Erdos-Renyi Poisson Process

- Binomial distribution and Poisson approximation
- Erdos-Renyi Poisson Process
- Suspiciousness Metrics

- Binomial distribution:
  - $1 = (p+q)^n = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x}$ $(p = 1-q)$.
  - $X \sim \text{Binomial}(n, p)$:
    $P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \dots, n$. $\begin{cases} E(X) = np \\ \text{Var}(X) = np(1-p) \end{cases}$

- Poisson distribution:
  - $e^{\lambda} = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \Rightarrow 1 = \sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!}$,
  - $X \sim \text{Poisson}(\lambda)$:
    $P(X=x) = e^{-\lambda} \frac{\lambda^x}{x!}$, $x = 0, 1, \dots$. $\begin{cases} E(X) = \lambda \\ \text{Var}(X) = \lambda \end{cases}$

- Both Binomial and Poisson are "counting processes".
  - discrete time
  - Binomial: finite time span;
    Poisson: infinite time span.

- Poisson approximation to Binomial distribution:
  - Suppose $np \to \lambda$ as $n \to \infty$. (np: mean of Binomial)
  - $P_B(X=x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$
    $\to \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x} = \frac{n(n-1)\cdots(n-x+1)}{x!} \frac{\lambda^x}{n^x} \left(1-\frac{\lambda}{n}\right)^{n-x}$
    $= \frac{n}{n} \frac{(n-1)}{n} \cdots \frac{n-x+1}{n} \frac{\lambda^x}{x!} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-x}$
    $\to e^{-\lambda} \frac{\lambda^x}{x!} = P_P(X=x)$.

- Additivity of independent Poisson r.v.'s:

$X_i \overset{ind}{\sim} \text{Poisson}(\lambda)$, $i = 1, \cdots, n$

$Y = \sum_{i=1}^{n} X_i \sim \text{Poisson}(n\lambda)$.

- CrossSpot Algorithm ~~with~~ under Erdos-Renyi Poisson Process:

- Given an n-length subvector $[X_{i1}, \cdots, X_{in}]$ in the N-length vector $[X_1, \cdots, X_N]$, $c \frac{n}{N}$

overall density $\lambda = \frac{C}{N}$ ↔ # tweets → # IPs

$Y_n = \sum_{j=1}^{n} X_{ij} \sim \text{Poisson}(n\lambda)$, since $X_{ij} \sim \text{Poisson}(\lambda)$; $\lambda = \frac{C}{N}$

$P(Y_n = c) = \frac{C^c}{c!}\left(\frac{n}{N}\right)^c e^{-\frac{Cn}{N}}$.

- For some subvector with (n, c), define the suspiciousness metric:

$f(n, c, N, C) = -\log P(Y_n = c)$

- Negative log likelihood: the larger the more suspicious.
- Given the overall density $\lambda = \frac{C}{N}$, estimate the probability of the subvector having density $\frac{c}{n}$.
- Further,

$$\begin{aligned} f(n, c, N, C) &= -\log P(Y_n = c) \\ &= -\log \frac{C^c}{c!}\left(\frac{n}{N}\right)^c e^{-\frac{Cn}{N}} \\ &= -\log \frac{C^c}{c!} \cdot e^{-c\log\frac{n}{N}} + C\frac{n}{N} \\ &= -c\log C + \log c! - c\log\frac{n}{N} + C\frac{n}{N}. \end{aligned}$$

By Stirling's Formula, $\log c! \approx c\log c - c$,

$$\begin{aligned} f(n, c, N, C) &= -c\log C + c\log c - c - c\log\frac{n}{N} + C\frac{n}{N} \\ &= c\left(\log\frac{c}{C} - 1\right) - c\log\frac{n}{N} + C\frac{n}{N}. \end{aligned}$$

- To obtain $\hat{f}(n, \rho, N, \lambda)$, $\rho = \frac{c}{n}$, $\lambda = \frac{C}{N}$, ~~just~~ just substitude (c, C) by (ρ, λ) in $f(n, c, N, C)$.

$$\begin{aligned} \hat{f}(n, \rho, N, \lambda) &= n\rho\left(\log\frac{n\rho}{N\lambda} - 1\right) - n\rho\log\frac{n}{N} + N\lambda\frac{n}{N} \\ &= n\rho\log\frac{\rho}{\lambda} - n\rho + n\lambda = n\left(-\rho + \rho\log\frac{\rho}{\lambda}\right) \\ &= n D_{KL}(\rho \| \lambda). \end{aligned}$$

• Kullback-Leibler Divergence:

$$D_{KL}(P|\lambda) = E_P\left(\log \frac{P}{\lambda}\right) = \sum_i P_P(i) \log \frac{P_P(i)}{P_\lambda(i)}.$$

- Suppose $Y_n \sim Poisson(\rho)$, $Y_n' \sim Poisson(\lambda)$

$$\begin{aligned} D_{KL}(P|\lambda) &= \sum_{c=0}^{\infty} e^{-\rho} \frac{\rho^c}{c!} \log\left[e^{-\rho}\frac{\rho^c}{c!} / e^{-\lambda}\frac{\lambda^c}{c!}\right] \\ &= \sum_{c=0}^{\infty} e^{-\rho}\frac{\rho^c}{c!}(-\rho + c\log\rho + \lambda - c\log\lambda) \\ &= \lambda - \rho + \sum_{c=0}^{\infty}(\log\frac{\rho}{\lambda})\, c \cdot e^{-\rho}\frac{\rho^c}{c!} \\ &= \lambda - \rho + \log(\frac{\rho}{\lambda})\, E(Y) \\ &= \lambda - \rho + \rho\log\frac{\rho}{\lambda}. \end{aligned}$$

- Hence we obtain the suspiciousness metric, $\hat{P}(n, \rho, N, \lambda)$.