Tweedie New Researcher Invited Lecture

Poisson Disorder Problems

Savas Dayanik
Princeton University

Ninth Meeting of New Researchers in Statistics and Probability ● Seattle, August 1-5, 2006
1. Problem Description

Let $X$ be a **compound Poisson process** whose rate $\lambda_0$ and jump distribution $\nu_0(\cdot)$ change to $\lambda_1$ and $\nu_1(\cdot)$, respectively, at some **unknown and unobservable** time $\theta$. 

![Diagram of a compound Poisson process](image)
1. Problem Description

Let $X$ be a compound Poisson process whose rate $\lambda_0$ and jump distribution $\nu_0(\cdot)$ change to $\lambda_1$ and $\nu_1(\cdot)$, respectively, at some unknown and unobservable time $\theta$.

(Only the process $X$ is observable.)
1. Problem Description

Let $X$ be a compound Poisson process whose rate $\lambda_0$ and jump distribution $\nu_0(\cdot)$ change to $\lambda_1$ and $\nu_1(\cdot)$, respectively, at some unknown and unobservable time $\theta$.

(Only the process $X$ is observable.)

Problem: Find a decision rule which
- detects the disorder time $\theta$ as quickly as possible,
- is adapted to the history of $X$. 
Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space supporting random variables \(\theta, Y_1, Y_2, \cdots\), a counting process \(N = \{N_t; t \geq 0\}\). Define

\[
X_t = X_0 + \sum_{k=1}^{N_t} Y_k \equiv X_0 + \int_{(0,t] \times \mathbb{R}^d} y \, p(ds, dy), \quad t \geq 0
\]

in terms of the point process describing jump times and sizes

\[
p((0, t] \times A) \triangleq \sum_{k=1}^{\infty} 1\{\sigma_k \leq t\} 1\{Y_k \in A\}, \quad t \geq 0, \ A \in \mathcal{B}(\mathbb{R}^d).
\]

and \(\sigma_k = \inf\{t > \sigma_{k-1} : X_t \neq X_{t-}\}, \ k = 1, 2, \ldots\ (\sigma_0 \equiv 0)\).

\[
(1) \quad \mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0} \quad \text{as the natural filtration of } X,
\]

\[
\mathbb{G} = \{\mathcal{G}_t\}_{t \geq 0}, \quad \mathcal{G}_t \triangleq \mathcal{F}_t \vee \sigma\{\theta\}.
\]

The disorder time \(\theta\) has the distribution

\[
(2) \quad \mathbb{P}\{\theta = 0\} = \pi \quad \text{and} \quad \mathbb{P}\{\theta > t|\theta > 0\} = e^{-\lambda t}, \quad t \geq 0.
\]

The counting process \(\{p(t, A) \triangleq p((0, t] \times A); t \geq 0\}\) is a non-homogeneous Poisson process with the \((\mathbb{P}, \mathbb{G})\)-intensity

\[
(3) \quad h(t, A) \triangleq \lambda_0 \nu_0(A) 1_{\{t < \theta\}} + \lambda_1 \nu_1(A) 1_{\{t \geq \theta\}}, \quad t \geq 0.
\]
Our **problem** is (i) to calculate the **minimum Bayes risk**

\[
V(\pi) \triangleq \inf_{\tau \in \mathcal{F}} \ R_\tau(\pi),
\]

\[(4) \quad R_\tau(\pi) \triangleq \mathbb{P}\{\tau < \theta\} + c \cdot \mathbb{E}\left[(\tau - \theta)^+\right], \quad \pi \in [0, 1),
\]

and (ii) to find an $\mathcal{F}$-stopping time $\tau$ where the infimum is attained (if exists, called a **minimum Bayes detection rule**).

The **Bayes risk** $R_\tau(\pi)$ in (4) associated with every $\mathcal{F}$-stopping time $\tau$ is the sum of

- the false alarm frequency $\mathbb{P}\{\tau < \theta\}$, and
- the expected detection delay cost $c \cdot \mathbb{E}[(\tau - \theta)^+]$. 


Our problem is (i) to calculate the minimum Bayes risk

\[ V(\pi) \triangleq \inf_{\tau \in \mathcal{F}} R_\tau(\pi), \] (4)

\[ R_\tau(\pi) \triangleq \mathbb{P}\{\tau < \theta\} + c \cdot \mathbb{E}[\tau - \theta]^+, \quad \pi \in [0, 1), \]

and (ii) to find an \( \mathcal{F} \)-stopping time \( \tau \) where the infimum is attained (if exists, called a minimum Bayes detection rule).

The Bayes risk \( R_\tau(\pi) \) in (4) associated with every \( \mathcal{F} \)-stopping time \( \tau \) is the sum of

- the false alarm frequency \( \mathbb{P}\{\tau < \theta\} \), and
- the expected detection delay cost \( c \cdot \mathbb{E}[\tau - \theta]^+ \).

Standard Bayes risks include

Linear delay penalty: \( R_\tau(\pi) = \mathbb{P}\{\tau < \theta\} + c \mathbb{E}[\tau - \theta]^+ \),

\[ R_\tau^{(\varepsilon)}(\pi) \triangleq \mathbb{P}\{\tau < \theta - \varepsilon\} + c \mathbb{E}[\tau - \theta]^+, \]

Expected miss: \( R_\tau^{(\text{miss})}(\pi) \triangleq \mathbb{E}[\theta - \tau]^+ + c \mathbb{E}[\tau - \theta]^+ \),

Expon. delay penalty: \( R_\tau^{(\text{exp})}(\pi) \triangleq \mathbb{P}\{\tau < \theta\} + c \mathbb{E}[e^{\alpha(\tau - \theta)^+} - 1]. \)
Where do the disorder problems arise?

**Insurance companies:** Recalculate the premiums for the future sales of insurance policies when the risk structure changes (e.g., the arrival rate of claims of certain size).

**Airlines, retailers of perishable products:** Adjust the prices when a change in the demand structure is detected (e.g., the arrival rate of a certain type of customers).

**Quality control and maintenance:** Inspect, recalibrate, or repair tools and machines as soon as a manufacturing process goes out of control.

**Fraud and computer intrusion detection:** Alert the inspectors for an immediate investigation as soon as abnormal credit card activity, cell phone calls, or computer network traffic are detected.
2. The Model

Let $(\Omega, \mathcal{F}, \mathbb{P}_0)$ be a p.s. with independent random elements:

- a Poisson process $N = \{N_t; t \geq 0\}$ with rate $\lambda_0$,
- iid $\mathbb{R}^d$-valued rv's $Y_1, Y_2, \ldots$ with distr. $\nu_0(\cdot)$ ($\nu_0(\{0\}) = 0$),
- a rv $\theta$ with the distribution

$$
\mathbb{P}_0\{\theta = 0\} = \pi \quad \text{and} \quad \mathbb{P}_0\{\theta > 0\} = (1 - \pi)e^{-\lambda t}, \ t \geq 0.
$$

A compound Poisson process with arrival rate $\lambda_0$ and jump distribution $\nu_0(\cdot)$ is defined by

$$
X_t = X_0 + \sum_{k=1}^{N_t} Y_k = X_0 + \int_{(0,t] \times A} y \ p(ds, dy), \ t \geq 0
$$

in terms of the point process on $(\mathbb{R}_+ \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}_+) \times \mathcal{B}(\mathbb{R}^d))$

$$
p((0, t] \times A) \triangleq \sum_{k=1}^{\infty} 1\{\sigma_k \leq t\} 1_A(Y_k), \quad t \geq 0, \ A \in \mathcal{B}(\mathbb{R}^d).
$$

Under $\mathbb{P}_0$ the process $\{p((0, t] \times A); t \geq 0\}$ is homogeneous Poisson process with the $\mathcal{F}$-intensity $\lambda_0 \cdot \nu_0(A)$. Each $\sigma_k$ is a jump time of $X$, and $\mathcal{F}$ is its history, and $\mathcal{G} = \mathcal{F} \vee \sigma\{\theta\}$. 
Let $\lambda_1$ be a constant, and $\nu_1(\cdot)$ be a probability measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ absolutely continuous wrt $\nu_0(\cdot)$ with RN-derivative

$$f(y) \triangleq \frac{d\nu_1(y)}{d\nu_0(y)}, \quad y \in \mathbb{R}^d.$$ 

Define locally a new probability measure $\mathbb{P}$ on $(\Omega, \vee_{t\geq 0}\mathcal{G}_t)$ by the Radon-Nikodym derivatives

$$\left.\frac{d\mathbb{P}}{d\mathbb{P}_0}\right|_{\mathcal{G}_t} = 1_{\{t<\theta\}} + 1_{\{t\geq \theta\}}e^{-(\lambda_1-\lambda_0)(t-\theta)} \prod_{k=N_\theta+1}^{N_t} \left[ \frac{\lambda_1}{\lambda_0} f(Y_k) \right], \quad t \geq 0. \tag{5}$$

Then every counting process $\{p((0, t] \times A); \; t \geq 0\}, \; A \in \mathcal{B}(\mathbb{R}^d)$ is a nonhomogeneous Poisson process with the $(\mathbb{P}, \mathcal{G})$-intensity

$$h(t, A) = \lambda_0\nu_0(A)1_{\{t<\theta\}} + \lambda_1\nu_1(A)1_{\{t\geq \theta\}}. \tag{3}$$

Since $\mathbb{P}_0 \equiv \mathbb{P}$ on $\mathcal{G}_0 = \sigma\{\theta\}$, the disorder time $\theta$ has the same distribution under $\mathbb{P}_0$ and $\mathbb{P}$.

Therefore, the model under the measure $\mathbb{P}$ of (5) has the same setup described in the beginning.
3. A Markovian sufficient statistic for detection problem

The Bayes risk $R_\tau(\pi) = \mathbb{P}\{\tau < \theta\} + \mathbb{E}\left[(\tau - \theta)^+\right]$, $\pi \in [0, 1)$ in (4) for every $\mathcal{F}$-stopping rule $\tau$ can be written as

$$R_\tau(\pi) = 1 - \pi + c(1 - \pi) \mathbb{E}_0 \left[\int_0^\tau e^{-\lambda t} \left(\Phi_t - \frac{\lambda}{c}\right) dt\right].$$

The expectation in (6) is taken under the ref. p.m. $\mathbb{P}_0$, and

$$\Phi_t \triangleq \frac{\mathbb{P}\{\theta \leq t|\mathcal{F}_t\}}{\mathbb{P}\{\theta > t|\mathcal{F}_t\}}, \quad t \in \mathbb{R}_+.$$

The process $\Phi$ is a piecewise-deterministic Markov process:

$$\begin{cases} \Phi_t = x(t - \sigma_{n-1}, \Phi_{\sigma_{n-1}}), & t \in [\sigma_{n-1}, \sigma_n) \\ \Phi_{\sigma_n} = \frac{\lambda_1}{\lambda_0} f(Y_n) \Phi_{\sigma_{n-1}} & \end{cases}, \quad n \geq 1.$$

The function $x(\cdot, \phi) = \{x(t, \phi); t \geq 0\}$ is the solution of

$$\frac{d}{dt} x(t, \phi) = \lambda + ax(t, \phi), \quad t \in \mathbb{R}, \quad \text{and} \quad x(0, \phi) = \phi; \quad \text{i.e.,} \quad x(t, \phi) = \phi_d + e^{at} [\phi - \phi_d], \quad t \in \mathbb{R}.$$

Here $a \triangleq \lambda - \lambda_1 + \lambda_0$, $\phi_d \triangleq -\lambda/a$. 
The min. Bayes risk in (4) of the Poisson disorder problem is
\[ U(\pi) = 1 - \pi + c (1 - \pi) \cdot V \left( \frac{\pi}{1 - \pi} \right), \quad \pi \in [0, 1). \]

The function \( V : \mathbb{R}_+ \mapsto (-\infty, 0] \) is the value function of the discounted optimal stopping problem

\[ V(\phi) \triangleq \inf_{\tau \in \mathcal{F}} \mathbb{E}_0 \left[ \int_0^\tau e^{-\lambda t} g(\Phi_t) \, dt \bigg| \Phi_0 = \phi \right] \]

with the running cost function
\[ g(\phi) \triangleq \phi - \frac{\lambda}{c}, \quad \phi \geq 0. \]

for the piecewise deterministic Markov process \( \Phi \).

[Left: sample paths of the process \( \Phi \)]
4. Successive approximations

Let us introduce the family of optimal stopping problems

\[ V_n(\phi) \triangleq \inf_{\tau \in \mathcal{F}} \mathbb{E}_0^{\phi} \left[ \int_0^{\tau \wedge \sigma_n} e^{-\lambda s} g(\Phi_s) ds \right], \quad \phi \in \mathbb{R}_+, \ n \geq 0, \]  

obtained from (8) by stopping the process \( \Phi \) at the \( n \)th jump time \( \sigma_n \) of the process \( X \).

**Proposition.** For every \( n \geq 0 \) and \( \phi \in \mathbb{R}_+ \), we have

\[ -\frac{1}{c} \cdot \left( \frac{\lambda_0}{\lambda + \lambda_0} \right)^n \leq V(\phi) - V_n(\phi) \leq 0. \]  

**Proof.** Due to the discounting and exponentially distributed jump interarrival times of \( X \) under \( \mathbb{P}_0 \).

**Lemma.** For every \( \mathcal{F} \)-stopping time \( \tau \) and \( n \geq 0 \), there is an \( \mathcal{F}_{\sigma_n} \)-measurable random variable \( R_n : \Omega \mapsto [0, \infty] \) such that

\[ \tau \wedge \sigma_{n+1} = (\sigma_n + R_n) \wedge \sigma_{n+1}, \quad \mathbb{P}_0\text{-a.s. on } \{ \tau \geq \sigma_n \}. \]
If for every bounded function $w : \mathbb{R}_+ \mapsto \mathbb{R}$ we define

$$Jw(t, \phi) = \int_0^t e^{-(\lambda + \lambda_0)u} (g + \lambda_0 \cdot Sw) (x(u, \phi)) du, \quad t \in [0, \infty]$$

where

$$Sw(x) \triangleq \int_{\mathbb{R}^d} w \left( \frac{\lambda_1}{\lambda_0} f(y) x \right) \nu_0(\text{d}y), \quad x \in \mathbb{R}.$$ 

then we can calculate the successive approximations $\{V_n(\cdot)\}_{n \geq 1}$ of the value function $V(\cdot)$ by

$$V_0(\cdot) \equiv 0, \quad \text{and} \quad V_n(\cdot) = J_0 V_{n-1}(\cdot) \triangleq \inf_{t \geq 0} JV_{n-1}(t, \cdot) \quad \forall n \geq 1.$$ 

Moreover

1. $V_n(\cdot) \searrow V(\cdot)$ (exponentially fast)

2. $V(\cdot) = J_0 V(\cdot)$ on $\mathbb{R}_+$. (Dynamic programming equation)

3. The value function $V(\cdot)$ is concave and nonpositive.

4. The stopping region $\Gamma = \{\phi \in \mathbb{R}_+ : V(\phi) = 0\}$ is in the form $\Gamma = [\xi, \infty)$ for some $0 < \xi < +\infty$. 
5. Examples

(a) Discrete jump distributions

(b) $\frac{\lambda_1}{\lambda_0} = \frac{1}{2}$

(c) $\frac{\lambda_1}{\lambda_0} = 1$

(d) $\frac{\lambda_1}{\lambda_0} = 2$

(e) Continuous jump distributions ($\mu = 2$)

(f) Gamma(2, $\mu$)

(g) Gamma(3, $\mu$)

(h) Gamma(6, $\mu$)

Parameters: $c = 0.2$, $\lambda = 1.5$, $\lambda_0 = 3$. 
\[ R^{(\text{linear})}_\tau (\pi) \triangleq \mathbb{P}\{\tau < \theta\} + c \mathbb{E}(\tau - \theta)^+, \]

\[ R^{(\varepsilon)}_\tau (\pi) \triangleq \mathbb{P}\{\tau < \theta - \varepsilon\} + c \mathbb{E}(\tau - \theta)^+, \]

\[ R^{(\text{miss})}_\tau (\pi) \triangleq \mathbb{E}(\theta - \tau)^+ + c \mathbb{E}(\tau - \theta)^+, \]

\[ R^{(\exp)}_\tau (\pi) \triangleq \mathbb{P}\{\tau < \theta\} + c \mathbb{E}[e^{\alpha(\tau - \theta)^+} - 1] \]

\[ \mathbb{E}_0^\phi \left[ \int_0^\tau e^{-\lambda t} (\Phi_t - k) \, dt \right] \]

Standard Poisson disorder problems:
6. Appendix

**Lebesgue decomposition of the measures.** Let $\nu_0(\cdot)$ and $\nu_1(\cdot)$ be probability measures on $(\Omega, \mathcal{B}(\mathbb{R}^d))$. Then there exist a Borel function $f : \mathbb{R}^d \mapsto [0, \infty]$ and a Borel set $H \subseteq \mathbb{R}^d$ such that

$$\nu_0(H) = 0,$$

$$\nu_1(B) = \int_B f(y) \nu_0(dy) + \nu_1(B \cap H), \quad B \in \mathcal{B}(\mathbb{R}^d).$$

If an observation $Y_n$ falls in $H$, then one cannot make any error by concluding that the change from $\nu_0(\cdot)$ to $\nu_1(\cdot)$ has happened.

In general, an alarm given for the first time by the simple rule above or the decision rule obtained in the previous sections by applying to the measures $\nu_0(\cdot)$ and

$$\tilde{\nu}_1(\cdot) = \int_{y \in \cdot} f(y) \nu_0(dy),$$

will be optimal for the linear penalty in (4).
References


Galchuk, L. I. and Rozovskii, B. L. (1971). The disorder