فرآيندهاي شمارشي و تصادفي[1]

روشهاي مبتني بر فرايندهاي تصادفي ، اولين بار توسط آلن[2] در سال ۱۹۷۵ و با استفاده از فرآيندهاي شمارشي بكار گرفته شدند . وي با تركيب عناصري از جمع هاي تصادفي و نظريه مارتينگلها در قالبي پيوسته _ زمان روشي را ابداع كرد كه به شمارش تعداد مرگها تا زمان t مي پردازد . فرآيند شمارشي مورد نظر او كه با $\{N(t) , t \geq 0\}$ نمايش داده ميشود داراي مشخصات زير است :

$$N(0) = 0$$

$$P[N(t) < \infty] = 1$$

$$N(t) = \sum_{i=1}^{n} N(t_i)$$

اين نظريه بعدها توسط اندرسن[3] و بورگان[4] در سال ۱۹۹۰ توسعه يافت و به يكي از روشهاي اصلي تحليل بقاء تبديل شد . عليرغم اينكه كاربرد اين روشها از پيچيدگي هاي بيشتري نسبت به روشهاي كاكس ، نلسون –آلن و كاپلان – ماير برخوردار است اما پاسخها دقيقتر بوده و موارد ناتمامي و بريدگي به خوبي در آنها مورد نظر قرار ميگيرند. در سال ۱۹۹۰ فلمينگ[5] و هارينگتون[6] كتابي در زمينه استفاده از فرآيندهاي شمارشي در عرصة آناليز بقاء منتشر ساختند و در آن تحليل بقاء را با كمك مارتينگلها تئوريزه نمودند .

در اين روشها فرآيند شمارشي $\{N(t) , t \geq 0\}$ بعنوان يك فرآيند تصادفي با $N(0)=0$ و $P[N(t) < \infty]=1$ در نظر گرفته ميشود . $N(t)$ از راست پيوسته با افزايشهاي يك واحدي است و نشان دهندة تعداد مرگها تا زمان t مي باشد .

---

[1] Counting and Stochastic Processes
[2] Aalen
[3] Andersen
[4] Borgan
[5] Fleming
[6] Harrington

این فرآیند به سادگي تعداد مرگهايي كه تا زمان t رخ داده اند را شمرده و موارد ناتمام قبل از زمان t را در نظر مي گيرد.

اطلاعات مربوط به متغيرهاي مستقل از زمان همچون سن ، جنس و تيمار در زمان صفر و نيز متغيرهاي وابسته به زمان را تا زمان t *تاریخچه* (History) فرآيند شمارشي در زمان t مینامند و آنرا با F_t نشان ميدهيم . با گذشت زمان بطور طبيعي اطلاعات بيشتري حاصل خواهد شد و لذا براي $s \le t$ نتيجه ميشود كه :

$$F_s \subseteq F_t$$

در حالتي كه داده ها از راست ناتمام هستند F_t (تاريخچه در زمان t) عبارتست از اطلاعات مربوط به زوج (T_i , δ_i) و F_{t^-} يعني تاريخچه تا زماني درست قبل از زمان t . تاريخچه $\{F_t , t \ge 0\}$ دنباله اي وابسته به مشاهدات اين فرآيند شمارشي است . براي داده هاي ناتمام از راست اگر X_i زمان مرگ (شكست) و C_i زمان ناتمامي و مستقل از يكديگر باشند شانس اينكه مرگ (شكست) در زمان t رخ دهد به شرطي كه تاريخچه تا زمان قبل از t در دست باشد عبارتست از :

$$\Pr = [t \le T_i \le t + dt, \delta_i = 1 | F_{t^-}] =$$

$$\begin{cases} \Pr[t \le X_i < t + dt, C_i > t + dt_i | X_i \ge t, C_i \ge t] = h(t)dt & ; \quad T_i \ge t \\ 0 & ; \quad T_i < t \end{cases}$$

براي هر فرآيند شمارشي ، $dN(t)$ را تغييرات $N(t)$ در يك زمان بسيار كوتاه $[t , t + dt]$ در نظر مي گيريم ، بنابراين :

$$dN(t) = N[(t + dt)^- - N(t^-)]$$

كه در آن $N(t^-)$ تعداد مرگ و مير درست قبل از زمان t است. پس در داده هاي ناتمام از راست $dN(t)$ برابر يك است اگر مرگي در t رخ داده باشد و در غير اينصورت صفر است . حال اگر $y(t)$ را بعنوان فرايند شمارشي تعداد افراد تحت مطالعه در زمان $T_i \ge t$ در نظر بگيريم آنگاه :

$E[dN(t) \mid Ft^{-}]$ +dt و t + dt تعداد مشاهدات با زمان تا تمامي بزرگتر از $\mid Ft^{-}] = y(t)\, h(t)\, dt$

$=E[t \le Xi \le t$

فرآیند $\lambda(t) = y(t)\, h(t)$ ، را فرآیند شدت یا شتاب فرآیند شمارشي $N(t)$ مینامیم[7]. $\lambda(t)$ خود یك فرآیند تصادفي وابسته به اطلاعات موجود در فرآیند تاریخچه است . حال فرآیند شدت تجمعي[8] را بشكل زیر تعریف مي كنیم :

$$\Lambda(t) = \int_0^t \lambda(s)ds \qquad ; \quad t \ge 0$$

كه در آن $E[\Lambda(t) \mid Ft^{-}] = \Lambda(t)$. فرآیند تصادفي $M(t) = N(t) - \Lambda(t)$ را مارتینگل فرآیند شمارشي مینامند و در اینصورت :

$$E[dM(t) \big| F_{t^-}] = E[dN(t) - d\Lambda(t) \big| F_{t^-}] = E[dN(t) \big| F_{t^-}] - E[\lambda(t)dt \big| F_{t^-}] = 0$$

زیرا میدانیم اگر در یك فرآیند تصادفي مقدار مورد انتظار آن در زمان t بشرط وجود اطلاعات در زمان s $(s<t)$ با مقدار آن در زمان s برابر باشد یك مارتینگل خوانده میشود . پس $M(t)$ یك مارتینگل است وقتیكه :

$$E[M(t) \big| F_s] = M(s); \qquad \forall s < t$$

یعني تغییرات M در فاصلة s تا t صفر است . $M(t)$ داراي دو جزء است یكي $N(t)$ كه یك تابع پلكاني غیر نزولي است و دیگري $\Lambda(t)$ كه یك تابع غیر نزولي هموار است و آنرا تابع متعادل كنندة[9] فرآیند شمارشي مینامیم . سرانجام در این روش با محاسبه واریانس $dM(t)$ برآوردهایي براي $M(t)$ و $N(t)$ و $\Lambda(t)$ بدست آمده و برآورد بقاء محاسبه میگردد :

$$\hat{S}(t) = \prod_{s=0}^{t} \left[1 - d\hat{H}(s)\right] = \prod_{s=0}^{t} \left[1 - \frac{dN(s)}{y(s)}\right]$$

و تابع درستنمایي براي n مشاهده تا زمان τ عبارتست از :

---

[7] Intensity Process
[8] Cumulative Intensity Process
[9] Compensator

$$L=[\prod_{j=1}^{n}\lambda_j(t)^{dN_j(t)}]\exp[-\sum_{j=1}^{n}\int_0^{\tau}\lambda_j(u)du]$$

که براي داده هاي ناتمام از راست و $\lambda_j(t) = y_j(t)\ h(\tau)$ ، براي $t \le t_j$ مقدار $y_j(t)$ برابر يك است و براي $t > t_j$ مقدار $y_j(t)$ برابر صفر است . و سرانجام :

$$L \propto \left[\prod_{j=1}^{n} h(t_j)^{\delta_j}\right]\exp[-\sum_{j=1}^{n} H(t_j)]$$

که δ_j مقادير صفر و يك را براي زندگی و مرگ مي گيرد .