Generating function #2

학술단체협의회 Seminar

임세진

Review..

- $A(x) = \sum_{n \geq 0} a_n x^n = \sum_{n=0}^{\infty} a_n x^n$

: $A(x)$ is a generating function of sequences of numbers a_n.

- If a generating function or its relations are well-known, we can infer the sequences of numbers.
- Examples..
  - Cauchy product formula
    - If $\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{m=0}^{\infty} b_m\right) = \sum_{k=0}^{\infty} c_k,\ \ c_k = \sum_{l=0}^{k} a_l b_{k-l}$.
    - For generating function case,
      - $\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{m=0}^{\infty} b_m x^m\right) = \sum_{k=0}^{\infty} c_k x^k,\ \ c_k = \sum_{l=0}^{k} a_l b_{k-l}$
  - If $A(x) = \sum_{n \geq 0} a_n x^n$, $C(x) = \sum_{k \geq 0} A(x)^k$. $C(x) = \frac{1}{1 - A(x)}$.

Let’s solve the real problem..

- 상황 : A와 B는 각각 자신만의 규칙을 세워서 메신저를 보낸다.
    - A는 상대방이 메세지를 보낸 이후 p의 확률로 메세지를 보낸다.
    - B는 상대방이 메세지를 보내고 될 수 있으면 빨리 답변을 하려고 노력한다.
    - A가 답변하는 시간은 푸아송 프로세스에서의 inter-event time 분포를 따른다.
    - B가 답변하는 시간은 exponent가 $\gamma + 1$인 Pareto 분포를 따른다.

- 문제 1 : A와 B가 서로 연락을 할 때, 이 연락을 한번 주고 받는 평균적인 주기를 구하시오.
- 문제 2 : “연락이 활발하다”는 것은 A와 B가 $\frac{1}{p}$분 이내에 연락을 연속해서 주고 받는 것이라 하자. 매 번 연락을 주고 받을 때마다 A와 B가 연락이 활발할 확률을 구하시오.

- 단 모든 단위 시간은 1분이다.

Relations of model

- 상황 : A와 B는 각각 자신만의 규칙을 세워서 메신저를 보낸다.
  - A는 상대방이 메세지를 보낸 시각으로부터 τ분 내외로 답변을 한다.
  - B는 상대방이 메세지를 보내고 최대한 빨리 답변을 하려고 노력한다.
  - A가 답변하는 시간은 푸아송 프로세스에서의 inter-event time 분포를 따른다.
  - B가 답변하는 시간은 exponent가 $\gamma + 1$인 Pareto 분포를 따른다.

- a_n : A가 n번째 답변하는 시각
- b_n : B가 n번째 답변하는 시각

- $a_n = b_{n-1} + \eta_a$, η_a : exponential 분포 $pe^{-p\eta_a}$ 를 따르는 랜덤 변수 (p : 매 시각 event가 발생할 확률)
- $b_n = a_{n-1} + \eta_b$, η_b : exponent가 $\gamma + 1$인 Pareto 분포를 따르는 랜덤 변수

문제 1

- $a_n = b_{n-1} + \eta_a$, η_a : exponential 분포 $pe^{-p\eta_a}$ 를 따르는 랜덤 변수 (p : 매 시각 event가 발생할 확률)
- $b_n = a_{n-1} + \eta_b$, η_b : exponent가 $\gamma + 1$인 Pareto 분포, $\gamma \eta_b^{-\gamma-1}$를 따르는 랜덤 변수

- 문제 1에서 알아내고 싶은 것

- $\tau_n = a_n + b_n$일 때, $\tau_n - \tau_{n-1}$의 평균 값 (first moment).
- $\tau_n - \tau_{n-1} = \eta_a + \eta_b$

- $\langle \tau_n - \tau_{n-1} \rangle = \langle \eta_a \rangle + \langle \eta_b \rangle = \frac{1}{p} + \frac{\gamma}{\gamma - 1}$, $\gamma > 1$

문제 2 : Survival probability

- n번째 연락을 주고 받을 때마다 A와 B가 연락이 활발할 확률은 결국 τ_n이 $\frac{1}{p}$분 이내로 생존할 확률을 의미한다.
- $\tau_n - \tau_{n-1} = \eta_a + \eta_b$
- τ_n이 $\frac{1}{p}$분 이내로 생존할 확률 : $q_-(n)$
- $q_-(n)$ 의 생성함수 : $\tilde{q}_-(z) = \sum_{n=0}^{\infty} q_-(n) z^n$
- Sparre-Andersen Theorem (랜덤 워커의 생존확률의 생성함수와 위치 분포 간의 관계식)
- $\tilde{q}_-(z) = \exp\left(\sum_{n=1}^{\infty} \frac{z^n}{n} \mathrm{Proba}[x_n < x_0]\right)$

문제 2 : Survival probability

- $\tau_n - \tau_{n-1} = \eta_a + \eta_b$
- $\tau_n - \tau_{n-1}$이 $\frac{1}{p}$분 이내로 생존할 확률 (τ_n 이 이 $\frac{n}{p}$분 이내로 생존할 확률) : $q_-(n)$
- Sparre-Andersen Theorem (랜덤 워커의 생존확률의 생성함수와 위치 분포 간의 관계식)
- $\tilde{q}_-(z) = \exp\left(\sum_{n=1}^{\infty} \frac{z^n}{n} \mathrm{Proba}[x_n < x_0]\right)$

- 문제 상황에서 $\tilde{q}_-(z)$는 다음과 같이 나타내어진다.
- $\tilde{q}_-(z) = \exp\left(\sum_{n=1}^{\infty} \frac{z^n}{n} \mathrm{Proba}\left[\tau_n < \frac{n}{p}\right]\right) = \exp\left(\sum_{n=1}^{\infty} \frac{z^n}{n} \left(\mathrm{Proba}\left[\eta_a < \frac{n}{p}\right] + \mathrm{Proba}\left[\eta_b < \frac{n}{p}\right]\right)\right)$
- $\mathrm{Proba}\left[\eta_a < \frac{n}{p}\right] = \int_{-\infty}^{n/p} pe^{-p\eta_a}\, d\eta_a = 1 - \int_{n/p}^{\infty} pe^{-p\eta_a}\, d\eta_a = 1 - e^{-n}$
- $\mathrm{Proba}\left[\eta_b < \frac{n}{p}\right] = \int_{-\infty}^{n/p} \gamma\eta_b^{-\gamma-1}\, d\eta_b = 1 - \int_{n/p}^{\infty} \gamma\eta_b^{-\gamma-1}\, d\eta_b = 1 - \left(\frac{n}{p}\right)^{-\gamma}, \quad \gamma > 0$

문제 2 : Survival probability

- 문제 상황에서 $\tilde{q}_-(z)$는 다음과 같이 나타내어진다.
- $\tilde{q}_-(z) = \exp\left(\sum_{n=1}^{\infty} \frac{z^n}{n} \text{Proba}\left[\tau_n < \frac{n}{p}\right]\right) = \exp\left(\sum_{n=1}^{\infty} \frac{z^n}{n}\left(\text{Proba}\left[\eta_a < \frac{n}{p}\right] + \text{Proba}\left[\eta_b < \frac{n}{p}\right]\right)\right)$
- $\text{Proba}\left[\eta_a < \frac{n}{p}\right] = \int_{-\infty}^{n/p} pe^{-p\eta_a}\, d\eta_a = 1 - \int_{n/p}^{\infty} pe^{-p\eta_a}\, d\eta_a = 1 - e^{-n}$
- $\text{Proba}\left[\eta_b < \frac{n}{p}\right] = \int_{-\infty}^{n/p} \gamma\eta_b^{-\gamma-1}\, d\eta_b = 1 - \int_{n/p}^{\infty} \gamma\eta_b^{-\gamma-1}\, d\eta_b = 1 - \left(\frac{n}{p}\right)^{-\gamma}, \quad \gamma > 0$

- $\tilde{q}_-(z) = \exp\left(\sum_{n=1}^{\infty} \frac{z^n}{n}\left(1 - e^{-n} + 1 - \left(\frac{n}{p}\right)^{-\gamma}\right)\right) = \exp\left(\sum_{n=1}^{\infty}\left(\frac{2z^n}{n} - \frac{z^n}{n}\left(e^{-n} + \left(\frac{n}{p}\right)^{-\gamma}\right)\right)\right)$
- $\tilde{q}_-(z) = \exp(-2\ln(1-z))\exp\left(\sum_{n=1}^{\infty}\left(-\frac{z^n}{n}\left(e^{-n} + \left(\frac{n}{p}\right)^{-\gamma}\right)\right)\right) = \frac{1}{(1-z)^2}\exp\left(\sum_{n=1}^{\infty}\left(-\frac{z^n}{n}\left(e^{-n} + \left(\frac{n}{p}\right)^{-\gamma}\right)\right)\right)$

문제 2 : Survival probability

- $\tilde{q}_-(z) = \exp(-2\ln(1-z))\exp\left(\sum_{n=1}^{\infty}\left(-\frac{z^n}{n}\left(e^{-n} + \left(\frac{n}{p}\right)^{-\gamma}\right)\right)\right) = \frac{1}{(1-z)^2}\exp\left(\sum_{n=1}^{\infty}\left(-\frac{z^n}{n}\left(e^{-n} + \left(\frac{n}{p}\right)^{-\gamma}\right)\right)\right)$

- $\exp\left(\sum_{n=1}^{\infty}\left(-\frac{z^n}{n}e^{-n}\right)\right) = \exp(-1 + \ln(e - z))$
- $\exp\left(\sum_{n=1}^{\infty}\left(-\frac{z^n}{n}\left(\frac{n}{p}\right)^{-\gamma}\right)\right) = \exp\left(-p^{\gamma}\mathrm{Li}_{\gamma+1}(z)\right)$

- $\tilde{q}_-(z) = \frac{1}{1-z}\frac{1}{e-z}\exp\left(-p^{\gamma}\mathrm{Li}_{\gamma+1}(z)\right), \quad \gamma > 0$

문제 2 : Survival probability

- A relation of generating function #1
  - $f_n = 1 + \sum_{m=1}^{n} g_m, \quad F(x) = \sum_{n=0}^{\infty} f_n x^n, \quad G(x) = \sum_{n=0}^{\infty} g_n x^n.$
  - $g_n = f_n - f_{n-1}, \quad G(x) = \sum_{n=0}^{\infty} g_n x^n = \sum_{n=0}^{\infty} (f_n - f_{n-1}) x^n. \quad f_{-1} = 1.$
  - $G(x) = F(x) - xF(x).$
  - $F(x) = \frac{1}{1-x} G(x).$

- A relation of generating function #2
  - $f_n = \sum_{m=1}^{n} \frac{1}{e^{n-m}} g_m, \quad F(x) = \sum_{n=0}^{\infty} f_n x^n, \quad G(x) = \sum_{n=0}^{\infty} g_n x^n.$
  - $g_n = e f_n - f_{n-1}, \quad G(x) = \sum_{n=0}^{\infty} g_n x^n = \sum_{n=0}^{\infty} (e f_n - f_{n-1}) x^n. \quad f_{-1} = 0.$
  - $G(x) = eF(x) - xF(x).$
  - $F(x) = \frac{1}{e-x} G(x).$

- $\tilde{q}_-(z) = \frac{1}{1-z} \frac{1}{e-z} \exp\left(-p^\gamma \mathrm{Li}_{\gamma+1}(z)\right), \quad \gamma > 1$

문제 2 : Survival probability

- $\tilde{q}_-(z) = \frac{1}{1-z}\frac{1}{e-z}\exp\left(-p^\gamma \mathrm{Li}_{\gamma+1}(z)\right), \quad \gamma > 1$

- $\exp\left(-p^\gamma \mathrm{Li}_{\gamma+1}(z)\right) \approx (?\,)$

- $\tilde{q}_-(z) = \frac{1}{1-z}\tilde{F}(z) = \frac{1}{1-z}\frac{1}{e-z}\tilde{G}(z)$
- $F(n) = \sum_{m=1}^{n}\frac{G(m)}{e^{n-m}}, \qquad \tilde{G}(z) = \exp\left(-p^\gamma \mathrm{Li}_{\gamma+1}(z)\right) \approx p^{-\gamma}n^{1+\gamma}z^{-n},$
- $q_-(n) = 1 + \sum_{m=1}^{n} F(m) = 1 + \sum_{m=1}^{n}\sum_{k=1}^{m}\frac{G(k)}{e^{m-k}}$

문제 2 : Survival probability

- “연락이 활발하다”는 것은 A와 B가 $\frac{1}{p}$분 이내에 연락을 연속해서 주고 받는 것이라 하자.
- A와 B가 연락을 시작하고 n 번째 연락까지 “연락이 활발할” 확률 $q_{-}(n)$은 다음과 같다.
- $q_{-}(n) = 1 + \sum_{m=1}^{n} F(m) = 1 + \sum_{m=1}^{n} \sum_{k=1}^{m} \frac{G(k)}{e^{m-k}}$
- 실제로(?) 위와 같은 과정으로 연락을 주고 받았을 때의 결과와 계산 결과가 같은지 알고 싶으므로, 시뮬레이션을 해보았다.

시뮬레이션 결과와 계산결과 비교

- 문제 1

Dashed line
: $\langle \tau_n - \tau_{n-1} \rangle = \frac{1}{p} + \frac{\gamma}{\gamma-1}, \quad \gamma > 1$

Scatter plot
: Simulation results