On the asymptotic behaviour of the number of renewals via translated Poisson

Aihua Xia (夏爱华)

School of Mathematics and Statistics

The University of Melbourne, VIC 3010

20 July, 2018

The 14th Workshop on Markov Processes and Related Topics

Some facts about renewal processes

- Let τ and τ_i's be independent and identically distributed inter-renewal times with mean μ and finite variance σ^2, define $S_n := \sum_{i=1}^{n} \tau_i$ and $S_0 = \sum_{i=1}^{0} \tau_i := 0$
- Let $N(t)$ be the number of renewals up to time t, that is,

$$N(t) = \max\{n : \; S_n \leq t < S_{n+1}\},$$

- $\mathbb{E}N(t) \sim \frac{t}{\mu}$ and $\mathrm{Var}(N(t)) \sim \frac{t\sigma^2}{\mu^3}$ as $t \to \infty$
  - Feller (1968), Vol 1, pp. 340–341; Vol. 2, p. 372.
- $W(t) := \frac{N(t) - t/\mu}{\sqrt{t\sigma^2/\mu^3}}$ is approximated by a standard normal distribution.
- **Question:** how fast?

Who cares?

- Renewal theory is one of indispensable topics in introductory courses of random processes.
- The behaviour of regenerative events is of significant interest in probability theory and related areas.

Speed of convergence?

- For the Kolmogorov distance, it must be easy!

$$\mathbb{P}(N(t) \geq n) = \mathbb{P}(S_n \leq t).$$

  - $S_n \sim N(n\mu, n\sigma^2)$.
  - Most likely values of $N(t)$ are within a few standard deviations of the mean, but the approximate normal is $N\left(\frac{t}{\mu}, \frac{t\sigma^2}{\mu^3}\right)$.
  - Do you still think it is easy?

Literature?

- Englund (1980):

$$\sup_n \left| \mathbb{P}(N(t) < n) - \Phi\left(\frac{(n\mu - t)\sqrt{\mu}}{\sigma\sqrt{t}}\right) \right| \leq 4\left(\frac{\gamma}{\sigma}\right)^3 \sqrt{\frac{\mu}{t}},$$

  where $\gamma^3 = \mathbb{E}(|\tau - \mu|^3)$.

  – $\Phi\left(\frac{(n\mu - t)\sqrt{\mu}}{\sigma\sqrt{t}}\right) = \mathbb{P}(X_t < n)$, with $X_t \sim N(t/\mu, t\sigma^2/\mu^3)$.

  – $d_{\mathrm{K}}(Q_1, Q_2) := \sup_{u \in \mathbb{R}} |Q_1(-\infty, u] - Q_2(-\infty, u]|$.

  – Under the Kol. distance, it is done!

  – Can someone use Stein's method to prove this bound?

More on literature

- Omey and Vesilo (2011): suppose the characteristic function of τ is integrable, then

$$\sup_t \left| \frac{\sigma\sqrt{n}}{\mu} \mathbb{P}(N(t) = n+1) - \Phi\left(\frac{t - n\mu}{\sigma\sqrt{n}}\right) \right| = o(1).$$

Moreover, if τ has finite third moment, then

$$\sup_t \left| \frac{\sigma\sqrt{n}}{\mu} \mathbb{P}(N(t) = n+1) - \Phi\left(\frac{t - n\mu}{\sigma\sqrt{n}}\right) \right| = O\left(n^{-1/2}\right).$$

- $\mathbb{P}(N(t) \in A)$ for any set $A \subset \mathbb{Z}_+$? Discretised normal?

Metrics?

How about the probabilities of values in any set?

- Total variation distance: for any probability measures Q_1 and Q_2 on $\mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}$,

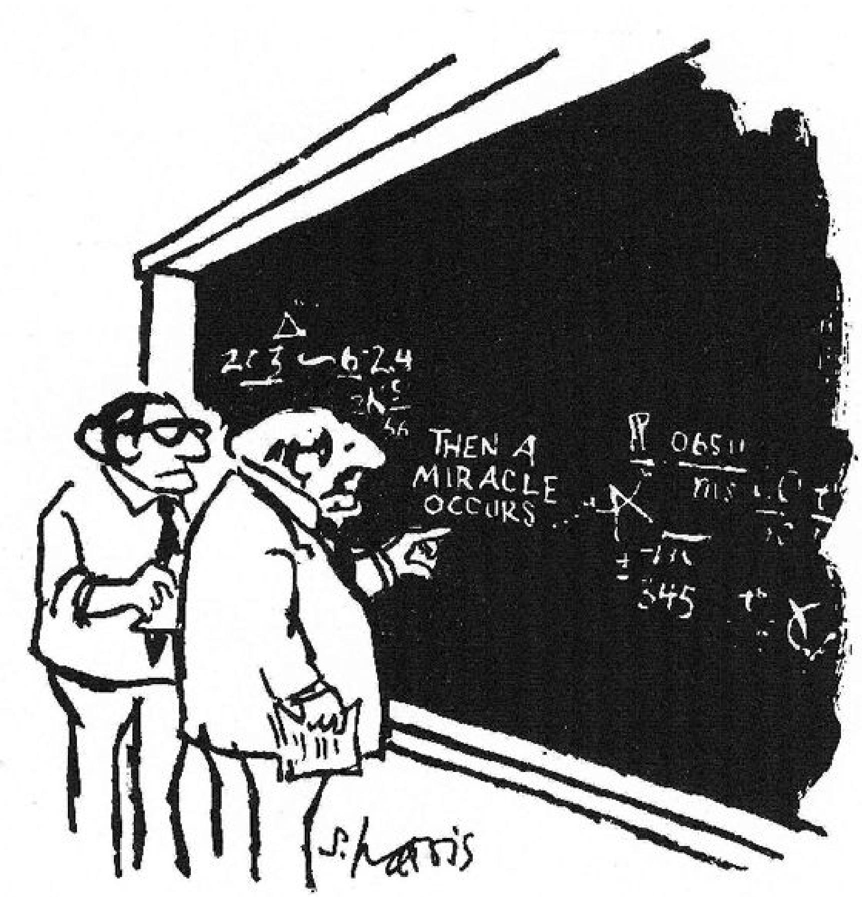
$$d_{\mathrm{TV}}(Q_1, Q_2) := \sup_{A \subset \mathbb{Z}} |Q_1(A) - Q_2(A)|.$$

- The Wasserstein distance: Q_1 and Q_2 on $\mathbb{R}$

$$d_{\mathrm{W}}(Q_1, Q_2) \quad := \quad \sup_{f \in \mathcal{F}_{\mathrm{Lip}(1)}} \left| \int f dQ_1 - \int f dQ_2 \right|,$$

where $\mathcal{F}_{\mathrm{Lip}(1)} = \{f : \mathbb{R} \mapsto \mathbb{R} : \ |f(x) - f(y)| \leq |x - y| \text{ for all } x, y \in \mathbb{R}\}$.

By handwaving: it should work!



"I think you should be more explicit here in step two."

- The cartoon is by Sidney Harris

How to do it?

- Characteristic functions: Englund (1980) uses
  - Berry–Esseen theorem for iid random variables,
  - Some technical adjustment to estimate the gaps amongst various normal distributions with different parameters.
- Coupling: challenging!
  - yet to see one bound using pure coupling but with the right order.
- Stein's method
  - No such work, have asked various people, from old to young!

Likely approximate distributions

- $N(t)$ is non-negative integer-valued, so if we consider d_{TV}, the approximate distribution must be non-negative integer-valued.
- If $\tau \sim \exp$, then $N(t)$ is Poisson (fixed point).
- If we use discretised normal, under what conditions can we get $N(t) \sim$ discretised normal for moderate t?
- Poisson(λ) is close to normal when λ is large.

Poisson is not enough!

- mean=variance, lack of flexibility.
- In general, for $a > 0$ and an integer b, a translated Poisson distribution is defined as $P_{a,b} := \mathrm{Pn}(a) * \delta_b$ (Röllin (2007)).
  - If N is close to a normal distribution, then it must be close to a translated Poisson distribution.
  - Suitable in the total variation distance so for all possible sets rather than intervals of the form $(-\infty, x]$.

Discretized normal

- $N^d(a+b, a)$ (Fang (2014)): having probability mass function at integer $z \in \mathbb{Z}$ as

$$\int_{z-1/2}^{z+1/2} \frac{1}{\sqrt{2\pi}a} e^{-\frac{(x-(a+b))^2}{2a^2}} dx.$$

  – Discretised normal can do the same job.

  – It does not offer the same interpretation as a translated Poisson.

Discrete CLT around $N(t)$

- Under the Kolmogorov distance, yes!
- Under d_{TV}: NO discrete CLT for $N(t)$!
  - If $\mathbb{P}(\tau = 1) = \mathbb{P}(\tau = 3) = 1/2$, then

$$\liminf_{t \to \infty} \min_{a,b} d_{\mathrm{TV}}(\mathscr{L}(N(t)), P_{a,b}) > 0.$$

  - Under what conditions can we have discrete CLT for $N(t)$?

The speed: $d_{\mathbf{W}}$

Let $a = t\sigma^2/\mu^3$ and $b = \lfloor \frac{t}{\mu}(1 - \frac{\sigma^2}{\mu^2})\rfloor$, where $\lfloor y \rfloor$ denotes the integer part of y. If $\mathbb{E}(\tau^3) < \infty$, then

$$d_{\mathrm{W}}(N(t), P_{a,b}) = O(1).$$

- If we standardise, then d_{W} is of order $O(t^{-1/2})$, hence Peccati, Solé, Taqqu & Utzet (2010) ensures

$$d_{\mathrm{W}}\left(\frac{N(t) - \mathbb{E}N(t)}{\sqrt{\mathrm{Var}(N(t))}}, N(0,1)\right) = O(t^{-1/2}).$$

The speed: $d_{\mathbf{TV}}$

- The Lebesgue decomposition theorem: for any distribution function G on $\mathbb{R}$ can be represented as

$$G = (1 - \alpha_G)G_s + \alpha_G G_a.$$

- A distribution function G on $\mathbb{R}$ is said to be non-singular if $\alpha_G > 0$.

The speed: $d_{\mathbf{TV}}$ (cont)

- a and b are as above. Assume $\mathbb{E}(\tau^3) < \infty$. Either of the following conditions ensures

$$d_{\mathrm{TV}}(N(t), P_{a,b}) = O\left(t^{-1/2}\right).$$

  - $0 < F(0) < 1$.
  - F is non-singular.
  - F is singular with $\mathrm{supp}(F) \cap \{c, 2c, 3c, \dots\} \neq \emptyset$ and $d_{\mathrm{TV}}(\tau, \tau + c) < 1$.

The Stein-Chen method

- $X \sim \mathrm{Pn}(\lambda)$ iff $\mathbb{P}(X+1=k) = \frac{k\mathbb{P}(X=k)}{\lambda}$, $k \in \mathbb{Z}_+$ iff $\mathbb{E}[\lambda g(X+1) - Xg(X)] = 0$ for a suitable class of g.

- Stein's identity for $\mathrm{Pn}(\lambda)$:

$$\lambda g(i+1) - ig(i) = f(i) - \mathrm{Pn}(\lambda)(f)$$

  for all suitable functions f.

  - $\mathrm{Pn}(\lambda)(f) = \mathbb{E}f(X)$ with $X \sim \mathrm{Pn}(\lambda)$.
  - By solving the equation recursively, g can be written in terms of f.

Stein's identity for $P_{a,b}$

Set $g(i) = 0$ for $i \le -1$ and write Stein's identity

$$\lambda g(i-b+1) - (i-b)g(i-b) = f(i-b) - \mathrm{Pn}(a)(f)$$

for all suitable functions f.

- There is a truncation problem at around $i = -1$.
- Write $\tilde{g}(j) := g(j-b)$ so that

$$\lambda \tilde{g}(i+1) - (i-b)\tilde{g}(i) \approx f(i-b) - \mathrm{Pn}(a)(f).$$

- Consider the stationary case $\mathcal{N}$ first: the first renewal needs adjustment.

Difficulty? a quick run of the proof

- By the Stein's identity for $P_{a,b}$:

$$
\begin{aligned}
& \mathbb{E}f(\mathcal{N}(t)-b)-\mathrm{Pn}(a)(f) \\
\approx \quad & a\mathbb{E}\tilde{g}(\mathcal{N}(t)+1)+b\mathbb{E}\tilde{g}(\mathcal{N}(t))-\mathbb{E}[\tilde{g}(\mathcal{N}(t))\mathcal{N}(t)]
\end{aligned}
$$

  – Need to work on $\mathbb{E}[\tilde{g}(\mathcal{N}(t))\mathcal{N}(t)]$.

From size biasing to Palm

- For a nonnegative integer-valued random variable X having positive finite mean μ, we consider $h(\cdot) = \delta_{\{k\}}(\cdot)$, then

$$\frac{\mathbb{E}[h(X)X]}{\mathbb{E}X} = \frac{k\mathbb{P}(X=k)}{\mu},$$

giving *size biased distribution*

$$\mathbb{P}(X^s = k) = \frac{k\mathbb{P}(X=k)}{\mu}.$$

From size biasing to Palm – cont

- $X^s \overset{\mathrm{d}}{=} X + 1$ iff $X \sim \mathrm{Pn}(\mu)$: the Stein-Chen method for Poisson approximation.
- Size biasing appears in various sampling contexts, e.g., in random digit dialing, it is proportionally more likely to dial households with more telephones than households with fewer phones.
- Size biasing of $\mathcal{N}(t)$ does not offer enough information.

From size biasing to Palm – cont

- We can expand

$$\mathbb{E}[\tilde{g}(\mathcal{N}(t))\mathcal{N}(t)] = \mathbb{E}\int_0^t \tilde{g}(\mathcal{N}(t))\mathcal{N}(d\alpha).$$

- It is possible to consider Radon-Nikodym derivative

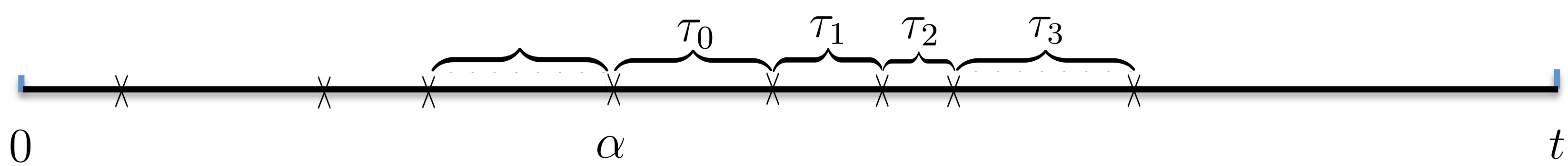
$$\frac{\mathbb{E}[\tilde{g}(\mathcal{N}(t))\mathcal{N}(d\alpha)]}{\mathbb{E}[\mathcal{N}(d\alpha)]} =: \mathbb{E}\tilde{g}(\mathcal{N}_\alpha(t)).$$

  – $\mathcal{N}_\alpha(t)$ is called a *Palm process* of $\mathcal{N}$ at α, its distribution is called the *Palm distribution*.

  – **Fact**: Palm distribution is the process version of size biasing.

The Palm for renewal process

- The Palm process at α: given there is a renewal at α, how the remaining part of the renewal process looks like?



- If $\mathcal{N}$ is a Poisson process, then $\mathcal{N}_\alpha \stackrel{\mathrm{d}}{=} \mathcal{N} + \delta_\alpha$: one additional observer at α, the rest is the same.

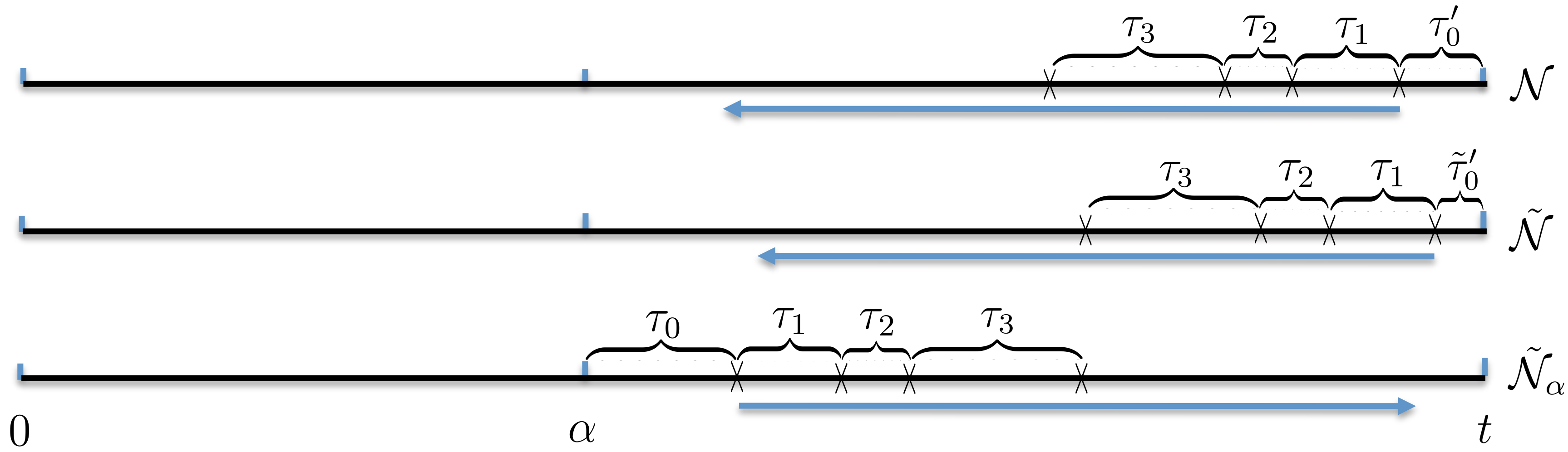
A quick run of the proof: cont

- Since $\mathbb{E}\mathcal{N}(d\alpha) = \frac{d\alpha}{\mu}$, we have

$$
\begin{aligned}
& \mathbb{E}f(\mathcal{N}(t)-b) - \mathrm{Pn}(a)(f) \\
\approx\ & a\mathbb{E}\tilde{g}(\mathcal{N}(t)+1) + b\mathbb{E}\tilde{g}(\mathcal{N}(t)) - \mathbb{E}\int_0^t \tilde{g}(\mathcal{N}(t))\mathcal{N}(d\alpha) \\
=\ & a\mathbb{E}\tilde{g}(\mathcal{N}(t)+1) + b\mathbb{E}\tilde{g}(\mathcal{N}(t)) - \frac{1}{\mu}\int_0^t \mathbb{E}\tilde{g}(\mathcal{N}_\alpha(t))d\alpha \\
=\ & a\mathbb{E}\Delta\tilde{g}(\mathcal{N}(t)) + (a+b)\mathbb{E}\tilde{g}(\mathcal{N}(t)) - \frac{1}{\mu}\int_0^t \mathbb{E}\tilde{g}(\mathcal{N}_\alpha(t))d\alpha \\
\approx\ & a\mathbb{E}\Delta\tilde{g}(\mathcal{N}(t)) - \frac{1}{\mu}\int_0^t \mathbb{E}[\tilde{g}(\mathcal{N}_\alpha(t)) - \tilde{g}(\mathcal{N}(t))]d\alpha,
\end{aligned}
$$

  - Δ is the forward difference operator and $a+b \approx t/\mu$,
  - $\mathcal{N}_\alpha$ is the Palm process of $\mathcal{N}$ at α.

A coupling that works!



Still unsatisfactory!

- Main difficulty in d_{TV}: $d_{\mathrm{TV}}(N(t), P_{a,b})$ has the same speed as $d_{\mathrm{TV}}(N(t), N(t)+1)$.
- What is $d_{\mathrm{TV}}(N(t), N(t)+1)$?
- If $\mathbb{P}(\tau = 1) = \mathbb{P}(\tau = 3) = 1/2$, then

$$\liminf_{t\to\infty} d_{\mathrm{TV}}(N(t), N(t)+1) \geq \sqrt{\frac{3}{8\pi}}e^{-8+O(t^{-1/2})}.$$

- Most regenerative events in Markov processes (both continuous and discrete time) have $d_{\mathrm{TV}}(N(t), P_{a,b}) = O(t^{-1/2})$.

Problems for further consideration

- If $\operatorname{supp}(F)$ contains an interval, then $d_{\mathrm{TV}}(N(t), P_{a,b}) = O(t^{-1/2})$.
- $d_{\mathrm{TV}}(N(t), P_{a,b}) = o(1)$ iff $d_{\mathrm{TV}}(N(t), P_{a,b}) = O(t^{-1/2})$.

Take home messages

For the distribution of the number of renewals approximated by a suitable translated Poisson (or discretized normal):

- d_{K}: order $O(t^{-1/2})$,
- d_{W}: $O(1)$,
- d_{TV}: $O(t^{-1/2})$ in most cases,
- the constants are too big and complicated.

Thank you!