نرخ خطر معکوس در توزیع های آمیخته

زهرا عرب برزو[۱] ، غلامرضا محتشمی برزادران [۲]

چکیده:

در این مقاله به معرفی اجمالی از نرخ خطر معکوس و توزیع های آمیخته پرداخته و سپس نرخ خطر معکوس در توزیع های آمیخته، زمان سپری شده از شکست در این توزیع ها را معرفی می کنیم، همچنین دو مدل جمعی و ضربی از نرخ خطر معکوس آمیخته را معرفی می کنیم و نشان می دهیم رفتار نرخ خطر معکوس آمیخته از k زیر جامعه با نرخ خطر معکوس افزایشی همواره افزایشی است.

واژه‌های کلیدی: نرخ خطر، نرخ خطر معکوس، توزیع آمیخته، زمان سپری شده از شکست .

۱ مقدمه

نرخ خطر معکوس یکی از شاخص های قابلیت اعتماد می باشد که در سال های اخیر مورد توجه بسیاری از دانشمندان واقع شده، کیلسون و سومیتا (۱۹۸۲) اولین افرادی بودند که نرخ خطر معکوس را به صورت نسبت تابع چگالی به تابع توزیع بیان کردند و آن را نرخ خطر دو گانه(RHR) نامیدند. که از این خصوصیت بیشتر در علوم قضایی، بیمه و داده های سانسور شده از چپ به منظور پیدا کردن زمان واقعی شکست استفاده می شود.
آلن (۱۹۹۸، ۱۹۹۲) وایز(۱۹۹۰) نشان دادند که علاوه بر عامل زمان، ناهمگنی هایی از جامعه که قابل مشاهده نیستند (فشار، دما، الکتریسیته) روی نرخ خطر تأ ثیر گذارند و در این راستا توزیع های آمیخته را معرفی کردند. بارلو و همکاران (۱۹۶۳) عمل آمیختگی را روی توزیع های مختلف مورد بررسی قرار دادند و در ادامه گورلند و سسارمن (۱۹۹۵) نشان دادند که نرخ خطرآمیخته از k زیر جامعه با نرخ خطر کاهشی همواره کاهشی است.

۲ نرخ خطر معکوس

فرض کنید T یک متغیر تصادفی طول عمر با تابع چگالی $f(t)$ و تابع توزیع $F(t)$، روی بازه (a,b) باشد که در آن $a = inf\{t : F(t) > ۰\}$ و $b = sup\{t : F(t) < ۱\}$ است. احتمال خرابی قطعه در فاصله زمانی $(t - \triangle t, t]$ به شرط این که در فاصله زمانی $[۰,t]$ شکست اتفاق افتاده باشد عبارتست از:

$$\begin{aligned} P(t-\triangle t < T < t \mid T < t) &= \frac{P(t-\triangle t<T<t)}{P(T<t)} \\ &= \frac{F(t)-F(t-\triangle t)}{F(t)}, \end{aligned}$$

متوسط احتمال فوق در فاصله زمانی به طول $\triangle t$ وقتی $\triangle t$ به سمت صفر میل می کند تابع نرخ خطر معکوس

---
[۱] گروه آمار، دانشگاه آزاد مشهد
[۲] دانشیار گروه آمار، دانشگاه فردوسی مشهد

نامیده می شود که آن را با $r(t)$ نشان می دهیم به طوری که

$$\begin{aligned} r(t) &= \lim_{\triangle t\to 0} \frac{P(t-\triangle t<T<t)}{P(T<t)\triangle t} \\ &= \lim_{\triangle t\to 0} \frac{F(t)-F(t-\triangle t)}{\triangle t F(t)} \\ &= \frac{f(t)}{F(t)}, \end{aligned}$$

تعبیر احتمالی تابع نرخ خطر معکوس بیانگر وقوع شکست در بازه زمانی $[t-\triangle t, t]$ است به شرط این که شکست قبل از زمان t اتفاق افتاده باشد.

بلوک و همکاران (۱۹۹۸) نشان دادند توزیع هایی که نرخ خطر افزایشی دارند مانند لگ نرمال، وایبل و گاما با پارامتر($\alpha > 1$) نرخ خطر معکوس آن ها کاهشی است. و اگر نرخ خطر توزیعی کاهشی باشد، نرخ خطر معکوس آن نیز همواره کاهشی است. همچنین نرخ خطر معکوس افزایشی برای یک متغیر تصادفی مثبت وجود ندارد.

۳ توزیع آمیخته

فرض کنید $f_0(.)$،$f_1(.)$،...،$f_n(.)$،... دنباله ای از توابع چگالی باشد که همگی توابع چگالی گسسته یا توابع چگالی احتمال اند. و ممکن است به پارامتر هایی بستگی داشته یا نداشته باشند و p_0،p_1،...،p_n، ... دنباله ای از پارامتر ها باشد که در شرط $\sum_{i=1}^{\infty} p_i = 1$ صدق کند، آن گاه تابع توزیع و تابع چگالی آمیخته را به ترتیب زیر داریم:

$$F(t) = \sum_{i=1}^{\infty} p_i F_i(t), \qquad p_i \geq 0.$$

$$f(t) = \sum_{i=1}^{\infty} p_i f_i(t), \qquad p_i \geq 0.$$

مفهوم آمیختگی را می توان تعمیم داد به طوری که اگر $f(t,z)$ خانواده ای از توابع چگالی با پارامتر z باشند، که روی فضای ν تعریف شده است (ممکن است متناهی یا نا متناهی باشد) آن گاه توزیع آمیخته به فرم:

$$F(t,z) = \int_{\nu} F(t,z) dG(z),$$

تعریف می شود که $dG(z)$ در حالت پیوسته تابع چگالی متغیر تصادفی Z است. برای حالت گسسته وقتی Z مقادیر متناهی $\{z_1, ..., z_k\}$ را اختیار کند داریم:

$$f(z) = \sum_{i=1}^{k} f(t, z_i)\pi(z_i),$$

که $\pi(z_i)$ جرم احتمال Z_i است. به طوری که $\sum_{i=1}^{k} p_i = 1$.

۴ نرخ خطر معکوس آمیخته

در قابلیت اعتماد برای برآورد زمان شکست از یک قطعه از توزیع های طول عمر استفاده می شود. در برآورد زمان شکست اگر علاوه بر عامل زمان عوامل دیگری مانند (دما، فشار، الکتریسیته) را در نظر بگیریم آن گاه نرخ خطر معکوس آمیخته با متغیر تصادفی Z را خواهیم داشت. یک حالت خاص از نرخ خطر معکوس آمیخته هنگامی است که $Z = z$ در نظر می گیریم و آن را با $r(t,z)$ نشان می دهیم. به طوری که

$$r(t \mid Z = z) = r(t,z).$$

در این حالت تابع توزیع آمیخته با پارامتر z به فرم:

$$\begin{aligned} F_m(t) &= E(F(t,z)) \\ &= \int_a^b F(t,z)g(z)dz, \end{aligned}$$

بیان می شود، نرخ خطر معکوس آمیخته را نیز به صورت زیر داریم:

$$\begin{aligned} r_m(t) &= \frac{\int_a^b f(t,z)g(z)dz}{\int_a^b F(t,z)g(z)dz} \\ &= \int_a^b r(t,z)g(z \mid t)dz, \end{aligned}$$

که در آن

$$\begin{aligned} g(z \mid t) &= g(z \mid T \leq t) \\ &= \frac{F(t,z)g(z)}{\int_a^b F(t,z)g(z)dz}, \end{aligned}$$

و تابع توزیع متناظر با آن به فرم:

$$\begin{aligned} G(z \mid t) &= P(Z \leq z \mid T \leq t) \\ &= \frac{\int_a^z F(t,u)g(u)du}{\int_a^b F(t,z)g(z)dz}, \end{aligned}$$

تعریف می شود. که همواره برای متغیر آمیخته Z داریم:

$$G(z) = \lim_{t\to\infty} G(z \mid t).$$

اگر $Z \equiv (Z \mid t)$ در نظر بگیریم آن گاه:

$$\begin{aligned} r_p(t) &= E(r(t,Z)) \\ &= \int_a^b r(t,z)g(z)dz, \end{aligned}$$

نرخ خطر معکوس غیر شرطی نامیده می شود.
ژیاهولی و همکاران (۲۰۱۰) نشان دادند که اگر $r(t,z)$ نسبت به z یکنوا باشد همواره $r_p(t) \geq r_m(t)$. همچنین اگر $r(t,z)$ برای همه $t \geq \circ$ نسبت به z افزایشی باشد و $\frac{\partial r(t,z)}{\partial z}$ برای همه $z \geq \circ$ نسبت به t کاهشی باشد آن گاه $r_p(t) - r_m(t)$ نسبت به $t \geq \circ$ کاهشی است.

۵ ویژگی هایی از مدل های نرخ خطر معکوس

امید شرطی از Z (با شرط $T < t$) را با $E(Z \mid t)$ نشان می دهیم و به صورت زیر تعریف می کنیم:

$$E(Z \mid t) = \int_\circ^\infty zg(z \mid t)dz,$$

لم۱: مشتق امید شرطی نسبت به Z برابر است با

$$E'(Z \mid t) = \frac{\int_\circ^\infty zf(t,z)g(z)dz}{\int_\circ^\infty F(t,z)g(z)dz} - r_m(t)E(Z \mid t).$$

اثبات: با مشتق گیری از امید شرطی Z (با شرط $T < t$) نسبت به t داریم:

$$E'(Z \mid t) = \int_\circ^\infty zg'(z \mid t)dz$$

که در آن

$$g'(z \mid t) = \frac{f(t,z)g(z)}{\int_\circ^\infty F(t,z)g(z)dz} - \frac{F(t,z)g(z)r_m(t)}{\int_\circ^\infty F(t,z)g(z)dz},$$

$$= \frac{f(t,z)g(z)}{\int_\circ^\infty F(t,z)g(z)dz} - r_m(t) \times g(z \mid t),$$

$$E'(Z \mid t) = \frac{\int_\circ^\infty zf(t,z)g(z)dz}{\int_\circ^\infty F(t,z)g(z)dz} - r_m(t)E(Z \mid t).$$

در این بخش دو مدل جمعی و ضربی از نرخ خطر معکوس را معرفی می کنیم.

- **مدل جمعی**

فرض کنید نرخ خطر معکوس آمیخته به فرم:

$$r(t,z) = r(t) + z,$$

که $r(t)$ یک تابع مثبت و پیوسته است و آن را نرخ خطر معکوس مبنا می نامیم، در این حالت نرخ خطر معکوس آمیخته را به صورت:

$$\begin{aligned} r_m(t) &= r(t) + \frac{\int_0^\infty zF(t,z)g(z)dz}{\int_0^\infty F(t,z)g(z)dz} \\ &= r(t) + E(Z \mid t). \end{aligned}$$

تعریف می کنیم.

لم۲: مشتق امید شرطی از Z برای یک مدل جمعی همواره تابعی افزایشی نسبت به t است به طوری که:

$$E'(Z \mid t) = var(Z \mid t).$$

اثبات: با مشتق گیری از امید شرطی Z (با شرط $T < t$) نسبت به t و جایگذاری عبارت

$$\begin{aligned} f(t,z) &= r(t,z)F(t,z) \\ &= (z + r(t))F(t,z), \end{aligned}$$

داریم:

$$\begin{aligned} E'(Z \mid t) &= \frac{\int_0^\infty [zr(t)F(t,z) + z^2F(t,z)]g(z)dz}{\int_0^\infty F(t,z)g(z)dz} \\ &- [r(t) + E(Z \mid t)]E(Z \mid t), \\ &= E(Z^2 \mid t)] - [E(Z \mid t)]^2 \\ &= var(Z \mid t) > 0, \end{aligned}$$

که تابعی افزایشی نسبت به t است.

- **مدل ضربی**

فرض کنید نرخ خطر معکوس آمیخته به فرم:

$$r(t,z) = zr(t),$$

تعریف شده باشد در این حالت نرخ خطر معکوس آمیخته را به صورت زیر داریم:

$$\begin{aligned} r_m(t) &= \int_0^\infty r(t,z)g(z \mid t)dz \\ &= r(t)E(Z \mid t), \end{aligned} \tag{1}$$

لم۳: مشتق امید شرطی از Z برای یک مدل ضربی برابر است با

$$E'(Z \mid t) = r(t)var(Z \mid t),$$

که همواره یک تابع افزایشی نسبت به t است.

اثبات: با مشتق گیری از امید شرطی Z(با شرط $T < t$) نسبت به t و جایگذاری

$$f(t,z) = r(t,z)F(t,z) = zr(t)F(t,z).$$

داریم:

$$\begin{aligned} E'(Z \mid t) &= [r(t) + E(Z \mid t)]E(Z \mid t) \\ &- \frac{\int_0^\infty z^2F(t,z)g(z)dz}{\int_0^\infty F(t,z)g(z)dz}, \\ &= r(t)[E^2(Z \mid t) - r(t)E(Z^2 \mid t)], \\ &= r(t)var(Z \mid t) > 0, \end{aligned}$$

که تابعی افزایشی نسبت به t است.

نتیجه ۱: با استفاده از لم ۳ و مشتق گیری از رابطه (۱) داریم:

$$\frac{r'(t)}{r^2(t)} \geq \frac{var(Z \mid t)}{E(Z \mid t)},$$

از این رو گشتاورهای شرطی و تابع نرخ خطر معکوس روی $IRFR$ و $DRFR$ بودن توزیع های آمیخته مؤثرند.

۶ زمان سپری شده از شکست

هم ارز طول عمر باقی مانده در نرخ خطر، فینکلستین (۲۰۰۸) زمان سپری شده از شکست[1] $(T_{w,x})$ را در نرخ

[1] waiting Time

خطر معکوس تعریف کرد که بیانگر مدت زمان سپری شده از شکست است به شرط آن که تا زمان x خرابی اتفاق افتاده باشد که x مقداری ثابت فرض می شود لذا:

$$\begin{aligned}\bar{F}_{w,x}(t) &= P(x-T>t \mid T\leq x)\\ &= \frac{F(x-t)}{F(x)},\end{aligned}$$

میانگین زمان سپری شده از شکست را با (MWT) نشان می دهیم و به صورت زیر تعریف می کنیم

$$\begin{aligned}\mu(x) &= E(x-T>t \mid T<x)\\ &= x-E(T \mid T<x)\\ &= x-\int_0^x \frac{tf(t)dt}{F(x)}\end{aligned}$$

که با انتگرال گیری جزء به جزء داریم:

$$\mu(x)=\int_0^x \frac{F(u)}{F(x)}du,$$

به طور مشابه میانگین زمان سپری شده از شکست در یک توزیع آمیخته را به صورت زیر معرفی می کنیم:

$$\mu(t,z)=\frac{\int_0^x F_m(u)du}{F_m(t)},$$

که با جایگذاری می توان نوشت:

$$\begin{aligned}\mu(t,z) &= \frac{\int_0^x\int_0^\infty F(t,z)g(z)dzdu}{\int_0^\infty F(t,z)g(z)dz}\\ &= \int_0^\infty \mu(t,z)g(z \mid t)dz,\end{aligned}$$

۷ نرخ خطر معکوس گسسته در توزیع های آمیخته

اگر Z یک متغیر تصادفی گسسته با تابع جرم احتمال $\pi(z_i)$ باشد که $Z=z_i$ آن گاه تابع توزیع، تابع جرم احتمال آمیخته و تابع نرخ خطر معکوس را به‌صورت زیر داریم:

$$\begin{aligned}F(t) &= \textstyle\sum_{i=1}^k F(t,z_i)\pi(z_i), \quad \text{تابع توزیع}\\ f(t) &= \textstyle\sum_{i=1}^k f(t,z_i)\pi(z_i), \quad \text{تابع جرم احتمال}\\ r(t) &= \frac{\sum_{i=1}^k f(t,z_i)\pi(z_i)}{\sum_{i=1}^k F(t,z_i)\pi(z_i)},\\ &= \textstyle\sum_{i=1}^k r(t,z_i)\pi(z_i \mid t), \quad \text{نرخ خطر معکوس}\end{aligned}$$

است. که $\pi(z_i \mid t)$ نشان دهنده تابع جرم احتمال $Z=z_i$ با شرط ($T<t$) است و به صورت زیر تعریف می شود:

$$\pi(z_i \mid t)=\frac{F(t,z_i)\pi(z_i)}{\sum_{i=1}^k F(t,z_i)\pi(z_i)}.$$

۸ نرخ خطر معکوس آمیخته از دو خانواده توزیع

یک حالت خاص از آمیختگی، آمیختگی روی توزیع های مختلف است فینکلستین (۲۰۰۸) نرخ خطر آمیخته از دو خانواده توزیع رامعرفی کرد و ما آن را برای نرخ خطر معکوس تعمیم می دهیم.

فرض کنید دو زیر جامعه با توابع توزیع F_1 ، F_2 و توابع چگالی f_1،f_2 داریم نرخ خطر معکوس آمیخته به فرم:

$$r(t)=\frac{pf_1(t)+(1-p)f_2(t)}{pF_1(t)+(1-p)F_2(t)},$$

تعریف می شود اگر معادله فوق رابرای k زیر جامعه تعمیم دهیم داریم:

$$r(t)=\frac{\sum_{i=1}^k p_if_i(t)}{\sum_{i=1}^k p_iF_i(t)},\ \sum_{i=1}^k p_i=1,\ k\geq 1, 0\leq p_i\leq 1.$$

آمیختگی از دو خانواده توزیع، خانواده توزیع جدیدی به وجود می آورد و ممکن است رفتار نرخ خطر معکوس روی این خانواده توزیع جدید متفاوت باشد.

مثال ۱ : فرض کنید خانواده توزیع های نمایی و وایبل را با توابع چگالی $f_{1}(t) = e^{-t}$ و $f_{2}(t) = 2te^{-t^{2}}$ و نرخ خطر معکوس کاهشی داشته باشیم با در نظر گرفتن $p = \frac{1}{2}$ به سادگی می توان نشان داد نرخ خطر معکوس آمیخته از این دو خانواده توزیع کاهشی است.

قضیه۱ : اگر T متغیر تصادفی منفی باشد. نرخ خطر معکوس آمیخته از k زیر جامعه با نرخ خطر معکوس افزایشی $(IRHR)$ همواره افزایشی $(IRHR)$ است.

اثبات: اگر $w_i(t) = \frac{p_i F_i(t)}{F(t)}$ در نظر بگیریم آن گاه

$r(t) = \sum_{i=1}^{k} w_i(t) r_i(t)$

$$r'(t) = \sum_{i=1}^{k} w_i'(t) r_i(t) + \sum_{i=1}^{k} r_i'(t) w_i(t),$$

از طرفی داریم:

$$\begin{aligned} w_i'(t) &= w_i(t)\left(r_i(t) - r(t)\right) \\ &= w_i(t)\left(\sum_{j=1}^{k} w_j(t)(r_j(t) - r_i(t))\right), \end{aligned}$$

با جایگذاری می توان نوشت

$$r'(t) = \sum_{i=1}^{k} w_i(t) r_i'(t) + \sum_{i<j} w_i(t) w_j(t)\left(r_i(t) - r_j(t)\right)^{2},$$

چون $0 \leq w_i(t) \leq 1$ و $r_i(t)$ افزایشی است رابطه فوق در کل افزایشی است.

۹ خلاصه و نتیجه گیری

امروزه نرخ خطر آمیخته اهمیت ویژه ای در شاخه مهندسی روی فرآیند های آبکاری [1] و سیستم هایی با حداقل تعمیرات [2] دارد. در این مقاله نرخ خطر معکوس و توزیع های آمیخته را در حالت یک متغیره معرفی کردیم و رفتار نرخ خطر معکوس آمیخته از دو خانواده توزیع را مورد بررسی قرار دادیم. برای مطالعه بیشتر در آینده می توانیم این مفاهیم را برای حالت دو متغیره و توزیع های وزنی بسط دهیم.

مراجع

[1] Aalen, O. O. (1998). Heterogrnrity in survival analysis. Statistical in Medicine 7, 1121-1137.

[2] Aalen, O. O. (1992). Modeling heterogeneity in survival analysis by the compound poisson distribution. The Annals of Applied Probability 2, 951-972.

[1] Burn-in
[2] Minimal Repair

[3] Barlow, R. E., Marshall, A. W. and Proschan, F. (1963). Properties of probability distributions with monotonic hazard rate. Annals of Statistics, 34(3), , PP 341-350.

[4] Block, H., Savits, T. and Singh, H. (1998). The reversed hazard rate function. Probability in the Engineering and Information Sciences 12:69-70.

[5] Finkelstein, M. (2008). Failure Rate Modeling for Reliability and Risk. Mathematical Statistics. Springer Series in Reliability Engineering.

[6] Gurland, J. and Sethuraman, J. (1995). How pooling filure data may reverse increasing failure rates. Journal of the American Statistical Association.

[7] Kilson, J. and Sumita, U. (1982). Uniform stochastic ordering and related inequalities. Canadian Journal of Statistics 10:181-198.

[8] Weiss. (1990). The biodemography of variation in human frailty. Demography 27, 106-185.

[9] Xiaohu Li, Gaofeng Da,. and Peng Zhao. (2010). On reversed hazard rate in general mixture models. Statistics and Probability Letters 80:654-661.