ISBN:1858-1382

J M S K

Jurnal Matematika, Statistika, & Komputasi
http://journal.unhas.ac.id/index.php/jmsk
Vol. 9, No.1, 33-38, Juli 2012

Estimasi *Hazard Rate Temporal Point Process*

Nurtiti Sunusi[1]

Abstrak

Point process adalah suatu model stokastik yang dapat menerangkan fenomena alam yang sifatnya acak baik dalam ruang maupun waktu. Point process dikarakterisasikan oleh intensitas bersyaratnya (*conditional Intensity*). Pada penelitian ini intensitas bersyarat proses titik temporal (*temporal point process*) dipandang sebagai proses pembaruan (*renewal process*) dimana selisih waktu sejak kejadian terakhir tidak tergantung pada selang sebelumnya. Intensitas bersyarat yang bersesuaian pada model ini disebut *hazard rate.* Untuk mengestimasi parameter hazard rate digunakan metode *Hazard Rate Single Decrement* (HRSD) yang diadaptasi dari metode estimasi dalam studi aktuaria yang dipakai dalam pembentukan tabel mortalita. Pada metode ini, satu individu diasosiasikan dengan satu kejadian. Jika informasi yang digunakan pada pembentukan tabel mortalita adalah tanggal lahir dan tanggal meninggal, maka pada *temporal point process* digunakan informasi waktu mulai dan berakhirnya suatu kejadian. Selanjutnya pada bagian akhir, ditinjau dua kasus yaitu estimasi hazard rate dengan waktu antar kejadian berdistribusi uniform dan eksponensial.

Kata Kunci: Point process, hazard rate, intensitas bersyarat, eksponensial.

Abstract

Point process is a stochastic model that can explain the natural phenomena that are random in both space and time. Point process characterized by the conditional Intensity. In this study, the conditional intensity of temporal point process is seen as a renewal process with is the elapsed time since the last occurrence does not depend on the previous interval. Conditional intensity corresponding to the model is called the hazard rate.To estimate the hazard rate parameters, we use hazard rate single decrement method which is adapted from the method in actuarial studies that used to formation the mortality table. In this method,one individual associated with an event. If the information is used to the formation of mortality table is date of birth and date of death, then information of event time when start and finally in temporal point process is considered. Furthermore, at the end of this research, we reviewed two cases, namely hazard rate with inter event time is uniform and exponentially distributed.

Keywords: Point process, hazard rate, intensitas bersyarat, eksponensial.

1. Pendahuluan

Point process adalah suatu model stokastik yang dapat menerangkan fenomena alam yang sifatnya acak baik dalam ruang maupun waktu. *Point process* dikarakterisasikan oleh intensitas bersyaratnya (*conditional Intensity*). Dalam *point process*, fungsi intensitas bersyarat yang didefinisikan sebagai turunan dari perubahan peluang kemunculan kejadian memegang peranan yang sangat penting seperti yang dikemukakan oleh Ogata pada tahun 1999 [4]. Hal ini disebabkan karena fungsi intensitas bersyarat (*Conditional Intensity*) disingkat CI mengkarakterisasikan *point process* yang bersesuaian. Misalnya, jika intensitas bersyarat hanya

[1] *Jurusan Matematika FMIPA Universitas Hasanuddin ,email:ntitisanusi@gmail.com*

bergantung pada waktu atau kejadian sekarang, maka bentuk intensitas bersyarat seperti ini mengkarakterisasikan proses Poisson tak stasioner. Sedangkan untuk intensitas bersyarat yang bergantung hanya pada selisih waktu sejak waktu kemunculan kejadian terakhir, maka hal ini merupakan proses pembaruan. Intensitas bersyarat yang bersesuaian dengan proses ini disebut *hazard rate*. *Hazard rate* didefinisikan sebagai peluang munculnya suatu kejadian sesaat setelah t_0 pada selang waktu $(t_0, t_0 + 1]$ dan diketahui belum ada kejadian hingga saat t_0.

Umumnya hazard rate gempa diestimasi menggunakan persamaan likelihood proses titik [2], [5], [6]. Persamaan tersebut merupakan persamaan non linear yang tidak mudah diselesaikan secara analitik sehingga seringkali diselesaikan secara numerik. Persamaan likelihood pada kasus waktu tunggu berdistribusi eksponensial dapat diselesaikan secara analitik. Namun, untuk kasus waktu tunggu lainnya diselesaikan secara numerik. Selain itu, metode ini masih terbatas pada estimasi hazard rate dalam selang waktu observasi. Dengan demikian, metode ini belum dapat digunakan untuk memprakirakan peluang kemunculan gempa pada tiap satu periode berikutnya. *Hazard rate* merupakan kunci dalam teori likelihood *point process*, jika *hazard rate* diketahui maka simulasi proses yang bersesuaian dapat dilakukan. Oleh karena itu, perlu untuk memperoleh model paramerik yang akurat untuk *hazard rate*.

2. Estimasi *Hazard Rate Single Decrement*

Pada bagian ini dikemukakan suatu metode untuk estimasi *hazard rate temporal point process* $\lambda(t|H_t)$ yang disebut HRSD. HRSD meliputi estimasi *likelihood single decrement* dan estimasi momen *single decrement* [3], [6]. Mengingat di aktuaria, hazard rate disimbolkan dengan μ_t, maka untuk pembahasan selanjutnya digunakan symbol μ_t.

Misalkan $X(t_0) = T - t_0$ menyatakan waktu tunggu hingga kemunculan gempa berikutnya, jika diketahui selisih waktu *t*$_0$ sejak kemunculan kejadian gempa yang terakhir dan T adalah waktu kemunculan kembali antara kemunculan dua gempa. Misal $\mu, S, \mathrm{dan} f$ berturut-turut menyatakan *hazard rate*, fungsi ketahanan (*survival function*), dan fungsi kepadatan peluang. *Hazard rate* μ_t dapat dinyatakan sebagai [1]:

$$\mu_{t_0} = \lim_{\Delta t_0 \to 0} \frac{P(t_0 < T \le t_0 + \Delta t_0 | T > t_0)}{\Delta t_0} \approx \frac{f(t_0)}{S(t_0)}. \quad (1)$$

Dengan mengintegralkan $-\mu_y dy = d \ln S(y)$ dengan batas dari t_0 sampai $t_0 + \Delta t_0$, diperoleh bahwa peluang belum ada kejadian hingga $t_0 + \Delta t_0$ jika diketahui belum ada kejadian hingga saat t_0, adalah [1]:

$$_{\Delta t_0}p_{t_0} = P(T > t_0 + \Delta t_0 | T > t_0) = e^{-\int_{t_0}^{t_0+\Delta t_0} \mu_y dy} = e^{-\int_0^{\Delta t_0} \mu_{t_0+s} ds}. \quad (2)$$

Misalkan $t_0 = 0$, yaitu sesaat setelah terjadi gempa, diperoleh

$$_{\Delta t_0}p_0 = S(t_0) = P(T > t_0) = \exp\left(-\int_0^{t_0} \mu(y) dy\right) \quad (3)$$

dimana $S(t_0)$ adalah fungsi ketahanan. Distribusi waktu kemunculan kembali **T** suatu kejadian dan waktu tunggu hingga kemunculan kejadian berikutnya $X(t_0)$ masing-masing dinyatakan sebagai berikut [1]:

$$T \sim {}_{\Delta t_0}p_{t_0} \mu_{t_0} \quad (4)$$

dan

$$X \sim {}_{\Delta t_0}p_{t_0}\mu_{t_0+\Delta t_0}. \tag{5}$$

Dalam ekspresi ini, ${}_{\Delta t_0}p_{t_0}\mu_{t_0+\Delta t_0}$ merupakan peluang bahwa suatu kejadian muncul antara t_0 dan $t_0 + \Delta t_0$ jika diketahui belum ada kejadian hingga saat t_0, dan

$$\int_0^{\infty} {}_{\Delta t_0}p_{t_0}\mu_{t_0+\Delta t_0} = 1; \qquad \frac{d}{dt}\,{}_{\Delta t_0}p_{t_0} = -{}_{\Delta t_0}p_{t_0}\mu_{t_0+\Delta t_0}. \tag{6}$$

3. Estimasi Likelihood dalam *Single Decrement*

Estimasi *hazard rate* menggunakan pendekatan *single decrement* melalui metode *Maximum Likelihood Estimate* (MLE) membutuhkan informasi *exit time*, yaitu waktu pada saat timbul kejadian. Misalkan d_{t_0} menyatakan jumlah kejadian yang terjadi pada selang $(t_0, t_0 + 1]$ dan $(n_{t_0} - d_{t_0})$ menyatakan jumlah kejadian pada selang $(t_0, t_0 + 1]$ dan $(n_{t_0} - d_{t_0})$ menyatakan jumlah kejadian setelah t_0. Karena waktu untuk masing-masing kejadian berbeda, maka kejadian dipandang secara sendiri-sendiri dan mengambil perkalian kontribusi masing-masing kejadian pada fungsi likelihood. Likelihood *L* untuk kejadian ke-*i* pada selang $(t_i, t_i + 1]$ diberikan oleh *pdf* untuk kemunculan kejadian pada selang tersebut jika diketahui tidak ada kejadian hingga saat t_0. Hal ini dapat dinyatakan sebagai berikut

$$L_i = f(t_0(i)|T > t_0(i)) = \frac{f(t_0(i))}{S(t_0)} = \frac{S(t_0(i))\mu(t_0(i))}{S(t_0)} \tag{7}$$

Yaitu kontribusi kejadian ke-*i* pada *L*. Jika dimisalkan $s_i = t_0(i) - t_0$ adalah waktu kejadian ke-*i* dalam selang $(t_0, t_0 + 1]$, dengan $0 < s_i \leq 1$, maka

$$L_i = \frac{S(t_0+s_i)\mu(t_0+s_i)}{S(t_0)} = {}_{s_i}p_{t_0}\mu_{t_0+s_i}. \tag{8}$$

Kontribusi banyaknya kejadian d_{t_0} pada *L* adalah $\prod_{i=1}^{d} {}_{s_i}p_{t_0}\mu_{t_0+s_i}$. Kontribusi $n_{t_0} - d_{t_0}$ kejadian yang muncul setelah $t_0 + 1$ adalah $(p_{t_0})^{n_{t_0}-d_{t_0}}$ dimana n_{t_0} merupakan banyaknya kejadian yang muncul pada saat t_0 atau setelahnya. Dengan demikian likelihood total *L* adalah

$$\begin{aligned} L &= (1 - q_{t_0})^{n_{t_0}-d_{t_0}} \prod_{i=1}^{d_{t_0}} {}_{s_i}p_{t_0}\mu_{t_0+s_i} \\ &= (p_{t_0})^{n_{t_o}-d_{t_0}} \prod_{i=1}^{d_{t_0}} {}_{s_i}p_{t_0}\mu_{t_0+s_i}. \end{aligned} \tag{9}$$

Untuk menyelesaikan persamaan (9) untuk $\hat{q}_{t_0}$ diperlukan asumsi bahwa distribusi ${}_{s_i}p_{t_0}\mu_{t_0+s_i}$ dinyatakan dalam bentuk q_{t_0}.

Berikut ini ditinjau dua kasus, yakni jika l_{t_0+s} yang menyatakan jumlah kejadian setelah $t_0 + s$ diasumsikan berdistribusi linear dan eksponensial yang diperlukan untuk menyatakan ${}_{s_i}p_{t_0}\mu_{t_0+s_i}$.

Kasus I. Asumsi Linear

Jika l_{t_0+s} menyatakan jumlah kejadian setelah $t_0 + s$ merupakan fungsi linear, maka

$$l_{t_0+s} = a + bs. \tag{10}$$

Untuk $s = 0$, diperoleh $l_{t_0+1} = a + b$, dan untuk $s = 1$ diperoleh $l_{t_0+1} = a + b$. Dengan demikian, dari persamaan (10) diperoleh

$$l_{t_0+s} = l_{t_0} + (-d_{t_0}).s \tag{11}$$

dengan d_{t_0} menyatakan jumlah kejadian pada $(t_0, t_0 + 1)$ dan

$$l_0 = \sum_{t_0=0}^{\omega-1} d_{t_0} = d_0 + d_1 + \cdots + d_{\omega-1}. \tag{12}$$

Untuk asumsi linear

$${}_s p_{t_0} = \frac{l_{t_0+s}}{l_{t_0}} = 1 - s\frac{d_{t_0}}{l_{t_0}} \text{dan} {}_s q_{t_0} = 1 - {}_s p_{t_0} = s.q_{t_0}. \tag{13}$$

Karena $\frac{d_{t_0}}{l_{t_0}} = q_{t_0}$, maka ${}_s p_{t_0} = 1 - s.q_{t_0}$ dan $\mu_{t_0+s} = \frac{-\frac{d}{ds} l_{t_0+s}}{l_{t_0+s}} = \frac{q_{t_0}}{1-s.q_{t_0}}$.
Jadi

$${}_{s_i} p_{t_0} \mu_{t_0+s_i} = q_{t_0}. \tag{14}$$

Dari persamaan (9) dan (10) diperoleh bahwa

$$\begin{aligned} L &= (1 - q_{t_0})^{n_{t_0} - d_{t_0}} . \prod_{\mathcal{D}} {}_{s_i} p_{t_0} \mu_{t_0+s_i} \\ &= (1 - q_{t_0})^{n_{t_0} - d_{t_0}} .(q_{t_0})^{d_{t_0}}. \end{aligned} \tag{15}$$

Log likelihood ℓ untuk persamaan (15) diperoleh:

$$\ell = \ln L = (n_{t_0} - d_{t_0}) \ln(1 - q_{t_0}) + d_{t_0} \ln(q_{t_0}). \tag{16}$$

Dengan menyelesaikan persamaan berikut

$$\frac{d\ell}{dq} = \frac{d_{t_0}}{q_{t_0}} - \frac{n_{t_0} - d_{t_0}}{1 - q_{t_0}} = 0, \tag{17}$$

Dapat diperoleh

$$\hat{q_{t_0}} = \frac{d_{t_0}}{n_{t_0}}. \tag{18}$$

Dari persamaan (17) diperoleh

$$\frac{d^2\ell}{dq_{t_0}^2} = -\frac{d_{t_0}}{dq_{t_0}^2} - \frac{n_{t_0}-d_{t_0}}{1-q_{t_0}^2}. \quad (19)$$

Karena $d_{t_0}, (n_{t_0} - d_{t_0}), q_{t_0}^2$, dan $(1 - q_{t_0}^2)$ positif, maka $\frac{d^2\ell}{dq_{t_0}^2} < 0$. Jadi persamaan (18) merupakan solusi MLE untuk persamaan (15).

Kasus II. Asumsi Eksponensial

Jika l_{t_0+s} menyatakan jumlah kejadian setelah $t_0 + s$ merupakan suatu fungsi eksponensial, maka

$$l_{t_0+s} = ab^s. \quad (20)$$

Untuk*s = 0* diperoleh $l_{t_0+1} = ab$. Sehingga $b = \frac{l_{t_0+1}}{a} = \frac{l_{t_0+1}}{l_{t_0}}$. Dari persamaan di atas (20) diperoleh

$$l_{t_0+s} = l_{t_0}\left(\frac{l_{t_0}+1}{l_{t_0}}\right) = (l_{t_0+1})^s \,.\, (l_{t_0})^{1-s}\,. \quad (21)$$

Atau ditulis

$$\begin{aligned} l_{t_0+s} &= (l_{t_0+1})^s . (l_{t_0})^{1-s} \\ &= (p_{t_0} . l_{t_0})^s (l_{t_0})^{1-s} = (p_{t_0})^s . l_{t_0}. \end{aligned} \quad (22)$$

Dengan demikian

$$(p_{t_0})^s = \frac{l_{t_0+s}}{l_{t_0}} = {}_s p_{t_0} \quad (23)$$

dan ${}_s q_{t_0} = 1 - {}_s p_{t_0} = 1 - (p_{t_0})^s = 1 - (1 - q_{t_0})^s$. Selanjutnya,

$$\begin{aligned} \mu_{t_0+s} &= \frac{-\frac{d}{ds} l_{t_0+s}}{l_{t_0+s}} = \frac{-l_{t_0}(p_{t_0})^s . \ln p_{t_0}}{l_{t_0} . (p_{t_0})^s} \\ &= -\ln p_{t_0} = \mu. \end{aligned} \quad (24)$$

Oleh karena itu, berdasarkan persamaan (23) dan (24) diperoleh total likelihood sebagai berikut:

$$\begin{aligned} L &= (1 - q_{t_0})^{n_{t_0} - d_{t_0}} \prod_{i=1}^{d_{t_0}} \mu_{t_0+s} (p_{t_0})^{s_i} \\ &= (p_{t_0})^{n_{t_0} - d_{t_0}} \prod_{i=1}^{d_{t_0}} \mu_{t_0+s} (p_{t_0})^{s_i} \\ &= \mu^{d_{t_0}} \exp\left(-\mu\left[(n_{t_0} - d_{t_0}) - \sum_{i=1}^{d_{t_0}} s_i\right]\right). \end{aligned} \quad (25)$$

Log likelihood untuk persamaan (25) adalah

$$\ell = \ln L = d_{t_0} \ln \mu - \mu\left[(n_{t_0} - d_{t_0}) + \sum_{i=1}^{d_{t_0}} s_i\right]. \quad (26)$$

Solusi persamaan $\frac{d\ell}{d\mu} = 0$ adalah

$$\hat{\mu} = \frac{d_{t_0}}{(n_{t_0}-d_{t_0})+\sum_{i=1}^{d_{t_0}} s_i}. \qquad (27)$$

Karena q_{t_0} berkorespondensi satu-satu dengan μ, maka dari persamaan (27) dapat diperoleh:

$$\hat{q_{t_0}} = 1 - \exp(-\hat{\mu}). \qquad (28)$$

yang merupakan estimator maksimum likelihood untuk q_{t_0}.

4. Kesimpulan

Dari hasil uraian yang diberikan di atas, diperoleh bahwa estimasi *hazard rate single decrement* menggunakan asumsi waktu tunggu berdistribusi eksponensial serupa dengan estimasi *hazard rate* menggunakan asumsi waktu tunggu berdistribusi uniform. Di samping itu, solusi yang menggunakan asumsi eksponensial lebih informatif dibandingkan dengan solusi yang menggunakan asumsi linear (uniform). Hal ini dapat dilihat bahwa hasil estimasi yang menggunakan asumsi eksponensial mengandung informasi waktu kejadian, banyak kejadian d_{t_0} pada selang waktu tunggu $(t_0, t_0+1]$, dan banyak kejadian $(n_{t_0} - d_{t_0})$ setelah $(t_0 - 1]$.

Daftar Pustaka

[1] Bowers N.L., Gerber H.U., Hickman J.C., Jones D.A., and Nesbitt C. J., 1986. *Actuarial Mathematics*. The Society of Actuaries.

[2] Daley D.J. and Vere-Jones D., 2003. *An Introduction to the Theory of Point Processes*. Springer, Berlin.

[3] Darwis S., Sunusi N., Triyoso W., dan Mangku I.W., 2009. Single Decrement Approach for Estimating Earthquake Hazard Rate. *Advances and Applicationsin Statistica*, 11(2), 229–237.

[4] Ogata Y., 1999. Seismicity Analysis Through Point Process Modeling: A Review. *Pure and Applied Geophysics*, 155, 471–507.

[5] Sunusi N., Darwis S., dan Triyoso W., 2008. Estimating Intensity of Point Processes Models Applied to Earthquake Prediction. *Mathematics Journal University Teknologi Malaysia*, 2, 405–411.

[6] Sunusi N., Darwis S., Triyoso W., dan Mangku I.W., 2010. Study of Earthquake Forecast through Hazard Rate Analysis. *International Journal of Applied Mathematics and Statistics* (IJAMAS), 17(J10), 96–103.