\end{aligned}
$$

at last notice that $T \overset{d}{=} T_1$ since the branching process starting from generation 0 with 1 individual is the same in distribution as the branching process starting from generation 1 with 1 individual, so the distribution of the total progeny in these two cases are the same. We conclude that

$$G_T(s) = s \cdot G(G_T(s)) \tag{31}$$

□

Remark. *By noticing the continuity of G_T and taking $s \to 1^-$, one may find that*

$$G_T(1) = G(G_T(1)) \tag{32}$$

*when $\eta = 1$, it's obvious that $G_T(1) = \mathbb{P}(T < \infty) = 1$. When $\eta < 1$, however, $G_T(1) < 1$ and is the fixed point of the generating function $G(s)$. Since in supercritical phase, the fixed point of $G(s)$ in $[0, 1)$ exists and is uniquely the extinction probability η, we conclude that $G_T(1) = \mathbb{P}(T < \infty) = \eta$. This provides **another perspective understanding the extinction probability**.*