بسم الله الرحمن الرحيم

١٢٩٨٠٢

دانشگاه شهید بهشتی

دانشکده علوم ریاضی

گروه آمار

پایان‌نامهٔ کارشناسی ارشد آمار بیمه

پیشگویی بهینهٔ فرایندهای پواسون آمیختهٔ مرکب

توسط

اکبر جسور قره‌باغ

استاد راهنما

دکتر محمدقاسم وحیدی اصل

استاد مشاور

دکتر محمد ذکایی

اردیبهشت ۱۳۸۸

تقدیم به

خانواده‌ام

آنهایی که از صمیم قلب دوست دارم

و تقدیم چه استاد بزرگوارم

جناب آقای دکتر وحیدی اصل

که درس‌های زیادی از ایشان آموختم.

قدردانی

با تشکر و سپاس فراوان از استاد ارجمند آقای دکتر وحیدی اصل که با راهنمایی‌های خویش مرا در انتخاب موضوع، تدوین این پایان‌نامه و سایر مراحل تحصیلی دورهٔ کارشناسی ارشد یاری نموده‌اند. از آقای دکتر ذکایی به خاطر راهنمایی‌های ارزنده‌شان تشکر می‌کنم. همچنین از اساتید گرانقدر آقای دکتر فریدروحانی و آقای دکتر پاینده نیز به سبب حضور در جمع داوران سپاسگزارم.

در پایان بر خود لازم می‌دانم از دوست خوبم آقای جلیلیان و همهٔ کسانی که در این راه مرا یاری کردند، قدردانی می‌کنم.

اکبر جسور قره‌باغ

تهران — اردیبهشت ۱۳۸۸

پیشگفتار

شناخت درست واقعیت نه تنها به انسان اجازه می‌دهد که از آن برای فعالیت علمی در زمینه‌های طبیعی و اجتماعی بهره‌گیری کند و خواست خود را تحقق بخشد، بلکه پیشگویی حوادث آینده را نیز ممکن می‌سازد.

با تکیه بر شناخت کافی موجود و قوانین مربوط می‌توان رویدادهای مختلف طبیعی را پیشگویی کرد. گردآوری اطلاعات، تجزیه و تحلیل آنها و نتیجه‌گیری علمی مستدل از آنها با تکیه بر قوانین رشتهٔ مربوط، امکان پیشگویی را فراهم می‌سازند.

شرکت بیمه‌ای را در نظر بگیرید که بیمه‌نامه‌های مختلف را ارائه می‌کند. هر یک از این نوع بیمه‌نامه‌ها یک داشتمان خاص برای بیمه‌گر محسوب می‌شود. طبیعی است که این شرکت در قبال اخذ وجهی به عنوان حق بیمه حاضر به قبول مخاطرهٔ بیمه‌گذار می‌شود و پذیرش مخاطره همراه با ادعاهایی است که در صورت وقوع اتفاق با توجه به بیمه‌نامهٔ صادره به بیمه‌گر می‌رسد. اما ادعاهای رسیده از دو جهت برای بیمه‌گر حایز اهمیت است: زمان ورود ادعاها و اندازهٔ ادعاها. شناخت کافی از این دو موضوع در نهایت منجر به مدلبندی فرایند تعداد ادعاها و در نتیجه مدلبندی صحیح فرایند مخاطرهٔ بیمه‌گر خواهد شد. البته کارهای سخت‌افزاری یک شرکت بیمه محدود به این موارد (اخذ حق بیمه، ارائهٔ بیمه‌نامه و پرداخت ادعاها) نمی‌شود بلکه مقداری از حق بیمه‌های اخذ شده برای به‌دست آوردن سود بیشتر، سرمایه‌گذاری می‌شود.

سؤالی که در اینجا به ذهن می‌رسد این است که شرکت چه میزان از این حق بیمه‌ها را باید سرمایه‌گذاری کند و چه مقداری را برای هزینه‌های جاری و پرداخت ادعاها نزد خود نگه دارد.

اگر فرایند مخاطرهٔ این شرکت به‌طور صحیح مدلبندی شود و یک پیشگوی بهینه از این فرایند داشته باشیم به آسانی می‌توان به این سؤال پاسخ داد. همچنین این پیشگوی بهینه را می‌توان برای پیشگویی مقدار سرمایهٔ شرکت در زمان‌های آتی و زمان ورشکستگی به‌کار برد.

پیشگویی بهینهٔ فرایندهای پواسون آمیختهٔ مرکب

چکیده

فرایند پواسون یکی از مهمترین فرایندهای تصادفی است که برای مدلبندی فرایند تعداد ادعاها به‌کار می‌رود. در این پایان‌نامه مشخصه‌های فرایند تعداد ادعاها را بیان می‌کنیم و همچنین نشان می‌دهیم که باید از فرایند پواسون آمیخته برای مدلبندی فرایند تعداد ادعاهایی که از یک دانشتمان ناهمگن به بیمه‌گر می‌رسد، استفاده کرد.

به‌علاوه، دو تعریف مختلف برای فرایند مخاطره که شامل فرایند مخاطرهٔ کلاسیک و فرایند مخاطره بر اساس فرایندهای نقطه‌ای نشان‌دار ارائه می‌کنیم و به پیشگویی این فرایندها می‌پردازیم.

بهترین پیشگوی نااریب خطی برای فرایند پواسون آمیختهٔ مرکب و ارائهٔ پیشگوی بهینهٔ صریح از جمله ویژگی‌های این پیشگویی خواهند بود.

واژه‌های کلیدی : فرایند پواسون، فرایند پواسون مرکب، فرایند پواسون آمیخته، فرایند نقطه‌ای نشان‌دار، اندازهٔ شدت، فرایند مخاطره، فرایند تعداد ادعاها، پیشگویی، پالایش، زمان ورشکستگی.

فهرست مندرجات

فهرست شکل‌ها

لیست اشکال

نمادها

توزیع متغیر تصادفی X	P_X
توزیع متغیر تصادفی X به شرط Y	$P_{X\|Y}$
تابع کوواریانس فرایند S_t	$R_S(s,t)$
تابع مشخصهٔ X	ϕ_X
تعداد ادعاها تا زمان t	N_t
توزیع پواسون با پارامتر s	$Poi(s)$
توزیع گاما با پارامترهای α و γ	$Ga(\alpha,\gamma)$
پارامتر ساختار فرایند پواسون آمیخته	Θ
زمان رسیدن nامین ادعا	T_n
زمان بین ادعای $n-1$ و n	W_n
فرایند مخاطرهٔ جمعی	S_t
فرایند مازاد	U_t
فضای احتمال پایه	$(\Omega,\mathcal{F},P)$
فضای فرایند نقطه‌ای نشان‌دار	$(E,\mathcal{E})$
مجموعهٔ اعداد حقیقی	$\mathbb{R}$
مجموعهٔ اعداد حقیقی نامنفی	$\mathbb{R}_+$
مجموعهٔ اعداد طبیعی	$\mathbb{N}$
$\mathbb{N}\cup\{0\}$	$\mathbb{N}_0$
$\mathbb{N}\cup\{0,\infty\}$	$\bar{\mathbb{N}}_0$
$\sigma(\sigma(\Theta)\cup\mathcal{F}_t^{\mathcal{N}})$	$\sigma(\Theta)\vee\mathcal{F}_t^{\mathcal{N}}$
$\sigma(\Theta)\cap\mathcal{F}_t^{\mathcal{N}}$	$\sigma(\Theta)\wedge\mathcal{F}_t^{\mathcal{N}}$

فصل ۱

تعریف‌ها و مفاهیم اولیه

در این فصل ابتدا فرایند تصادفی را تعریف می‌کنیم و سپس به معرفی فرایند پواسون همگن و ناهمگن و همچنین فرایند پواسون مرکب می‌پردازیم و در بخش آخر زمان ورشکستگی را توضیح خواهیم داد.

۱.۱ فرایند پواسون

فرایند تصادفی، خانواده‌ای از متغیرهای تصادفی $\{X_t, t \in T\}$ است که بر یک فضای احتمال $(\Omega, \mathcal{F}, P)$ تعریف می‌شوند.

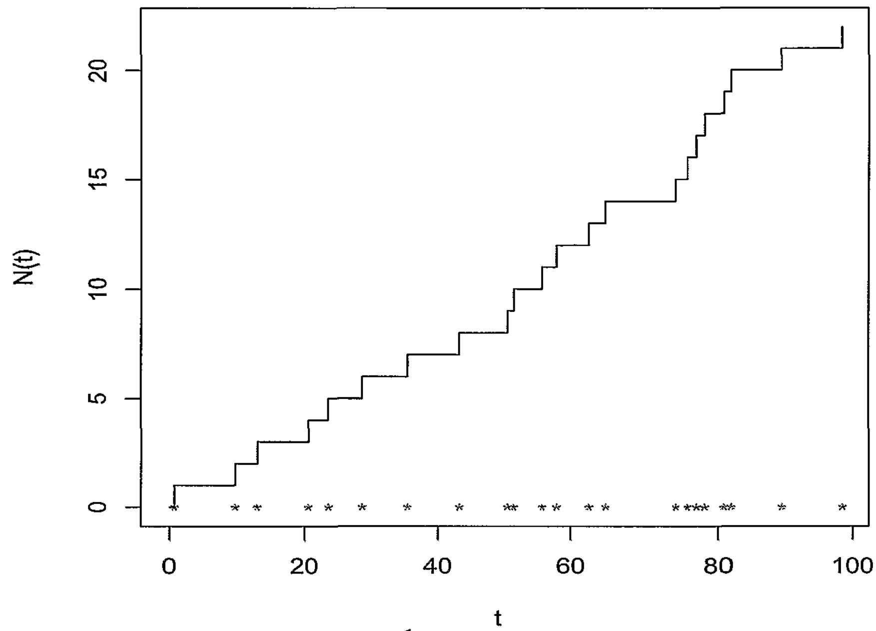
فرایند تصادفی $\{N_t, t \in \mathbb{R}_+\}$، فرایند شمارشی است هرگاه N_t نشان دهندهٔ تعداد پیشامدهایی از نوع خاص باشد که تا زمان t رخ داده‌اند. لذا فرایند شمارشی $\{N_t, t \in \mathbb{R}_+\}$ باید شرایط زیر را داشته باشد.

i) $N_t \geq 0$.

ii) N_t مقادیر صحیح بگیرد.

iii) برای هر $s < t$، $N_s \leq N_t$.

iv) برای $s < t$، $N_t - N_s$ برابر با پیشامدهایی باشد که در بازهٔ $[s, t]$ روی می‌دهند.



شکل ۱.۱. ۱: فرایند پواسون همگن با نرخ $\lambda = ۰.۲$

فرایند شمارشی $\{N_t,\ t \in \mathbb{R}_+\}$ دارای نموهای مستقل است اگر تعداد پیشامدهایی که در بازه‌های زمانی جدا از هم روی می‌دهند مستقل از یکدیگر باشند؛ یعنی

$$P[N_{t+s} - N_s = k_1 | N_r = k_2,\ 0 \leq r \leq s] = P[N_{t+s} - N_s = k_1]$$

و دارای نموهای مانا است هرگاه توزیع تعداد پیشامدهایی که در هر بازهٔ زمانی روی می‌دهند، فقط به طول زمانی بازهٔ مورد نظر بستگی داشته باشد؛ یعنی

$$P[N_{t+s} - N_t = k] = P[N_{l+s} - N_l = k].$$

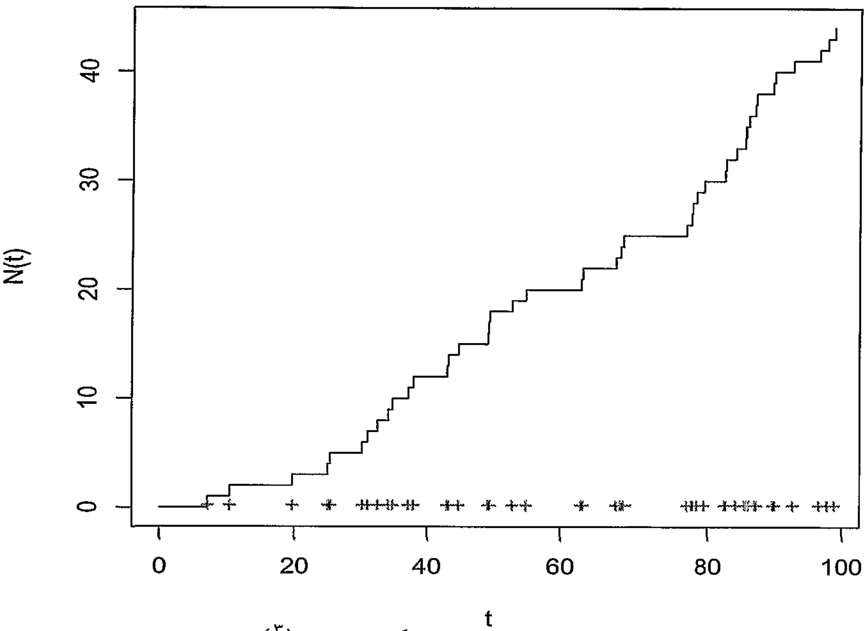
تعریف ۱.۱.۱ فرایند شمارشی $\{N_t, t \geq 0\}$ فرایند پواسون با نرخ λ است (شکل ۱.۱.۱)، هرگاه

i) به‌ازای هر t، $t \geq ۰$، $N_t \geq ۰$.

ii) $\{N_t, t \geq ۰\}$ دارای نموهای مستقل و مانا است.

iii) توزیع تعداد پیشامدها در هر بازه‌ای به طول t، پواسون با میانگین λt است؛ یعنی، برای هر $t, s \geq ۰$ و $n = ۰, ۱, \dots$،

$$P[N_{t+s} - N_s = n] = \frac{e^{-\lambda t}(\lambda t)^n}{n!}.$$



شکل ۱.۱ـ ۲: فرایند پواسون ناهمگن با نرخ $\lambda(t) = t^{(\frac{۳}{۲})}$

قضیهٔ ۱.۱.۱ (کینگمن [۱۹۹۳]) اگر N_iها فرایندهای پواسون مستقل با نرخ λ_i باشند، آنگاه $\sum_{i=۱}^{n} N_i$ نیز فرایند پواسون با نرخ $\lambda = \sum_{i=۱}^{n} \lambda_i$ است.

۱.۱.۱ فرایند پواسون ناهمگن

تعریف ۲.۱.۱ فرایند شمارشی $\{N_t, t \geq ۰\}$ یک فرایند پواسون ناهمگن یا نامانا با تابع شدت $\lambda(t)$ نامیده می‌شود(شکل ۲.۱.۱) هرگاه

i) به‌ازای هر t، $N_t \geq ۰$.

N_t(ii دارای نموهای مستقل است.

$P[N_{t+h} - N_t = ۱] = \lambda(t).h + o(h)$(iii یا $\lim_{h\to ۰} \frac{P[N_{t+h}-N_t=۱]}{h} = \lambda(t)$.

$P[N_{t+h} - N_t \geq ۲] = o(h)$ (iv یا $\lim_{h\to ۰} \frac{P[N_{t+h}-N_t\geq ۲]}{h} = ۰$.

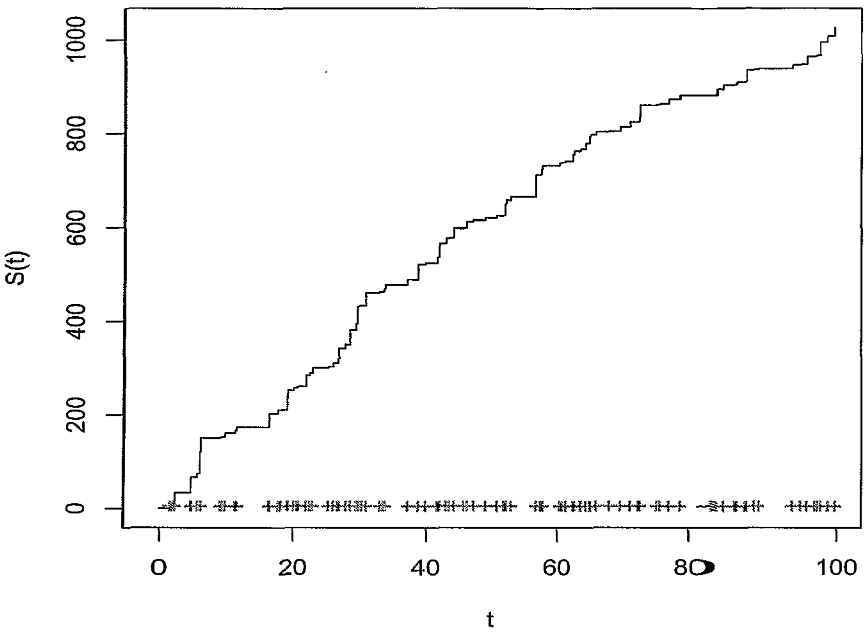
قضیهٔ ۲.۱.۱ اگر $\{N_t, t \geq ۰\}$ فرایند پواسون ناهمگن با تابع شدت $\lambda(t)$ باشد و $m(t)$ را به‌صورت $m(t) = \int_۰^t \lambda(x)dx$ تعریف کنیم، آنگاه به‌ازای هر n، $n = ۰, ۱, ۲, \ldots$

$$P[N_{t+s} - N_t = n] = \frac{e^{-[m(t+s)-m(t)]}[m(t+s) - m(t)]^n}{n!}.$$

۲.۱ فرایند پواسون مرکب

تعریف ۶.۲.۱ اگر $\{X_k\}_{k=۱}^{\infty}$ دنباله ای از متغیرهای تصادفی مستقل و هم‌توزیع با تابع توزیع F و $\{N_t, t \geq ۰\}$ یک فرایند پواسون همگن (ناهمگن) با پارامتر $\lambda > ۰$ و مستقل از دنبالهٔ X_k باشد، آنگاه فرایند تصادفی $S_t = \sum_{k=۱}^{N_t} X_k$ یک فرایند پواسون مرکب نامیده می‌شود (شکل ۳.۱.۱).

از ویژگی های فرایند پواسون مرکب می‌توان به مستقل و مانا بودن نموهای این فرایند اشاره کرد (کارلین و تیلور[۱۹۸۱]). فرایندهای پواسون مرکب برای مدلبندی بسیاری از پدیده‌های فیزیکی به کار می رود. به‌طور مشخص اکثر پیشامدها مطابق با فرایند پواسون رخ



شکل ۱.۱.۳: فرایند پواسون مرکب
(N_t فرایند پواسو_ن با نرخ $\lambda = ۱$ و X_iها iid از توزیع نمایی با پارامتر ۰.۱)

می‌دهند و هر پیشامد با مقداری تصادفی تعیین می‌شود.

نمونه‌هایی از کار_برد این فرایند عبارتند از

- فرض کنید که N_t تعداد ادعاهای رسیده به شرکت بیمه تا زمان t بوده و X_k مقدار kامین ادعا باشد . در این‌صورت S_t مجموع ادعاهای رسیده تا زمان t است . دربارهٔ این مورد در آینده بیشتر بحث خواهد شد.

- فرض کنید که N_t تعداد معاملات در یک بورس عمومی تا لحظهٔ t و X_k تعداد سهام فروخته شده در kامین معامله باشد، در این‌صورت S_t تعداد کل سهام فروخته شده تا لحظهٔ t است.

- فرض کنید که N_t تعداد شوک های اتفاق افتاده در یک سیستم تا لحظهٔ t و X_k مقدار خسارت یا فرسایش ایجاد شده به‌وسیلهٔ kامین شوک باشد، دراین‌صورت S_t مجموع خسارت‌های وارده به سیستم تا لحظهٔ t است.

۱ـ ۲.۱ تابع مشخصهٔ فرایند پواسون مرکب

چون N_t از یک توزیع پواسون با پارامتر λt تبعیت می کند، تابع مولد احتمال آن به‌صورت $g_{N_t}(s) = e^{-\lambda t(1-s)}$ است. همچنین با توجه به اینکه S_t مجموعی تصادفی از متغیرهای تصادفی است، لذا

$$\phi_{S_t}(u) = g_{N_t}[\phi(u)]$$

$$= exp\{-\lambda t[1-\phi(u)\}, \qquad \infty < u < +\infty$$

که $\phi(u) = E[e^{iuX}]$ تابع مشخصهٔ متغیر تصادفی مربوط به اندازهٔ ادعاها است و همچنین تابع مشخصهٔ توأم نیز به‌ازای $0 < u_1 < u_2 < ... < u_n$ به صورت زیر است.

$$\phi_{S_{t_1},...,S_{t_n}}(u_1,...,u_n) = E[\exp(iu_1 S_{t_1} + ... + iu_n S_{t_n}]$$

$$= E[\exp(i\sum_{k=1}^{n} a_k[S_{t_k} - S_{t_{k-1}}])] \qquad (a_k = u_k + ... + u_n)$$

$$= \prod_{k=1}^{n} E[\exp\{ia_k[S_{t_k} - S_{t_{k-1}}]\}] = \prod_{k=1}^{n} E[\exp\{ia_k[S_{(t_k - t_{k-1})}]\}]$$

$$= \prod_{k=1}^{n} \phi_{S(t_k - t_{k-1})}(a_k).$$

البته توجه شود که چون X_iها مستقل و هم‌توزیع هستند این رابطه برقرار است.

۱ـ ۲.۲ تابع توزیع S_t

تابع توزیع S_t را می‌توان بعد از شرطی کردن بر روی مقادیر N_t به طور صریح نوشت:

$$P(S_t \le s) = P\left(\sum_{k=1}^{N_t} X_k \le s\right) = \sum_{n=0}^{\infty} P\left(\sum_{i=1}^{N_t} X_i \le s \,|N_t = n\right) \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

$$= \sum_{n=0}^{\infty} \frac{(\lambda t)^n e^{-\lambda t}}{n!} F^{(n)}(s)$$

(چون N_t از X_k مستقل است)

که

$$F^{(n)}(s) = P(X_1 + ... + X_n \leq s) \qquad , \qquad F^{(0)}(s) = \begin{cases} 1; s \geq 0 \\ 0; s < 0 \end{cases} .$$

۳.۲.۱ مجموع فرایندهای پواسون مرکب

فرایندهای پواسون مرکب دارای ویژگی های مطلوب زیادی هستند. ما اینجا یکی از آنها را شرح می دهیم .

مجموع دو فرایند پواسون مرکب و مستقل، خود نیز فرایند پواسون مرکب است. برای k، $k = 1, 2$، فرض کنیم که S_t^k ، فرایندهای پواسون مرکب باشند و ϕ_k تابع مشخصه مربوط به kامین فرایند باشد. با فرض $S_t = S_t^1 + S_t^2$، ادعا می کنیم که S_t یک فرایند پواسون مرکب با شدت $\lambda = \lambda_1 + \lambda_2$ وتابع مشخصهٔ زیر است.

$$\phi(u) = \frac{\lambda_1}{\lambda_1 + \lambda_2}\phi_1(u) + \frac{\lambda_2}{\lambda_1 + \lambda_2}\phi_2(u) .$$

واضح است که S_t دارای نموهای مستقل مانا است، زیرا هر دو فرایند S_t^1 و S_t^2 دارای نموهای مستقل مانا هستند . لذا تنها لازم است که تابع مشخصهٔ S_t محاسبه شود.

$$\phi_{S_t}(u) = \phi_{S_t^1}(u)\,\phi_{S_t^2}(u) = \exp\{-\lambda t[1 - P_1\phi_1(u) - P_2\phi_2(u)]\}$$

که $\lambda = \lambda_1 + \lambda_2$ و $P_i = \frac{\lambda_i}{\lambda}$، پس S_t یک فرایند پواسون مرکب است.

۳.۱ زمان ورشکستگی

فرایند ادعاهای جمعی یا کل ادعاهای رسیده به بیمه‌گر تا زمان t را با S_t نشان داده و به صورت

$$S_t = \sum_{k=1}^{N_t} X_k$$