ADVANCED STOCHASTIC PROCESSES - 12° LECTURE

2. Chernoff bounds. They are obtained as an application of the exponential Markov inequality for r.v.'s X having finite moment generating function:

1. $\mathbb{P}(X \geq a) \overset{\forall t>0}{=} \mathbb{P}(e^{Xt} \geq e^{at}) \leq e^{-at} \cdot \mathbb{E}(e^{Xt})$
2. $\mathbb{P}(X \leq a) \underset{\forall t<0}{=} \mathbb{P}(e^{Xt} \geq e^{at}) \leq e^{-at} \mathbb{E}(e^{Xt})$

↳ Markov inequality on e^{Xt}

Then, minimizing over $t>0$ in 1. (and over $t<0$ in 2.), we get

$$\mathbb{P}(X \geq a) \leq \inf_{t>0} \{e^{-at} \cdot \mathbb{E}(e^{Xt})\} = e^{-\sup_{t>0}\{at - \log \mathbb{E}(e^{Xt})\}}$$

$$\mathbb{P}(X \leq a) \leq \inf_{t<0} \quad '' \quad '' \quad = e^{-\sup_{t<0}\{at - \log \mathbb{E}(e^{Xt})\}}$$

(H) Then, let us consider $(X_k)_{k\in\mathbb{N}}$ iid r.v.'s with $\mathbb{E}(e^{tX_k}) < \infty$ and let $S_n = \sum_{k=1}^{n} X_k$ so that $\mathbb{E}(S_n) = \mu \cdot n$

Remark:

- From LLN, $\forall \varepsilon > 0$: $\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = \mathbb{P}(|S_n - \mu n| > \varepsilon n) \longrightarrow 0$

  $\quad = \mathbb{P}(S_n > (\mu+\varepsilon)n) + \mathbb{P}(S_n < (\mu-\varepsilon)n)$

- From CLT, $\forall \varepsilon > 0$: $\mathbb{P}\left(\left|\frac{S_n - \mu n}{\sqrt{n}}\right| > \varepsilon\right) = \mathbb{P}(|S_n - \mu n| > \varepsilon\sqrt{n}) \longrightarrow 2(1 - \phi(\varepsilon)) \neq 0$

  ϕ: distr. of $N(0,1)$

$\Longrightarrow \forall \varepsilon > 0, \forall \gamma > \frac{1}{2}$: $\mathbb{P}(|S_n - \mu n| > \varepsilon n^{\gamma}) \longrightarrow 0$

↳ Slutsky's Lemma: $\begin{cases} X_n \xrightarrow{d} X \\ Y_n \xrightarrow{P} c \end{cases} \Rightarrow X_n \cdot Y_n \xrightarrow{d} cX$

As a conclusion, the event $\{|S_n - \mu n| > \varepsilon n^{\gamma}\}$ is a rare event or large fluctuation, having asymptotic 0 probability.

How fast does its probability converges to 0?

↓ Under suitable hypotheses, this is exponentially fast ↓

Theorem [Chernoff bounds]

In the setting given in (H), it holds ...

Theorem [Chernoff bounds]

In the setting given in (H), it holds

* if $a > \mu \implies \mathbb{P}(S_n \geq an) \leq e^{-n I_x(a)}$
* if $a < \mu \implies \mathbb{P}(S_n \leq an) \leq e^{-n I_x(a)}$

where $I_x(a) = \sup_{t \in \mathbb{R}} \{ ta - \lg \mathbb{E}(e^{tX_1}) \}$

Hence, $\forall \varepsilon > 0$ $\mathbb{P}(|S_n - \mu n| \geq \varepsilon n) \leq 2 e^{-n \cdot C(\varepsilon)}$

find the value of $C(\varepsilon)$

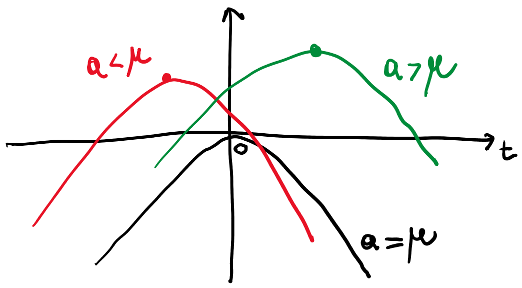
Proof:

1. I is the Legendre transform of $\lg \mathbb{E}(e^{tX_1})$ (review its properties!)
2. Let $g(t) = ta - \lg \mathbb{E}(e^{tX_1})$, so that $I(a) = \sup_{t \in \mathbb{R}} g(t)$

Then it turns out that

* $g(t)$ is concave (from properties of moment gen. fct)
* $g'(t) = a - \dfrac{\mathbb{E}(X_1 e^{tX_1})}{\mathbb{E}(e^{tX_1})} = \Big|_{t=0} a - \mu \begin{cases} < 0 & \text{if } a < \mu \to g \downarrow \text{ in } 0 \\ = 0 & \text{if } a = \mu \to \max \text{ in } 0 \\ > 0 & \text{if } a > \mu \to g \uparrow \text{ in } 0 \end{cases}$

and drawing a qualitative graph of g, we get



$\implies$ • if $a = \mu$, the bounds in Thm hold, but useless as $I(\mu) = 0$

• if $a \neq \mu$, then $I(a) > 0$

In particular $\sup_{t \in \mathbb{R}} g(t) = \begin{cases} \sup_{t>0} g(t) & \text{if } a > \mu \\ \sup_{t<0} g(t) & \text{if } a < \mu \end{cases}$.

3. As an application of the exponential Markov inequality, also recalling (from iid hp.) $\mathbb{E}(e^{tS_n}) = e^{n \lg \mathbb{E}(e^{tX_1})}$

* if $a > \mu$: $\mathbb{P}(S_n \geq an) \leq e^{-n \overset{``}{\underset{t>0}{\sup}} \{ \underbrace{at - \lg \mathbb{E}(e^{tX_1})}_{g(t)} \}} \leq e^{-n I_x(a)}$ (where $\sup_{t>0}$ "= $\sup_{t \in \mathbb{R}}$")

and similarly for $a < \mu$. #

Example

1. Let $S_n \sim Poi(n\lambda)$. $S_n \overset{d}{=} \sum_{k=1}^{n} X_k$, with (X_k) iid $\sim Poi(\lambda)$

Recall that $\mathbb{E}(e^{tX_1}) = e^{-\lambda(1-e^t)}$, $\forall t \in \mathbb{R}$

$$\Rightarrow I_{Poi(\lambda)}(a) = \sup_{t\in\mathbb{R}} \{ta + \lambda(1-e^t)\} = \lambda - a - a\lg\frac{\lambda}{a}$$

$\hookrightarrow$ verify

2. Let $S_n \sim Bin(n,p)$. $S_n \overset{d}{=} \sum_{k=1}^{n} X_k$, with (X_k) iid $\sim Be(p)$

Recall that $E(e^{tX_1}) = pe^t + 1-p = 1-p(1-e^t) \overset{*}{\leq} e^{-p(1-e^t)}$

$$\Rightarrow I^{Bin(n,p)}(a) = a\lg\frac{p}{a} + (1-a)\lg\frac{1-p}{1-a} \quad \text{(verify!)}$$

$$\overset{*}{\leq} I^{Poi(np)}(a)$$

In the context of Galton-Watson trees, we get:

$$\begin{cases} \text{if } \mu < 1 \quad \mathbb{P}(T > n) \leq \mathbb{P}\left(\sum_{k=1}^{n} X_k > n-1\right) = \mathbb{P}\left(\sum_{k=1}^{n} X_k \geq n\right) \leq e^{-n I_x(1)} \\ \text{if } \mu > 1 \quad \mathbb{P}(T \leq n) \leq \sum_{k=0}^{n-1} \mathbb{P}\left(\sum_{j=1}^{k} X_j \leq k\right) \leq \sum_{k=0}^{n-1} e^{-k I_x(1)} = \dfrac{e^{-n I_x(1)}}{1 - e^{I_x(1)}} \end{cases}$$

where $I_x : \mathbb{R} \to \overline{\mathbb{R}}$ s.t. $I_x(x) = \sup_{t\in\mathbb{R}} \{tx - \lg\mathbb{E}(e^{tX_k})\}$

Notice that if $\mu = 1 \Rightarrow I_x(1) = 0$, and the two bounds above become useless.

With some refined computations one can prove that when $\mu = 1$, $\mathbb{P}(T = n) \propto n^{-\frac{3}{2}}$ (and hence $\mathbb{P}(T \geq n) \propto n^{-1/2}$) for $n \gg 1$.

POISSON PROCESS

Def [Counting process]: A process $(N_t)_{t\geq 0}$ with value on $\mathbb{N}_0$ is a counting process if

$N_0 = 0$, $N_s \leq N_t$, $\forall\, s \leq t$ (non-decreasing)

Comment: N_t models the number of events of given type occurred in $[0,t]$ related to some random observation. (arrivals)

In modelizations there are two relevant properties that a counting process may satisfy:

1. (N_t) has independent increments if, $\forall\, 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \dots \leq t_n$, [ordered sequence of times in $\mathbb{R}^+$]

$(N_{t_k} - N_{s_k})_{k=1,\dots,n}$ are independent r.v's

2. (N_t) has stationary increments if the law of an increment only depends on its length:

$$N_{t+s} - N_s \overset{d}{=} N_t - N_0 = N_t \quad , \forall\, t,s \geq 0$$

Def [Poisson Process - homogeneous]

A counting process $(N_t)_{t\geq 0}$ is a (homogeneous) Poisson Process with intensity $\lambda > 0$ (denoted $PP(\lambda)$) if it has indep. and stat. increments s.t

3. $\forall\, s < t$: $N_t - N_s \sim Poi(\lambda \cdot (t-s))$

Remark: Note that $3. \iff 3'.$ $N_t \sim Poi(\lambda t)$ (from stationarity and $N_0 = 0$)

Q: Amplitude of jumps? How many arrivals at same time t?

On one hand $\mathbb{P}(N_h \geq 2) = 1 - e^{-h\lambda} - \lambda h e^{-h\lambda} \xrightarrow[h\to 0]{} 0$. (at most one jump in a infinitesimal interval of time)

But more than this:

The trajectories of $PP(\lambda)$ perform jumps of amplitude 1. In particular:

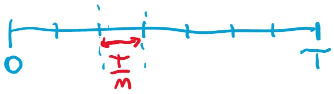
Proposition:

For any given $T \in \mathbb{R}^+$: $\mathbb{P}\left(\exists\, s \in [0,T] : \lim_{h\to 0} (N_{s+h} - N_s) \geq 2 \right) = 0$

Proposition:

For any given $T \in \mathbb{R}^+$: $\mathbb{P}\left(\exists\, s \in [0,T] : \lim_{h \downarrow 0}\left(N_{s+h} - N_s\right) \geq 2\right) = 0$

Proof: For $n \in \mathbb{N}$, consider a partition of $[0,T]$ in n intervals of length T/n : $\left(\left[k\cdot\frac{T}{n} ; (k+1)\cdot\frac{T}{n}\right]\right)_{k=0,\dots,n-1}$

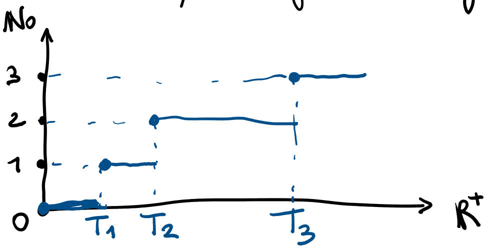


Then, considering the complementary event:

$$\mathbb{P}\left(\bigcap_{k=0}^{n-1}\left\{\left(N_{\frac{(k+1)T}{n}} - N_{\frac{kT}{n}}\right) \leq 1\right\}\right) \underset{\text{indep, stat.}}{=} \mathbb{P}\left(N_{\frac{T}{n}} \leq 1\right)^n = e^{-T\cdot\lambda}\left(1 + \lambda\frac{T}{n}\right)^n$$

$$\xrightarrow[n\to\infty]{} e^{-T\lambda} \cdot e^{-\lambda T} = 1 \qquad \#$$

If we draw the graph $t \to N_t$, we get a trajectory with jumps as the following:



- Arrival and inter-arrival times

For a $PP(\lambda)$, $(N_t)_{t \geq 0}$ let us consider the sequence of arrival times

* $(T_n)_{n \geq 1}$: $T_n := \inf\{t \geq 0 : N_t = n\} \in \overline{\mathbb{R}}_+$ (with $\inf\{\phi\} = +\infty$)

and notice that $\mathbb{P}(T_1 > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t} \Rightarrow T_1 \sim Exp(\lambda)$

Moreover, we consider the seq. of inter-arrival times $(W_n)_{n \geq 1}$ s.t

* $W_n := T_n - T_{n-1}$, $\forall n \geq 2$ so that $T_n = \sum_{k=1}^{n} W_k \quad \forall n \in \mathbb{N}$
$W_1 = T_1$

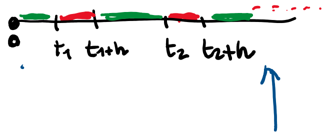
Proposition: $(W_n)_{n \geq 1}$ is a sequence of iid $Exp(\lambda)$ r.v's.

Proof: For a given $n \in \mathbb{N}$, we first compute the joint prob. density of $(T_1, \dots, T_n)$, and then use that $(W_1, \dots, W_n) = F(T_1, \dots, T_n)$ to compute its density and derive the result.
$$F = \begin{pmatrix} 1 & & 0 \\ -1 & 1 & \\ & \ddots & \ddots \\ 0 & -1 & 1 \end{pmatrix}$$

to compute its density and derive the result.

$$F = \begin{pmatrix} -1 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & -1 & 1 \end{pmatrix}$$

- Consider n disjoint intervals of length $h>0$, $[t_k, t_k+h]$ with $t_k+h < t_{k+1}$, $\forall\, k=1,\dots,n$.



Then:

$$\mathbb{P}\Big(\bigcap \{T_k \in (t_k, t_k+h]\}\Big) = \mathbb{P}\big(N_{t_1}=0,\, N_{t_1+h}-N_{t_1}=1,\, N_{t_2}-N_{t_1+h}=0, \dots\big)$$

indep. of increments

$$= e^{-\lambda t_1} \cdot \lambda h e^{-\lambda h}\, e^{-\lambda(t_2-(t_1+h))} \cdot \lambda h e^{-\lambda h} \cdot \ldots \cdot \lambda h e^{-\lambda h}$$

↳ n-th arrival

$$= e^{-\lambda(t_n+h)} \cdot \lambda^n \cdot h^n$$

Then

$$\frac{\mathbb{P}\Big(\bigcap \{T_k \in (t_k, t_k+h]\}\Big)}{h^n} \xrightarrow[h\to 0]{} \lambda^n e^{-\lambda t_n} = p_{T_1,\dots,T_n}(t_1,\dots,t_n)$$

density of $(T_1,\dots,T_n)$

- Since $(W_1,\dots,W_n) = F(T_1,\dots,T_n)$ with $F^{-1} = \begin{pmatrix} 1 & & 0 \\ 1 & 1 & \\ \vdots & & \ddots \\ 1 & \cdots & 1 \end{pmatrix}$, we get that $(W_1,\dots,W_n)$ has density

$$p_{W_1,\dots,W_n}(s_1,\dots,s_n) = \lambda^n e^{-\lambda(s_1+\dots+s_n)} = \prod_{k=1}^{n} \lambda e^{-\lambda s_k} \Rightarrow (W_n) \text{ iid}, \ \mathrm{Exp}(\lambda)$$

#

Corollary: $T_n = \sum_{k=1}^{n} W_k \sim \Gamma(n,\lambda)$

It also holds a converse result of the above Proposition, which indeed provides an alternative construction of $(N_t)_{t\geq 0}$, and shows that $PP(\lambda)$ exists!

Construction: Let $(W_n)_{n\in\mathbb{N}}$ iid rv's $\sim \mathrm{Exp}(\lambda)$

Define $T_n := \sum_{k=1}^{n} W_k$, $\forall\, n \in \mathbb{N}$.

Set $N_t := \#\{n \in \mathbb{N}_0 : T_n \leq t\}$

Theorem: The process $(N_t)_{t\geq 0}$ is a $PP(\lambda)$

Proof (idea): Use $\{N_t = k\} = \{T_k \leq t < T_{k+1}\}$, and the joint density of $(T_1,\dots,T_n)$ to show that. (See Klenke - section 5.5) joint density of $(T_1, \dots, T_m)$ to show that, (See Klenke - section 5.5)

$$\mathbb{P}\left(N_{t_j} - N_{t_{j-1}} = k_j,\ \forall j = 1, \dots, m\right) = \prod_{j=1}^{m} \left[e^{-\lambda(t_j - t_{j-1})} \cdot \frac{(\lambda(t_j - t_{j-1}))^{k_j}}{k_j!} \right]$$

$\forall\, t_1 < t_2 < \dots < t_m,\ \forall\, k_1, \dots, k_m \in \mathbb{N}_0,\ \forall m$

Decomposition and superposition of Poisson Process

- Superposition

Let $(N^k)_{k \in M}$ a countable family of independent Poisson processes s.t

$$N^k = (N^k_t)_{t \geq 0} \text{ is a } PP(\lambda_k)$$

Interpretation: Each process counts events of a given mark $k \in M$

Let $N_t := \sum_{k \in M} N^k_t$, $\forall t \geq 0$. If $\lambda = \sum_{k \in M} \lambda_k < \infty$

$\Longrightarrow (N_t)_{t \geq 0}$ is $PP(\lambda)$

Indeed: $\cdot\ (N_t)$ is a counting process $\begin{cases} N_0 = 0 \\ N_t \nearrow \text{ since } N^k_t \nearrow\ \forall k \in M \end{cases}$

- indep. of increments follows from indep. of increments of $(N^k)_{k \in M}$
- for any $0 < s < t$:

$$N_t - N_s = \sum_{k \in M} (N^k_t - N^k_s) \overset{d}{=} \sum_{k \in M} Poi((t-s)\lambda_k) \overset{\text{indep. of } N^k,\ k \in M}{=} Poi((t-s)\lambda)$$

#

- Decomposition: Marked Point Process

Let $(N_t)_{t \geq 0}$ be a $PP(\lambda)$ and assume that each arrival event, indep. on the other, has a random mark in a countable set M with density $P = (P_k)_{k \in M} \in \mathcal{P}(M)$. $\longrightarrow$ Also represented as: $(T_n, m_n)_{n \geq 1}$, where m_n: mark of n-th arrival $\sim P$, and T_n: n-th arrival time $\sim \Gamma(n, \lambda)$

Let $N^k_t = \#$ arrivals of mark k in $[0, t]$

$\forall t \geq 0$ and $\forall k \in M$

$\Longrightarrow N^k = (N^k_t)$ is a $PP(\lambda P_k)$ and $(N^k)_{k \in M}$ are independent

Indeed: $\cdot\ N^k$ are all counting processes by construction

Indeed: • N^n are all counting processes by construction

• N^k have all indep. increments, since N has indep. increments

Moreover, for any $0 < s < t$, $\forall (m_k)_{k \in M}$ s.t $\sum_{k \in M} m_k = n < \infty$

$$\mathbb{P}(N_t^k - N_s^k = m_k, \forall k \in M) = \underbrace{\mathbb{P}(N_t^k - N_s^k = m_k, \forall k \in M \mid N_t - N_s = n)}_{\text{combinatorial term}} \underbrace{\mathbb{P}(N_t - N_s = n)}_{\frac{(\lambda(t-s))^n}{n!} e^{-\lambda(t-s)}}$$

$$= \frac{\cancel{n!}}{\prod_{k \in M} m_k!} \left(\prod_{k \in M} p_k^{m_k} \right) \frac{[\lambda (t-s)]^n}{\cancel{n!}} e^{-\lambda(t-s)}$$

$$= \prod_{k \in M} \left[\frac{(p_k \cdot \lambda(t-s))^{m_k}}{m_k!} \cdot e^{-\lambda p_k (t-s)} \right]$$

$\begin{cases} n = \sum_{k \in M} m_k \\ 1 = \sum_{k \in M} p_k \end{cases}$

which implies that (N^k) are independent, and all with stationary increments s.t. $N_t^k - N_s^k \sim Poi(\lambda \cdot p_k (t-s))$. #