APLIKASI METODE KERNEL DALAM ESTIMASI FUNGSI HAZARD

Oleh :

Bambang Avip Priatna Martadiputra

Jurusan Pendidikan Matematika
FPMIPA Universitas Peendidikan Indonesia

ABSTRACT

Let T be a nonnegative random variable representing the lifetimes of individuals in some population. Let f(t) denote the probability density function of T and F(t) denote the distribution function of T, the hazard function of T defined as

$$h(t) = \frac{f(t)}{S(t)} \text{ where } S(t) = 1 - F(t)$$

If equation (1) integrated we have cumulative hazard function H (t).
This paper describes application of kernel method for estimation of hazard function h (.) based censoring data. And then we will show that the hazard estimator is unbiased asymptotically, consistent, and normal asymptotically.

Key word: *kernel methods, estimation hazard function.*

FUNGSI SURVIVAL, FUNGSI HAZARD, DAN FUNGSI HAZARD KUMULATIF

Untuk meningkatkan kualitas suatu produk hasul industri dan dalam mengembangkan metode pengobatan diperlukan uji hidup. Uji hidup berguna untuk melakukan pengujian tentang daya tahan dan kehandalan suatu produk hasil industri (menunjukkan waktu kegagalan) dan atau pengukuran tentang tahan hidup penderita dalam pengujian yang menyangkut pengobatan suatu penyakit (menunjukkan waktu survival).

Cara uji hidup terbaik adalah dengan menguji sebanyak n benda pada kondisi (tegangan, stress) operasinya yang normal sampau semua benda/objek mati (sample lengkap) atau percobaan dihentikan sebelum semua benda mati (sample disensor). Sensor dimaksudkan untuk memperpendek waktu pengujian terutama bagi benda-bendda yang cukup handal. Beberapa macam sensor telah dipelajari orang, antara lain sensor tipe I : jika pengujian dihentikan setelah dicapai waktu tertentu (waktu penyensoran), dan sensor tipe II : jika pengujian dihentikan setelah kematian benda ke-r < n. Tahan hidup benda yang mati diamati dan digunakan untuk estimasi misalnya tahan hidup rata-rata dan kehandalam benda itu.

Misalkan T adalah variabel random tunggal nonnegative kontinu pada interval $[0, \infty)$ yang menunjukkan tahan hidup suatu benda, komponen, sistem, atau individu dalam suatu populasi, Jika f(t) dan F(t) masing-masing adalah fungsi kepadatan peluang dan fungsi distribusi dari T, maka fungsi hazard didefinisikan sebagai

$$h(t) = \frac{f(t)}{S(t)} \text{ dengan } S(t) = 1 - F(t) \qquad \dots (1)$$

dengan S(t) adalah fungsi survival.

Jika persamaan (1) diintegralkan maka akan diperoleh fungsi hazard kumulatif, yaitu

$$H(t) = \int_t^{\infty} h(x)dx = -\ln S(t) \qquad \dots (2)$$

Dari persamaan (2) akan diperoleh hubungan

$$S(t) = \exp(-\int_t^{\infty} h(x)dx) \qquad \dots (3)$$

PROSES MENGHITUNG (COUNTING PROCESS)

Suatu pendekatan alternatif untuk mengembangkan prosedur –prosedur inferensi untuk data tersensor adalah dengan menggunakan metode proses menghitung (counting process). Pendekatan ini pertama kali dikembangkan oleh Aalen (1975) dengan mengkombinasikan elemen-elemen dari integrasi stokastik, teori martingale waktu kontinu dan teori proses menghitung kedalam suatu metodologi yang lebih mudah untuk mengembangkan teknik-teknik inferensi besaran-besaran survival berdasarkan data tersensor.

Definisi

Proses menghitung N(t), $t \geq 0$ didefinisikan sebagai suatu proses stokastik yang mempunyai sifat : (a) N(0) = 0; (b) dengan peluang 1, $N(t) < \infty$; (c) lintasan-lintasan sample dari N(t) kontinu kanan dan konstan dengan lompatan berujuran 1.

Definisi.

Suatu proses $\lambda(t) = Y(t)h(t)$ … (4)

disebut proses intensitas untuk proses menghitung, yaitu suatu proses stokastik yang tergantung pada informasi yang dimuat dalam suatu proses histori F_t seluruh Y(t).

Definisi

Proses intensitas kumulatif didefinisikan sebagai $A(t) = \int_0^t \lambda(s)ds$ … (5)

yang memenuhi sifat $E\left[N(t)\middle|F_{f^-}\right] = E\left[A(t)\middle|F_{t^-}\right] = A(t)$ … (6)

Definisi

Suatu proses M(t) = N(t) – A(t) … (7)

disebut martingale proses menghitung, yang memenuhis sifat

$$E\left[dM(t)\middle|F_{t^-}\right] = E\left[dN(t) - dA(t)\middle|F_{t^-}\right] = E\left[dN(t)\middle|F_{t^-}\right] - E\left[dA(t)\middle|F_{t^-}\right] = 0 \qquad \ldots (8)$$

Definisi

Martingale adalah suatu proses stokastik yang mempunyai sifat bahwa nilai ekspektasi pada waktu t yang memberi histori pada waktu s < t sama dengfan nilai martingale pada waktu s. Dengan kata lain, M(t) adalah sebuah martingale jika $E\left[M(t)\middle|F_s\right] = M(s)$ … (9)

untuk semua s ≤ t.

Definisi

Pada suatu interval [0,t] suatu integral stokastik untuk suatu proses berkenaan dengan suatu matringale dinyatakan sebagai $\int_0^t K(u)dM(u)$ … (10)

yang mempunyai sifat bahwa integral-integral stokastik tersebut adalah martingale-martingale sebagai suatu fingsi dari t.

Proses variasi yang dapat diramalkan dapat diperoleh dari suatu proses variasi yang daoat diramalkan untuk suatu martingale asli, yaitu

$$\left\langle \int_0^t K(u)dM(u) \right\rangle = \int_0^t K^2(u)d\langle M\rangle(u) \qquad \ldots (11)$$

dengan $d\langle M\rangle(t) = \mathrm{var}\left(dM(t)\middle|F_{t^-}\right) = E\left[(dM(t))^2\middle|F_{t^-}\right]$ … (12)

MODEL INTENSITAS MULTIFIKATIF AALEN

Dalam bentuk-bentuk estimasi fungsi intensitas, teori nonparametrik memusatkan pada fungsi intensitas/hazard kumulatif. Hasil pekerjaan Aalen (1978) diketahui bahwa estimasi fungsi hazard dapat digambarkan dalam suatu kontek inferensi untuk suatu model intensitas multifikati proses menghitung yang diberikan oleh persamaan (4) yaitu $\lambda(t) = Y(t)h(t)$ untuk $t \geq 0$, dengan h(.) adalah fungsi nonstokastik tak diketahui yang disebut fungsi intensitas (fungsi hazard), dan Y(.) adalah proses stokastik terobservasi.

Jika persamaan (4) dan (5) disubstitusikan pada persamaan (7) maka akan diperoleh

$$N(t) = M(t) + \int_0^t h(s)Y(s)ds \qquad \ldots (13)$$

Selanjutnya jika persamaan (13) didiferensialkan maka diperoleh

$$h(t)Y(t)dt = dN(t) - dM(t) \qquad \ldots (14)$$

dengan dM(t) diartikan sebagai noise.

Jika setiap ruas persamaan (14) dibagi dengan Y(t) ≠ 0 dan dikalikan dengan suatu indikator dari Y(t), yaitu J(t) = I{Y(t)>0} kemudian dituliskan dalam bentuk integral maka diperoleh

$$\int_0^t J(s)h(s)ds = \int_0^t \frac{J(s)}{Y(s)} dN(s) - \int_0^t \frac{J(s)}{Y(s)} dM(s) \qquad \ldots (15)$$

Selanjutnya

$$\hat{H}(t) = \int_0^t \frac{J(s)}{Y(s)} dN(s) \qquad \ldots (16)$$

dikenal dengan estimator Nelson-Aalen untuk fungsi hazard kumulatif H(t).

Jika $T_{(1)}$, $T_{(2)}$, … adalah statistik terurut yang menyatakan waktu-waktu kematian maka estimator Nelson Aalen untuk H(t0 dapat dimyatakan sebagai

$$\tilde{H}(t) = \begin{cases} 0 & \text{, jika } t \le T_{(i)} \\ \sum_{T_{(i)} \le t} \frac{d_i}{Y_n} T_{i(i)} & \text{, jika } t \ge T_{(i)} \end{cases} \qquad \ldots (17)$$

dengan

d_i adalah indikator kematian untuk individu ke-i

$Y_n(T_{(1)},)$ adalah banyaknya individu yang hidup sebelum $T_{(1)}$. (dalam Klein (1997, h.86).

Sedangkan

$$W(t) = \int_0^t \frac{J(s)}{Y(s)} dM(s) \qquad \ldots (18)$$

adalah integral stokastik untuk proses J(u)/Y(u) yang dapat diramalkan berkenaan dengan suatu martingale. Jadi W(t) juga martingale, sehingga W(t) dapat dipandang sebagai noise random atau suatu statistik yang tidak diperlukan dalam estimasi (dalam Klein (1997), h.76).

Sementara itu

$$H^*(t) = \int_0^t J(s)h(s)ds \qquad \ldots (19)$$

adalah besaran random untuk data tersensor yang sama dengan H(t) dalam suaru range dimana data diperoleh dengan mengabaikan W(t) sehingga $\hat{H}(t)$ adalah estimator nonparametrik untuk besaran random $H^*(t)$. (dalam Klein (1997), h.76).

Karena W(t) = $\hat{H}(t) - H^*(t)$ sebuah martingale maka E($\hat{H}(t)$)= E($H^*(t)$). Proses variasi yang dapat diramalkan untuk W(t) digunakan dengan menggunakan persamaan (11), yaitu

$$\langle W\rangle(t) = \int_0^t \left(\frac{J(s)}{Y(s)}\right)^2 d\langle M\rangle(s) = \int_0^t \left(\frac{J(s)}{Y(s)}\right)^2 Y(s)h(s)ds = \int_0^t \left(\frac{J(s)}{Y(s)}\right)^2 h(s)ds \quad \ldots (20)$$

FUNGSI-FUNGSI KERNEL

Definisi

Fungsi kernel $K : R \rightarrow R$ adalah fungsi yang simetrik terhadap titik pusat 0 dan mempunyai sifat

$$\int_{-\infty}^{\infty} K(x)dx = 1 \quad \ldots (21)$$

Dari definisi 4.1, jika K(.) adalah fungsi nonnegative maka K(.) dapat diartikan sebagai suatu fungsi padat peluang (fungsi densitas).

Fungsi-fungsi kernel K(.) mempunyai support [-1,1] dan memenuhi syarat-syarat berikut:

$$(1) \int_{-1}^{1} \mathrm{K(u)du} = 1, \quad (2) \int_{-1}^{1} \mathrm{uK(u)du} = 0, \quad (3) \int \mathrm{u^2K(u)du} = \alpha \neq 0, \quad (4) \int_{-1}^{1} \mathrm{K^2(u)du} < \infty \quad \ldots (22)$$

(dalam Eubank (1988, h.111).

Beberapa macam fungsi kernel yang telah dikenal dan sering digunakan dalam estimasi statistik nonparametrik adalah sebagai berikut :

$$\begin{array}{lll} \text{1. Uniform} & : K(u) = \tfrac{1}{2} I(|u| \le 1) & \\ \text{2. Triangle} & : K(u) = (1 - |u|) I(|u| \le 1) & \\ \text{3. Epanechnikov} & : K(u) = \tfrac{3}{4}(1-u^2) I(|u| \le 1) & \\ \text{4. Quantic} & : K(u) = \tfrac{15}{16}((1-u^2))^2 I(|u| \le 1) & \dots (23) \\ \text{5. Triweigh} & : K(u) = \tfrac{35}{32}((1-u^2))^3 I(|u| \le 1) & \\ \text{6. Gaussian} & : K(u) = \dfrac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}u^2) & \\ \text{7. Cosinus} & : K(u) = \tfrac{\pi}{4} \cos\left(\dfrac{\pi u}{2}\right) I(|u| \le 1) & \end{array}$$

(dalam Hardle (1991), h.45).

Dalam estimasi statistik nonparametric menggunakan kernel K(.) biasa digunakan estimator-estimator dengan pembobotan, yaitu

$$K_b(x) = \frac{1}{b} K\left(\frac{x}{b}\right) \qquad \dots (24)$$

dengan b adalah parameter smoothing yang mengatur tingkat kemulusan untuk kernel-smoorhers yang disebut bandwidth b. Jika fungsi kernel K(u) mempunyai support [-1,1] maka fungsi K(u/b) mempunyai support [-b,b].

ESTIMATOR KERNEL UNTUK FUNGSI HAZARD

Estimator Nelson-Aalen $\hat{H}(t)$ seperti dinyatakan dalam persamaan (17) digunakan untuk mengestimasi fungsi hazard kumulatif H(t). Slope dari estimator ini ternyata memberikan suatu estimasi yang kasar untuk fungsi hazard h(t). (dalam Klein (1997). H. 152). Oleh karena itu, pada bagian ini akan digunakan kernel yang diperhalus untuk mengestimasi fungsi hazard h(t).

Estimator-estimator kernel yang diperhalus untuk h(t) didasarkan pada estimator Nelson-Aalen $\hat{H}(t)$ dan varian $\hat{V}(\tilde{H}(t))$ untuk data tersensor. Suatu estimator kernel diperhalus untuk h(t) adalah rata-rata terbobot untuk $\Delta\tilde{H}(t_j)$ atas waktu-waktu kejadian yang mendekati t. Kedekatan ditentukan oleh suatu bandwidth b sedemikan hingga waktu-waktu kejadian dalam [t-b,t+b) termuat dalam suatu rata-rata terbobot yang mengestimasi h(t). Suatu bandwidth dipilih untuk meminimalkan ukuran kesalahan kuadrat rata-rata (MSE) atau untuk memberikan derajat kemulusan.

Sebuah estimator kernel diperhalus untuk fungsi hazar h(t) adalah didefinisikan untuk semua titik-titik waktu $t \ge 0$. Untuk titik-titik waktu t dengan $b \le t \le T_j - b$. Sebuah estimator kernel diperhalus untuk h(t) berdasarkan fungsi kernel K(.) diberikan oleh

$$\hat{h}(t) = \tfrac{1}{b}\sum_{j=1}^{J} K\left(\frac{t-T_j}{b}\right)\Delta\tilde{H}(t_j) \qquad \ldots (25)$$

Varians untuk $\hat{h}(t)$ diestimasi dengan

$$\hat{\sigma}^2(\hat{h}(t)) = \tfrac{1}{b^2}\sum_{j=1}^{J} K\left(\frac{t-T_j}{b}\right)^2 \Delta\hat{V}(\tilde{H}(t_j)) \qquad \ldots (26)$$

(dalam Klein (1997), h. 153).

Jika persamaan (17) disubstitusikan pada persamaan (25) dan (26) maka diperoleh

$$\hat{h}(t) = \tfrac{1}{b}\sum_{j=1}^{J} K\left(\frac{t-T_j}{b}\right)\frac{d_j}{Y_n(T_j)} \qquad \ldots (27)$$

dan

$$\hat{\sigma}^2(\hat{h}(t)) = \tfrac{1}{b^2}\sum_{j=1}^{J} K\left(\frac{t-T_j}{b}\right)^2 \frac{d_j}{Y_n^2(T_j)} \qquad \ldots (28)$$

ESTIMATOR KERNEL UNTUK FUNGSI HAZARD PROSES MENGHITUNG

Ramlau-Hansen (1983) mengestimasi fungsi hazard proses menghitung menggunakan fungsi kernel untuk estimator Nelson-Aalen nonparametrik diperhalus untuk fungsi hazard kumulatif dengan mengasumsikan bahwa model intensitas Aalen berlaku untuk t $\in$ [0,1],

Contoh

Misalkan T_1, T_2, …, T_n adalah waktu-waktu kegagalan atau kematian berdistribusi identik independen dengan nilai-nilai $T_i \in [0,1]$ dan h adalah fungsi hazard dengan

$$F(1) = 1 - \exp\left(-\int_0^t h(s)ds\right) < 1.$$

Misalkan C_1, C_2, …, C_n adalah waktu-waktu sensor berdistribusi identik independen yang berkorespondensi dengan fungsi distribusi G serta asumsikan bahwa waktu-waktu sensor independen dengan waktu-waktu kegagalan atau kematian dan G(1-) < 1.

Banyaknya kegagalan $N_n(t) = \sum_{j=1}^{n} I(T_j \le t, T_j \le C_j)$ adalah suatu proses menghitung, karena memenuhi ketiga syarat dari definisi 2.1.

Proses intensitas $\lambda_n(t) = h(t)Y_n(t)$ dan $Y_n(t) = \sum_{j=1}^{n} I\{T_j > t, C_j < t\}$ menyatakan banyaknya individu yang hidup sebelum waktu t.

Suatu estimator untuk fungsi hazard kumulatif $H_n^*(t) = \int_0^t h(s)I\{Y_n(s) > 0\}ds$ menjadi estimator Nelson-Aalen $H_n(t) = \int_0^t \frac{J_n(s)}{Y_n(s)} dN_n(s) = \sum_{T_j \le t} \frac{d_j}{Y_n(T_j)}$ dengan d_j adalah indicator kematian untuk individu ke-j, yaitu $d_j = \begin{cases} 1, & \text{jika individu ke - j mati} \\ 0, & \text{jika individu ke - j tersensor} \end{cases}$.

Estimator kernel untuk fungsi hazard yang berkorespondensi adalah

$$\hat{h}_n(t) = \frac{1}{b_n}\int_0^t K\left(\frac{t-s}{b_n}\right) d\hat{H}_n(s) = \frac{1}{b_n}\sum_{j=1}^{n} K\left(\frac{t-T_j}{b_n}\right)\frac{d_j}{Y_n(T_j)} \qquad \ldots (27)$$

Definisi

Misalkan T_1, T_2, …, T_n adalah waktu-waktu kegagalan atau kematian berdistribusi identik independen dengan nilai-nilai $T_i \in [0,1]$ dan h adalah fungsi hazard dimana $F(1) = 1 - \exp\left(-\int_0^t h(s)ds\right) < 1$. Misalkan C_1, C_2, …, C_n adalah waktu-waktu sensor berdistribusi identik independen yang berkorespondensi dengan fungsi distribusi G serta asumsikan bahwa waktu-waktu sensor independen dengan waktu-waktu kegagalan atau kematian dan $G(1-) < 1$.

Jika K(.) adalah sebuah fungsi terbatas dengan integal 1 dan b adalah parameter positif maka estimator kernel untuk fungsi hazard yang berkorespondensi didefinisikan sebagai

$$\hat{h}(t) = \tfrac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right) d\hat{H}(s) \qquad \ldots (28)$$

dengan $\hat{H}(t)$ adalah estimator Nelson-Aalen untuk fungsi hazard kumulatif H.

SIFAT-SIFAT ESTIMATOR KERNEL UNTUK FUNGSI HAZARD

Teorema 1

Misalkan

$$K_b(s) = \frac{1}{b}K\left(\frac{s}{b}\right) \text{ dan } \mathrm{j(s) = E(J(s))} \text{ maka } \mathrm{E(\hat{h}(t)) = E(h}^*(t)) = K_b * (hj)(t) \qquad \dots (29)$$

dengan * menyatakan konvolusi terurut.

Selanjutnya

$$\hat{\sigma}^2(t) = E\left[\hat{h}(t) - h^*(t)\right]^2 = \tfrac{1}{b}\int_{-1}^{1} K^2(u)h(t-bu)E\left(\frac{J(t-bu)}{Y(t-bu)}\right)du \text{ untuk } \mathrm{t} \in [\mathrm{b}, 1-\mathrm{b}] \qquad \dots (30)$$

Estimator takbias untuk $\sigma^2(t)$ adalah $\hat{\sigma}^2(t) = \dfrac{1}{b^2}\displaystyle\int_{-1}^{1} K^2\left(\frac{t-s}{b}\right)\frac{J(s)}{Y^2(s)}dN(s)$ … (31)

Bukti :

(1) Akan dibuktikan $E(\hat{h}(t)) = E(h^*(t)) = K_b * (hj)(t)$

$$E(\hat{h}(t)) = E\left(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right)d\hat{H}(s)\right)$$

$$= \mathrm{E}\left(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right)d\left(\int_0^s \frac{J(t)}{Y(t)}dN(t)\right)\right)$$

$$= \mathrm{E}\left(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right)\frac{J(s)}{Y(s)}dN(s)\right)$$

$$= \mathrm{E}\left(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right)d\left(\int_0^s \frac{J(t)}{Y(t)}dN(t)\right)\right)$$

Jika persamaan (13) disubstitusikan kedalam $E(\hat{h}(t))$ maka akan diperoleh

$$E(\hat{h}(t)) = E\left(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right)J(s)h(s)\right)ds + E\left(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right)\frac{J(s)}{Y(s)}dM(s)\right)$$

Karena $\dfrac{1}{b}\displaystyle\int_0^1 K\left(\frac{t-s}{b}\right)\frac{J(s)}{Y(s)}dM(s)$ adalah martingal*e* rata - rata nol, maka

$$E(\hat{h}(t)) = E(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right) J(s)h(s))ds$$

Karena $K_b(s) = \frac{1}{b}K\left(\frac{s}{b}\right)$ dan $\mathrm{j(s) = E(J(s))}$ maka $\mathrm{E}(\hat{\mathrm{h}}(\mathrm{t})) = \mathrm{K_b} * (ht)(j)$ … (*)

$$E(h^*(t)) = E(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right) dH^*(s)$$

$$= \mathrm{E}(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right) d(\int_0^s h(t)J(t)dt)$$

$$= \mathrm{E}(\frac{1}{b}\int_0^1 K\left(\frac{t-s}{b}\right) h(s)J(s)ds)$$

Karena $K_b(s) = \frac{1}{b}K\left(\frac{s}{b}\right)$ dan $\mathrm{j(s) = E(J(s))}$ maka $\mathrm{E(h^*(t)) = K_b} * (ht)(j)$ … (**)

Dari (*) dan (**) terbukti bahwa $E(\hat{h}(t)) = E(h^*(t)) = K_b * (hj)(t)$

(2) Akan dibuktikan

$$\hat{\sigma}^2(t) = E\left[\hat{h}(t) - h^*(t)\right]^2 = \tfrac{1}{b}\int_{-1}^{1} K^2(u)h(t-bu)E(\frac{J(t-bu)}{Y(t-bu})du \ \text{ untuk } t \in [b, 1-b]$$

$$\hat{\sigma}^2(t) = E\left[\hat{h}(t) - h^*(t)\right]^2$$

$$= \frac{1}{b}\int_{-1}^{1} K\left(\frac{t-s}{b}\right) d(\hat{H} - H^*)(s))^2$$

Karena $\hat{\mathrm{H}} - \mathrm{H}^*$ adalah martingale integrabel kuadrat dengan varians proses

$$\left\langle \hat{H} - H^* \right\rangle(t) = \int_0^t h(s)\frac{J(s)}{Y(s)}ds$$

maka

$$E\left(\hat{h}(t) - h^*(t)\right)^2 = E\left[\frac{1}{b^2}\int_{-1}^{1} K^2\left(\frac{t-s}{b}\right)h(s)J(s)ds\right]$$

$$\hat{\sigma}^2(\mathrm{t}) = E\left[\frac{1}{b}\int_{-1}^{1} K^2(u)h(t-ub)E\left[\frac{J(t-ub)}{Y(t-ub)}\right]du\right]$$

(3) Akan dibuktikan bahwa estimator takbias untuk $\sigma^2(t)$ adalah

$$\hat{\sigma}^2(t) = \frac{1}{b^2}\int_{-1}^{1} K^2\left(\frac{t-s}{b}\right)\frac{J(s)}{Y^2(s)}dN(s)$$

Bukti:

$$E\left[\hat{\sigma}^2(t)\right] = E\left[\frac{1}{b^2}\int_0^1 K^2\left(\frac{t-s}{b}\right)\frac{J(s)}{Y^2(s)}dN(s)\right]$$

$$= E\left[\frac{1}{b^2}\int_0^1 K^2\left(\frac{t-s}{b}\right)\frac{J(s)}{Y^2(s)}(h(s)Y(s)ds + dM(s)\right]$$

$$= E\left[\frac{1}{b^2}\int_0^1 K^2\left(\frac{t-s}{b}\right)\frac{J(s)}{Y^2(s)}(h(s)Y(s)ds\right] + E\left[\frac{1}{b^2}\int_0^1 K^2\left(\frac{t-s}{b}\right)\frac{J(s)}{Y^2(s)}dM(s)\right]$$

Karena $\frac{1}{b^2}\int_0^1 K^2\left(\frac{t-s}{b}\right)\frac{J(s)}{Y^2(s)}dM(s)$ adalah martingale rata - rata nol, maka

$$E\left[\hat{\sigma}^2(t)\right] = E\left[\frac{1}{b^2}\int_0^1 K^2\left(\frac{t-s}{b}\right)\frac{J(s)}{Y(s)}h(s)ds\right]$$

$$= E\left[\frac{1}{b}\int_0^1 K^2(u)\frac{J(t-ub)}{Y(t-ub)}h(t-ub)du\right]$$

$$= \sigma^2(t)$$

Teorema

(a) Jika fungsi hazard h kontinu pada titik t dan $j_n = E(J_n) \to 1$ secara seragam pada persekitaran dari t maka $E[h_n(t)] \to h(t)$ bila $n \to \infty$

(b) Jika (1) $nE\left[\frac{J_n(t)}{Y_n(t)}\right] \to \frac{1}{\tau(t)}$ secara seragam dalam persekitaran dari t dan

(2) h dan τ kontinu pada titik t

maka $\sigma^2(t) = E\left[\hat{h}_n(t) - h^*(t)\right]^2 = \frac{1}{nb_n}\left[\frac{h(t)}{\tau(t)}\int_{-1}^{1} K(u)du + \circ\left(\frac{1}{nb_n}\right)\right]$ bila $n \to \infty$

Bukti:

(1) Akan dibuktikan $E\left[\hat{h}_n(t)\right] \to h(t)$

Dari bukti teorema 8.1 diketahui bahwa

$$E(\hat{h}(t)) = E(\frac{1}{b_n}\int_0^1 K\left(\frac{t-s}{b_n}\right) J_n(s)h(s))ds$$

$$= E\left[\int_{-1}^{1} K(u)h(t - ub_n) J_n(t - ub_n)du\right]$$

Jika $n \to \infty$ dan $b_n \to 0$ maka $E(\hat{h}(t)) = h(t)E\left[J_n(t)\int_{-1}^{1} K(u)du\right] = h(t)(1)(1) = h(t)$

(2) Akan dibuktikan bahwa

$$\sigma^2(t) = E\left[\hat{h}_n(t) - h^*(t)\right]^2 = \frac{1}{nb_n}\left[\frac{h(t)}{\tau(t)}\int_{-1}^{1} K(u)du + \circ\left(\frac{1}{nb_n}\right)\right] \text{ bila } n \to \infty$$

Bukti:

$$\sigma^2(t) = E\left[\hat{h}_n(t) - h^*(t)\right]^2$$

$$= \frac{1}{nb_n}\left[\int_{-1}^{1} K^2(u)h(t - ub_n)E\left[\frac{J_n(t - ub_n)}{Y_n(t - ub_n)}\right]du\right]$$

$$= \frac{1}{nb_n}\left[h(t - ub_n)nE\left[\frac{J_n(t - ub_n)}{Y_n(t - ub_n)}\right]\right]\int_{-1}^{1} K^2(u)du$$

$$\text{Karena } \mathrm{nE}\left[\frac{\mathrm{J_n}(t)}{Y_n(t)}\right] \to \frac{1}{\tau(t)} \text{ maka } \sigma^2(t) = \frac{1}{nb_n}\left[h(t)\left(\frac{1}{\tau(t)}\right)\right]\int_{-1}^{1} K^2(u)du + \circ\left(\frac{1}{nb_n}\right)$$

Teorema

Jika (1) untuk setiap $\varepsilon > 0$ berlaku $\int_0^1 \mathrm{T_n}(s)^2 I\{|T_n(s)| > \varepsilon\}\lambda_n(s)ds \xrightarrow{p} 0$ dan

(2) $\int_0^1 \mathrm{T_n^2}(s)\lambda_n(s)ds \xrightarrow{p} 1$ bila $\mathrm{n} \to \infty$

maka $\mathrm{M_n}(t) \xrightarrow{d} Z \sim N(0,1)$

Bukti : (lihat Liptser dan Shiryayev (1980))

Teorema

Jika diasumsikan bahwa (1) $\mathrm{n}\frac{\mathrm{J_n}}{\mathrm{Y_n}} \to \frac{1}{\tau}$ seragam dalam persekitaran dari t bila $\mathrm{n} \to \infty$

(2) fungsi-fungsi h dan τ kontinu pada titik t

maka $\sqrt{\mathrm{nb_n}}\left(\hat{h}_n(t) - h_n^*(t)\right)$ konvergen dalam distribusi pada distribusi normal dengan

rata-rata = 0 dan varians $= \frac{\mathrm{h(t)}}{\tau\mathrm{(t)}}\int_{-1}^{1} K^2(u)du$ bila $\mathrm{n} \to \infty$, $\mathrm{b_n} \to 0$, dan $\mathrm{nb_n} \to \infty$

Bukti:

Karena $\sqrt{nb_n}\left(\hat{h}_n(t) - h_n^*(t)\right) = \int_0^1 T_n(s)dM_n(s) = \widetilde{M}_n(1)$ maka untuk membuktikan teorema 7.4 cukup ditunjukkan bahwa kedua syarat dari teorema 7.3 di atas dipenuhi. *Karena* $T_n(s) = \sqrt{nb_n}K\left(\frac{t-s}{b_n}\right)\frac{J_n(s)}{Y_n(s)}$ maka untuk sebarang $\varepsilon > 0$ berlaku

$$|T_n(s)| > \varepsilon \Leftrightarrow \left|K\left(\frac{t-s}{b_n}\right)\right| n\frac{J_n(s)}{Y_n(s)} > \sqrt{nb_n}\varepsilon$$

Karena $\mathrm{b_n} \to 0$, $\mathrm{nb_n} \to \infty$, dan $\mathrm{n}\frac{\mathrm{J_n}}{\mathrm{Y_n}} \xrightarrow{p} \frac{1}{\tau}$ seragam dalam persekitaran dari t dan $\frac{1}{\tau}$

terbatas pada persekitaran dari t, maka $\mathrm{I}\{|\mathrm{T_n}(s)| > \varepsilon\} \xrightarrow{p} 0$ seragam pada [0,1].

Dari definisi $\mathrm{T_n}(s)$ dikatahui bahwa $\mathrm{E}\left[\int_0^1 \mathrm{T_n^2}(s)\lambda_n(s)ds\right] < \infty$ dan $\lambda_n(s) = Y_n(\mathrm{s})\mathrm{h}(\mathrm{s})$. berarti

Akibatnya $\int_0^1 \mathrm{T_n^2}(s)\,\mathrm{I}\{|\mathrm{T_n}(s)| > \varepsilon\} Y_n(\mathrm{s})\mathrm{h}(\mathrm{s})\mathrm{ds} \xrightarrow{\mathrm{p}} 0$

syarat (1) dari teorema 7.3 dipenuhi.

Karena $\int_0^1 T_n^2(s)\lambda_n(s)ds = \frac{1}{b_n}\int_0^1 K^2\left(\frac{t-s}{b_n}\right) n \frac{J_n(s)}{Y_n(s)} ds$ maka

$$\int_0^1 T_n^2(s)\lambda_n(s)ds = \int_0^1 K^2(u)h(t-ub_n)n\frac{J_n(t-ub_n)}{Y_n(t-ub_n)}du$$

Karena $\mathrm{n} \to \infty, \mathrm{b_n} \to 0, \mathrm{nb_n} \to \infty$ maka $\int_0^1 \mathrm{T_n^2}(s)\lambda_n(s)ds \xrightarrow{p} \frac{h(t)}{\tau(t)}\int_{-1}^1 K(u)du.$

Ini berarti syarat kedua dari teorema 7.3 dipenuhi.

Dengan kata lain terbukti bahwa $\sqrt{\mathrm{nb_n}}\left(\hat{h}_n(t) - h_n^*(t)\right)$ konvergen dalam distribusi

pada distribusi normal dengan rata - rata 0 dan varians $= \frac{\mathrm{h(t)}}{\tau(\mathrm{t})}\int_{-1}^1 K(u)du$ bila

$\mathrm{n} \to \infty, \mathrm{b_n} \to 0, \mathrm{nb_n} \to \infty$

DAFATAR PUSTAKA

Eubank, R. L. (1988). *"Spline smoothing and Nonparametric Regression"*. New York : Marcel Dekker.

Gasser, Th., Muller, H. G. (1979). "Kernel estimation of regression functions". In *Smoothing Teqniques for Curve Estimation.* Gesser, Th. Dan Rosenblatt, M. (eds.). Heidelberg : Springer.

Klein, J. P,. Moeschberger, M.L. (1997). *"Survival Analysis"* : Techniques for Cencored and Trucated Data". New York : Springer-Verlag.

Lawless, J.F. (1982). *"Statistical Models and Methods for Lifetime Data"*. New York: John Wiley & Sons.

Ramlau-Hansen, H. (1983). "Smoothing Counting Processes by Mean of Kernel Function". *The Annals of Statistics, 11, 453-466.*