(52) another example of a point process:

Def. 5.19 (Poisson random measure (PRM))
Let μ be a Random measure on $\mathcal{A}^* = \mathcal{B}(E)$, that is $\mu(K) < \infty$ for all compact $K \in \mathcal{A}^*$. Then a point process N is called Poisson random measure with mean measure μ (PRM(μ)) if
(i) $N(A) \sim \mathrm{Pois}(\mu(A))$ for all $A \in \mathcal{A}^*$
(ii) $N(A_1), \ldots, N(A_n)$ indep. r.v.'s, if $A_i \cap A_j = \emptyset$, $i \neq j$, $i = 1, \ldots, n$

Ex. 5.20 (homogeneous Poisson proc.)
$E = [0,\infty)$, mean measure $\mu(A) := \tau \cdot \lambda(A)$, $A \in \mathcal{B}([0,\infty))$ ← Lebesgue-Borel measure
$A = [0,t] \longrightarrow$ $N(\omega, A) = N(t)$, $t \geq 0$ Poisson process with intensity τ
(exercise)

We now want to discuss an example of a point process given by the cluster point process in connection with the general cluster model:

Def. 5.21 (General cluster model)
The general cluster model is described by the following conditions:
(i) Claims arrive at (non-delayed) reporting times $0 < T_1 < T_2 < \ldots$, $i \geq 1$
(ii) Each claim at T_i triggers a cluster of payments $X_{ij} \geq 0$ (X_{ij} r.v.) to the insured at time
$$T_{ij} = T_i + \sum_{k=1}^{j} Y_{ik}$$
for all $1 \leq j \leq K_i$, where $K_i \in \{1, 2, \ldots\}$ and Y_{ij}, $j \geq 1$ are positive r.v.'s.

$\longrightarrow$ $\sum_{k=1}^{j} X_{ik}$ sum of all payments to the insured $(T_i, T_{ij}]$

Let us now introduce the following point process, called cluster point process w.r.t.
$$E = [0,\infty) \times [0,\infty)^{\mathbb{N}} \times [0,\infty)^{\mathbb{N}} \times \mathbb{N}, \qquad (5.3)$$
where $[0,\infty)^{\mathbb{N}} \overset{\text{def}}{=} \{(a_n)_{n \geq 1} : a_j \in [0,\infty) \text{ for all } j \geq 1\}$

$\longrightarrow$ E complete separable metric space

(53) with metric d given by

$$d(x,y) = |t_1 - t_2| + \rho(Y_1, Y_2) + \rho(X_1, X_2) + |K_1 - K_2|$$

for $x = (t_1, Y_1, X_1, K_1)$, $y = (t_2, Y_2, X_2, K_2) \in E$, where ρ is a metric on $[0,\infty)^{\mathbb{N}}$ defined by

$$\rho(X_1, X_2) = \sum_{j\geq 1} 2^{-j} \frac{|x_j^{(1)} - x_j^{(2)}|}{1 + |x_j^{(1)} - x_j^{(2)}|}, \quad X_1 = (x_j^{(1)})_{j\geq 1}, \; X_2 = (x_j^{(2)})_{j\geq 1} \in [0,\infty)^{\mathbb{N}}$$

Define $Z_i : \Omega \to E$ r.v. (w.r.t. $\mathcal{A}^* = \mathcal{B}(E)$) by

$$Z_i(\omega) = (T_i(\omega), \underbrace{(Y_{ij}(\omega))_{j\geq 1}, (X_{ij}(\omega))_{j\geq 1}, K_i(\omega)}_{=: A_i}), \quad i \geq 1$$

$\longrightarrow$ $$N(\omega, A) := \sum_{i\geq 1} \varepsilon_{Z_i(\omega)}(A) = \sum_{i\geq 1} \varepsilon_{(T_i(\omega), A_i(\omega))}(A), \quad (5.4)$$

$A \in E$ is a <u>point process</u> (marked by $(A_i)_{i\geq 1}$), if for all ω $Z_i(\omega) \in K$ for finitely many i, $K \subseteq E$ compact

$\longrightarrow$ <u>Application: Chain ladder method</u>

$\longrightarrow$ popular method in insurance for estimating reserves

$\longrightarrow$ based on the <u>Chain ladder model</u>

$\longrightarrow$ <u>model for claim numbers / total claim amounts</u> based on the cluster process model

<u>Notation</u>: periods of time (e.g. in years)

$$C_i := (i-1, i], \quad C_{i,i+j} := (i-1, i+j]$$

$\longrightarrow$ $N_{i,i+j} \overset{\text{def}}{=}$ number of payments in $C_{i,i+j}$ arising from reporting times T_n in C_i

$$= \#\{(n,\ell) : T_n \in C_i, T_{n\ell} \in C_{i,i+j}, n \geq 1, 1 \leq \ell \leq K_i\}$$

$$= \sum_{n\geq 1} \sum_{\ell=1}^{K_i} \mathbb{1}_{\{T_n \in C_i, T_{n\ell} \in C_{i,i+j}\}}$$

$$= \sum_{n\geq 1} \sum_{\ell=1}^{K_i} \mathbb{1}_{\{T_n \in C_i, T_n + \sum_{j=1}^{\ell} Y_{nj} \in C_{i,i+j}\}}$$

$$\overset{(5.2)}{=} \int_E f(\overbrace{t, (Y_j)_{j\geq 1}, (X_j)_{j\geq 1}, K}^{=z}) \, N(d\underbrace{(t, (Y_j), (X_j), K)}_{=z}), \quad (5.5)$$

where $f(z) = f(t, (Y_j)_{j\geq 1}, (X_j)_{j\geq 1}, K) := \sum_{\ell=1}^{K} \mathbb{1}_{\{t \in C_i, t + Y_1 + \dots + Y_\ell \in C_{i,i+j}\}}$ is a (measurable) function on E

(54) $S_{i,i+j} \overset{\text{def}}{=}$ sum of all payments in $C_{i,i+j}$ triggered at reporting times $T_n \in C_i$

$$= \sum_{n\geq 1} \sum_{\ell \geq 1} X_{n\ell} \, 1\!\!1_{\{T_n \in C_i, \, T_{n\ell} \in C_{i,i+j}\}}$$

$$\overset{(5.2)}{=} \int_E g(z) \, N(dz), \qquad (5.6)$$

where $g(z) = g(t,(y_j)_{j\geq 1},(x_j)_{j\geq 1},K) := \sum_{\ell=1}^{K} x_\ell \, 1\!\!1_{\{t\in C_i, \, t+y_1+\dots+y_\ell \in C_{i,i+j}\}}$

$\longrightarrow$ chain ladder given by $(N_{i,i+j}, S_{i,i+j})$, $i=1,\dots,m$, $1\leq i+j \leq m$:

$$\begin{array}{cccc} (N_{1,1}, S_{1,1}) & (N_{1,2}, S_{1,2}) & \dots & (N_{1,m}, S_{1,m}) \\ & (N_{2,2}, S_{2,2}) & \dots & (N_{2,m}, S_{2,m}) \\ & & \ddots & \vdots \\ & & & (N_{m,m}, S_{m,m}) \end{array} \qquad (5.7)$$

$\longrightarrow$ If C_m is the present year, then the chain ladder provides the complete annual information on payments in present and past years

$\xrightarrow[\text{year}]{C_m \text{ present}}$ central problem

How can we predict or estimate the future claim numbers $N_{i,m+1}$ or total claim amounts $S_{i,m+1}$, $i=1,\dots,m$ based on past and present observations $N_{i,i+j}$, $1\leq i\leq m$, $1\leq i+j\leq m$ or $S_{i,i+j}$, $1\leq i\leq m$, $1\leq i+j\leq m$ in (5.7), respectively?

$\longrightarrow$ possible solution: best predictor/estimator $\hat{\mu}$ of $N_{i,m+1}$ given by the minimizer of $E[(N_{i,m+1}-Y)^2]$ w.r.t. r.v.'s Y which are functions of $N_{i,i+j}$, $1\leq i\leq m$, $1\leq i+j\leq m$

$\xrightarrow{\text{L.4.1.4}}$ $\hat{\mu} = E[N_{i,m+1} \mid N_{\ell,\ell+j}, 1\leq \ell\leq m, 1\leq \ell+j\leq m]$

reasonable assumption: $N_{i,i+j}$, $i\geq 1$ independent r.v.'s for each j

$\xrightarrow[\text{Prob. 2}]{\text{Proof of Exer. 4,}}$ $\hat{\mu} = E[N_{i,m+1} \mid N_{i,i}, \dots, N_{i,m}]$ (5.8)

55

$\xrightarrow{(4.3)}$ $\hat{\mu} = g(N_{i,i}, \dots, N_{i,m})$, g function

$\longrightarrow$ problem: g non-linear in general and difficult or impossible to compute

$\longrightarrow$ possible solution: "approximation" of $\hat{\mu}$ by a linear function of $N_{i,i}, \dots, N_{i,m}$ or if we additionally require memorylessness of $N_{i,i}, \dots, N_{i,m}$ by a linear function of $N_{i,m}$, that is

$$\hat{\mu} \approx c \cdot N_{i,m} \qquad (5.9)$$

for a constant c

$\xrightarrow[\text{arguments for } S_{i,i+j}]{(5.8), (5.9), \text{ same}}$ Mack's model:

Def. 5.22 (Mack's model)

In addition to (i) and (ii) in Def. 5.21 we assume that

(iii) The r.v.'s $(N_{i,i+j}, S_{i,i+j})_{j \geq 0}$ are i.i.d. w.r.t. $i = 1, 2, \dots$

(iv) $N_{i,i+j}, S_{i,i+j} > 0$ with prob. 1, $i \geq 1, j \geq 0$ and for all j there ex. real numbers f_j, g_j s.t. for all i:

$$E[N_{i,i+j+1} \mid N_{i,i}, \dots, N_{i,i+j}] = f_j N_{i,i+j} \qquad (5.10)$$

and

$$E[S_{i,i+j+1} \mid S_{i,i}, \dots, S_{i,i+j}] = g_j \cdot S_{i,i+j} \qquad (5.11)$$

Rem. 5.23: $N_{i,i+j}, S_{i,i+j} > 0$ is a serious restriction, since this excludes e.g. that $N_{i,i+j}$ is Poisson distributed.

$\longrightarrow$ next aim: estimation of f_j, g_j in (5.10), (5.11) from present and past observations:

W.l.o.g. let us only consider this problem in the sequel for f_j w.r.t. claim numbers

Using (iii) in Def. 5.22 and the SLLN for each fixed j we get the following sample versions of $E[N_{i,i+j}] \overset{(iii)}{=} E[N_{1,1+j}]$, $Var[N_{i,i+j}] \overset{(iii)}{=} Var[N_{1,1+j}]$ and $Cov[N_{i,i+j_1}, N_{i,i+j_2}] \overset{(iii)}{=} Cov[N_{1,1+j_1}, N_{1,1+j_2}]$:

56

$$\overline{N}_m^{(j)} := \frac{1}{m-j} \sum_{i=1}^{m-j} N_{i,i+j} \xrightarrow[m\to\infty]{} E[N_{1,1+j}], \qquad (5.12)$$

$$[S_m^{(j)}]^2 := \frac{1}{m-j-1} \sum_{i=1}^{m-j} (N_{i,i+j} - \overline{N}_m^{(j)})^2 \xrightarrow[m\to\infty]{} Var[N_{1,1+j}] \quad (5.13)$$

and

$$\hat{\gamma}_m(j_1,j_2) := \frac{1}{m-j_2} \sum_{i=1}^{m-j_2} (N_{i,i+j_1} - \overline{N}_m^{(j_1)})(N_{i,i+j_2} - \overline{N}_m^{(j_2)})$$

$$\xrightarrow[m\to\infty]{} Cov[N_{1,1+j_1}, N_{1,1+j_2}], \; j_1 < j_2 \quad (5.14)$$

with prob. 1, if $E[N_{1,1+j}], Var[N_{1,1+j}] < \infty$.

By applying (5.10) and $E[E[X|A]] = E[X]$, we see that

$$E[N_{1,1+j+1}] = E[E[N_{1,1+j+1} | N_{1,1}, \dots, N_{1,1+j}]]$$

$$= E[f_j N_{1,1+j}] = f_j \cdot E[N_{1,1+j}] \overset{>0}{\underset{\text{cond. (iv)}}{}}$$

$$\Rightarrow \quad f_j = \frac{E[N_{1,1+j+1}]}{E[N_{1,1+j}]} \qquad (5.15)$$

$\xrightarrow{(5.12)}$ for large $m-j$

$$\hat{f}_j^{(m)} := \frac{\overline{N}_m^{(j+1)}}{\overline{N}_{m-1}^{(j)}} = \frac{\sum_{i=1}^{m-j-1} N_{i,i+j+1}}{\sum_{i=1}^{m-j-1} N_{i,i+j}} \approx f_j \qquad (5.16)$$

($\hat{f}_j^{(m)}$ chain ladder estimator of f_j)

$\longrightarrow$ Prop. 5.24 (Asymptotic properties of $\hat{f}_j^{(m)}$)

Assume Mack's model. Then

(i) $\hat{f}_j^{(m)}$ is a strongly consistent estimator of f_j, that is

$$\hat{f}_j^{(m)} \xrightarrow[m\to\infty]{} f_j \text{ with prob. 1}$$

for all $j \geq 0$

(ii) If $Var[N_{1,1+j}] < \infty$ for all j, then $\hat{f}_j^{(m)}$ is asymptotically normal, that is

$$\sqrt{m}\,(\hat{f}_j^{(m)} - f_j) \xrightarrow[m\to\infty]{d} Z, \quad Z \sim N(0, \sigma^2),$$

where

$$\sigma^2 = \frac{E[(N_{1,1+j+1})^2] - f_j^2 E[(N_{1,1+j})^2]}{E[(N_{1,1+j})^2]} \qquad (5.17)$$