Poisson Clumping Heuristic

It's a technique for deriving Prob. Approxi.

<u>Def</u>: For random sets $C_1, C_2 \cdots C_n \cdots$. $0 < Y_1 < \cdots < Y_n \cdots$ (Y_n) is Poisson process with rate λ

$C_n + Y_n$ the translated set is called clump.

Then we call $\cup(Y_k + C_k)$ mosaic process with clump rate λ.

① <u>Example</u>:

Consider Ornstein - Uhlenbeck Process (Standard stationary, $m(x) = -x$, $\sigma^2(x) = 2$) $(X_t)_{t \geq 0}$.

Next, we will find approxi. for $P(\max_{[0,a]} X_t \geq b)$

$S = \{t \mid X_t \geq b\}$ random set. behaves like a mosaic process:

i) First reach high level b is the origin of first clump.

ii) Locally at level b. X_t behaves like a $(-b, 2)$ - BM. The set of $\{t : X_t \geq b\}$ after the first origin is a clump C_1.

iii) Then X_t will be pulled back to level b then 0. Repeatly, it will reach level b again.

iv) Set Y_1 is during time from origin to first reach at level b. The during time between clump origins form $(Y_n)_{n\geq 2}$.

v) $(Y_n) \sim Exp(\lambda)$ approximately is reasonable since the time between clumps is memoryless.

Next we will find rate λ:

$$p(\max_{[0,a]} X_t \geq b) = p(S \cap [0,a] \neq \emptyset) \approx p(Y_1 \leq a) = 1 - e^{-\lambda a}.$$

Suppose the stationary dist. of X is $Z \sim N(0,1)$.

We have relation : $\pi[b,\infty) = \lambda E(C)$ in long time.

by : $\pi[b,\infty)t \approx$ Time spend on $\{X_t \geq b\}$.

$=$ Time spend on clumps $= \lambda t E(C)$

$E(C) \approx E(T_{[0,\infty)})$. $T_{[0,\infty)}$ is sojourn time of $(-b,2)$ - SBM. By scale : $E(C) \approx E(T'_{(-\infty,0]})$. $T'_{(-\infty,0]}$ is sojourn time of $(\frac{b}{\sqrt{2}}, 1)$ - SBM below 0.

$\Rightarrow E(C) \approx 1/2(\frac{b}{\sqrt{2}})^2 = 1/b^2$. So $\lambda \approx b^2(1-\phi(b))$

We obtain $p = 1 - e^{-b^2(1-\phi(b))a}$.

② <u>Procedure:</u>

i) Transform the random extrema problem into sparse random sets.

e.g. Find $P(\max_{[0,n]} X_t \geq b)$. Set $S = \{\max_{[0,n]} X_t \geq b\}$

$$p \approx P(S \cap [0,n] \neq \emptyset)$$

ii) Sparse random set often consists of i.i.d random clumps thrown down randomly.

e.g. Assert S is approxi a Mosaic process with rate λ.

iii) Estimate the clump size λ.

e.g. From $\pi[b,\infty) \approx \lambda E(C)$. find $E(C)$.

iv) λ can be estimated by a simpler process.

e.g. The sojourn time problem of SBM.