بررسی ساختار وابستگی مدل های بی ثبات

معصومه بزرگ - محمد امینی

گروه آمار دانشگاه فردوسی مشهد

چکیده: مدل های بی ثبات در تجزیه و تحلیل بقا برای تشخیص ناهمگنی های (پنهان) بین افراد کاربرد فراوانی دارد. علاوه بر این وابستگی بین متغیرهایی تصادفی ، که در تحلیل بقا، نظریه قابلیت اعتماد و زمینه های دیگر آمار و احتمال نقش مهمی دارد مورد توجه محققان می باشد. در این مقاله به بررسی ساختار وابستگی بین یک متغیر پایه با متغیر بی ثباتی، و همچنین تاثیر ترتیب تصادفی دو متغیر پایه به روی متغیرهای بی ثباتی متناظر می پردازیم.

واژه‌های کلیدی: مدل بی ثبات، ترتیب تصادفی، وابستگی، مقایسه تصادفی

۱ مقدمه

مدل های بی ثبات اولین بار در سال ۱۹۷۹ توسط واپل [۱] ، برای مطالعه برخی ناهمگنی که بوسیله متغیرهای کمکی قابل اندازه گیری نمی باشد معرفی شد. کاربرد این مدل ها در تحلیل بقا برای ناهمگنی مشاهده نشده (پنهان) بین افراد، با فرض اینکه نرخ خطر افراد متاثر از یک کمیت مشخص به نام بی ثباتی و یک نرخ خطر پایه در نظر گرفته شود، مورد توجه بسیاری از محققان بوده است. در سال های اخیر نیز مطالعات بسیاری روی این مدل ها انجام شده است. به عنوان مثال گوپتا و کرمانی[۲] (۲۰۰۶) مقایسه های تصادفی در مدل های بی ثبات را از طریق تاثیر مقایسه دو متغیر تصادفی پایه بر روی متغیرهای بی ثباتی متناظر مطالعه کرده و نتایج مفیدی بدست آوردند. ژو و لی[۳] (۲۰۰۸) نیز جنبه های وابستگی بین متغیر پایه و متغیر بی ثباتی در این مدل ها را بررسی کردند. علاوه بر این مولفان دیگری روی مدل های بی ثبات تحقیق کردند. به عنوان مثال آلن[۴] (۱۹۹۲، ۱۹۸۸، ۱۹۸۷)، هوگارد[۵] (۲۰۰۰، ۱۹۹۵، ۱۹۹۱، ۱۹۸۴)،

۱ Vaupel
۲ Gupta and kirmani
۳ Maochao Xu, Xiaohu Li
۴ Aalen
۵ Hougaard

بارلو و پروسچان[۶] (۱۹۸۱)، نلسن[۷] (۱۹۹۹) و مولر و استون[۸] (۲۰۰۲).
در این مقاله ، هدف مطالعه جنبه های وابستگی متغیر پایه با متغیر بی ثباتی و بررسی تاثیر ترتیب تصادفی دو متغیر پایه بر روی متغیرهای بی ثباتی متناظر می باشد.
فرض کنید (X, V) یک بردار تصادفی از متغیرهای تصادفی نامنفی باشد ، به قسمی که توزیع شرطی $(X|V = v)$ مطلقا پیوسته با نرخ خطر (مخاطره)

$$\lambda(t|v) = v\lambda_\circ(t) \qquad \forall \ t \geq \circ \tag{۱}$$

باشد. که در آن $\lambda_\circ(t)$ تابع نرخ خطر متغیر تصادفی پایه X و مستقل از V است. در واقع اگر X معرف عمر یک مولفه باشد، آن گاه $\lambda(t|v)$ معرف نرخ مخاطره در لحظه t برای مولفه ای با بی ثباتی V خواهد بود. مدل (۲) بیان می کند نرخ مخاطره یک مولفه (فرد) برابر است با حاصلضرب کمیت معلوم V و نرخ مخاطره پایه $\lambda_\circ(t)$ که تاثیر سن را تشریح می کند. متغیر تصادفی V را به عنوان متغیر تصادفی بی ثباتی در جامعه معرفی کردند .
اگر متغیرهای تصادفی نامنفی X و Y را به ترتیب با توزیع های F و G در نظر بگیریم، گوپتا و کرمانی (۲۰۰۶) نشان دادند اگر $\bar{G}(t) = P[Y > t]$ توزیع پایه با نرخ مخاطره پایه $\lambda_\circ(t)$ باشد، آن گاه تابع مخاطره جامعه را با $\lambda(t) = -\frac{d}{dt}\ln \bar{F}(t)$ نشان می دهیم. و در نتیجه

$$\bar{F}(t) = E[\bar{G}^V(t)] \qquad \forall \ t \geq \circ \tag{۲}$$

روابط (۲) و (۵) نشان می دهند که متغیر بی ثباتی V یک ارتباط منطقی بین X و Y برقرار می کند.علاوه بر این ، اگر $\theta(t) = E[V|X > t]$ آن گاه گوپتا و گوپتا (۱۹۹۶) یک رابطه مفید بین $\lambda(t)$ تابع نرخ مخاطره جامعه و $\lambda_\circ(t)$ تابع نرخ مخاطره پایه به صورت زیر بدست آوردند:

$$\lambda(t) = \lambda_\circ(t)\theta(t) \qquad \forall \ t \geq \circ \tag{۳}$$

با توجه به روابط (۲) و (۵) و (۵) ، بررسی ساختار وابستگی بین X و Y با حضور متغیر بی ثباتی V ، قابل توجیه می باشد.
نمادها و تعریف های زیر در ادامه مورد نیاز است :
فرض کنید X و Y متغیرهای تصادفی پیوسته به ترتیب با توابع توزیع F و G و توابع چگالی f و g باشند و برای هر $t \geq \circ$ ، $X_t = (X - t|X > t)$ طول عمر باقیمانده

[۶] Barlow and Proschan
[۷] Nelsen
[۸] Muller and Stoyn

در لحظه t باشد. ۱) Y در ترتیب درستنمایی از X کوچکتر است اگر $\frac{f(x)}{g(x)}$ نسبت به $x \geq 0$ صعودی باشد. (که به صورت $Y \geq_{lr} X$ یا $F \geq_{lr} G$ نشان می دهیم.)

۲) Y در ترتیب نرخ خطر از X کوچکتر است اگر $\frac{\bar{F}(x)}{\bar{G}(x)}$ نسبت به $x \geq 0$ صعودی باشد. (که به صورت $Y \geq_{hr} X$ یا $F \geq_{hr} G$ نشان می دهیم)

۳) Y در ترتیب نرخ خطر معکوس از X کوچکتر است اگر $\frac{F(x)}{G(x)}$ نسبت به $x \geq 0$ صعودی باشد. (که به صورت $Y \geq_{rh} X$ یا $F \geq_{rh} G$ نشان می دهیم.)

۴) Y در ترتیب تصادفی معمولی از X کوچکتر است اگر برای هر $x \geq 0$ داشته باشیم $\bar{F}(x) \geq \bar{G}(x)$. (که به صورت $Y \geq_{st} X$ یا $F \geq_{st} G$ نشان می دهیم.)

۵) Y در ترتیب مقعر صعودی از X کوچکتر است اگر برای هر $t \geq 0$ داشته باشیم: $\int_0^t \bar{F}(x)\ dx \geq \int_0^t \bar{G}(x)\ dx$ (که به صورت $Y \geq_{icv} X$ یا $F \geq_{icv} G$ نشان می دهیم.)

۶) X دارای نسبت درستنمایی نزولی است (DLR)، اگر X_t نسبت به $t \geq 0$ نزولی باشد، به معنی ترتیب نسبت درستنمایی.

۷) X دارای نرخ شکست نزولی است (DFR)، اگر X_t نسبت به $t \geq 0$ نزولی باشد، به معنی ترتیب تصادفی معمولی.

۸) X دارای متوسط نرخ شکست نزولی است (DFRA)، اگر $\frac{-\ln \bar{F}(t)}{t}$ نسبت به $t \geq 0$ نزولی باشد.

۹) X را NWU (NBU) گوییم اگر برای هر $t \geq 0$ داشته باشیم: $X \leq_{st} (\geq_{st}) X_t$

۱۰) X را NWU(2) (NBU(2)) گوییم اگر برای هر $t \geq 0$ داشته باشیم: $X \leq_{icv} (\geq_{icv}) X_t$

۱۱) متغیر تصادفی Y به طورتصادفی در دم راست برحسب X نزولی است اگر برای همه t ها $P(Y > t | X > x)$ نسبت به x نزولی باشد.

۱۲) اگر X و Y دارای چگالی توام $f(x,y)$ باشند آن گاه X و Y بطور منفی وابسته نسبت درستنمایی هستند ، NLRD(X,Y) ، اگر برای هر $y_1 < y_2$: $\frac{f(x,y_2)}{f(x,y_1)}$ تابعی نزولی نسبت به x باشد.

مفاهیم متنوعی از وابستگی ها معرفی شده است که برای مطالعه بیشتر در این زمینه می توان به مراجعه کرد. وابستگی مورد مطالعه در این مقاله در تعریف ۱۲ بیان شده است ، قوی ترین نوع وابستگی بین دو متغیر تصادفی می باشد.

۲ نتایج اصلی

در بخش ۱.۲ نشان می دهیم که X و V دارای وابستگی نسبت درستنمایی منفی هستند همچنین ویژگی های DFR و DFRA و DLR در دو متغیر تصادفی X و Y با حضور متغیر بی ثباتی V بررسی می شود.علاوه بر این ، تاثیر مقایسه متغیر تصادفی پایه بر متغیر V در بخش ۲.۲ مطالعه می شود.

۱.۲ ساختار وابستگی

قضیه زیر نشان می دهد که متغیرهای تصادفی X و V دارای قوی ترین نوع وابستگی منفی هستند.

قضیه ۱.۲ : X و V به طور منفی وابسته نسبت درستنمایی هستند.

اثبات : اگر $h(v),\ \ v \geq 0$ تابع چگالی متغیر بی ثباتی V باشد، آن گاه گوپتا و کرمانی (۲۰۰۶) نشان دادند که

$$f(x,v) = v\lambda_0(t)(\bar{G}(x))^v h(v), \quad x,v \geq 0$$

در نتیجه برای هر $v_2 > v_1$ و $x \geq 0$ داریم

$$\begin{aligned}\frac{f(x,v_2)}{f(x,v_1)} &= \frac{v_2\lambda_0(t)(\bar{G}(x))^{v_2}h(v_2)}{v_1\lambda_0(t)(\bar{G}(x))^{v_1}h(v_1)} \\ &= \frac{v_2 h(v_2)}{v_1 h(v_1)}(\bar{G}(x))^{v_2-v_1}\end{aligned}$$

دیده می شود که $\frac{f(x,v_2)}{f(x,v_1)}$ نسبت به x نزولی است. بنابراین NLRD(X,V) . □

نتیجه ۱.۲ : برای هر $v \geq 0$ و $t \geq s \geq 0$ و $u \geq v \geq 0$ داریم

الف) $P(V > v|X > t) \leq P(V > v|X > s)$

ب) $P(X > t|V > u) \leq P(X > t|V > v)$

اگر متغیر تصادفی X معرف طول عمر و V به عنوان یک شاخص ریسک در نظر گرفته شود، آن گاه نامساوی (ب) بیان می کند که با افزایش ریسک ، عمر کوتاه تر می شود. اگر متغیر تصادفی V دارای ویژگی های NBU(2) و NWU(2) باشد ،آن گاه ژو و لی (۲۰۰۸) کران های بالا و پایین برای تابع بقا زمانی که ریسک بالا باشد، ارائه دادند.

قضیه ۲.۲: فرض کنید که $E(V) \leq 1$. در اینصورت

الف) اگر V دارای ویژگی NBU(2) باشد آن گاه برای هر $t \geq 0$ و $\xi \geq 0$ داریم:

$$P(X > t | V > \xi) \geq \bar{G}^{\xi+1}(t) \tag{۴}$$

ب) اگر V دارای ویژگی NWU(2) باشد آن گاه برای هر $t \geq 0$ و $\xi \geq 0$ داریم:

$$P(X > t | V > \xi) \geq \bar{G}^{\xi}(t) \tag{۵}$$

علاوه بر این آنها ثابت کردند :

گزاره ۲.۲: اگر Y از کلاس DFR (DFRA) باشد، آن گاه X هم از کلاس DFR (DFRA) است.

قضیه ۳.۲: اگر Y دارای خاصیت DFR باشد و $\lambda_0(x)$ ، logconvex باشد، آن گاه X دارای خاصیت DLR است.

۲.۲ مقایسه های تصادفی

برای انتخاب توزیع احتمال متغیرهایی تصادفی بی ثباتی در زمینه های کاربردی هیچ پایه ی نظری ثابتی وجود ندارد. بنابراین بررسی اینکه توزیع احتمال متغیر تصادفی اصلی چه واکنشی نسبت به تغییرات توزیع احتمال بی ثباتی نشان می دهد از اهمیت خاصی برخوردار است. در قضیه های زیر بیان می شود که ترتیب نسبت درستنمایی، ترتیب نرخ خطر، و ترتیب نرخ خطر معکوس بین دو متغیر بی ثباتی، ترتیب نسبت درستنمایی و ترتیب نرخ خطر معکوس را برقرار می سازد.

قضیه ۴.۲: فرض کنید V_1 و V_2 دو متغیر تصادفی بی ثباتی باشند به قسمی که $V_1 \leq_{lr} V_2$. همچنین فرض کنید که $\bar{F}_i(t) = E((\bar{G}(t))^{V_i}), i = 1, 2$ تابع بقای متغیر تصادفی بی ثبات جامعه اصلی با متغیر تصادفی بی ثباتی V_i و تابع بقای پایه $\bar{G}(t)$ باشد.آن گاه $\frac{\bar{F}_2(t)}{\bar{F}(t)}$ نسبت به $t \geq 0$ نزولی ست.

قضیه ۵.۲: اگر $V_1 \leq_{lr} V_2$ آن گاه $X_1 \geq_{lr} X_2$.

قضیه ۶.۲: اگر $V_1 \leq_{rh} V_2$ آن گاه $X_1 \geq_{hr} X_2$.

مثال ۱.۲: فرض کنید $\{N(t), t \geq 0\}$ یک فرآیند پواسون ناهمگن با تابع میانگین Λ (یعنی $\Lambda(t) \equiv E[N(t)], t \geq 0$) ، و همچنین فرض کنید که $T_1, T_2, \ldots$ زمان های آغاز موفقیت باشد. تابع بقای T_n به صورت، $P\{T_n > t\} = \Sigma_{i=0}^{n-1} \frac{(\Lambda(t))^i}{i!} e^{-\Lambda(t)}, t \geq 0$ و تابع چگالی T_n به صورت، $f_n(t) = \lambda(t) \frac{(\Lambda(t))^{n-1}}{(n-1)!} e^{-\Lambda(t)}, t \geq 0$ داده شده است. که در آن $\lambda(t) \equiv \frac{d}{dt}\Lambda(t), n = 1, 2, \ldots$.

به سادگی دیده می شود که $\frac{f_{n+1}(t)}{f_n(t)}$ برای $n = 1, 2, \ldots$ نسبت به t صعودی است. بنابراین

$$T_n \leq_{lr} T_{n+1}, \quad n = 1, 2, \ldots.$$

و از این نتیجه می گیریم که اگر $V_1 = T_n$ و $V_2 = T_{n+1}$ در نظر بگیریم، بنابراین شرایط قضیه ۵.۲ برقرار می شود.

مثال ۲.۲: برای حالت خاصی از مثال فوق فرض کنید که V_1 و V_2 دارای تابع چگالی های زیر باشند :

$$h_1(v) = \frac{1}{\Gamma(\alpha_1)} \beta_1^{\alpha_1} V^{\alpha_1 - 1} e^{-\beta_1 V}$$
$$h_2(v) = \frac{1}{\Gamma(\alpha_2)} \beta_2^{\alpha_2} V^{\alpha_2 - 1} e^{-\beta_2 V}$$

داریم

$$\frac{h_2(v)}{h_1(v)} = \frac{\frac{1}{\Gamma(\alpha_2)} \beta_2^{\alpha_2} V^{\alpha_2 - 1} e^{-\beta_2 V}}{\frac{1}{\Gamma(\alpha_1)} \beta_1^{\alpha_1} V^{\alpha_1 - 1} e^{-\beta_1 V}}$$
$$= \frac{\Gamma(\alpha_1)\beta_2^{\alpha_2}}{\Gamma(\alpha_2)\beta_1^{\alpha_1}} \; V^{\alpha_2 - \alpha_1} \; e^{-V(\beta_2 - \beta_1)}$$

طبق تعریف ترتیب درستنمایی شرط لازم برای اینکه $V_1 \leq_{lr} V_2$ برقرار باشد باید $\frac{h_2(v)}{h_1(v)}$ نسبت به v صعودی باشد، و این در حالات زیر برقرار می شود :

الف) اگر $\beta_2 > \beta_1$ و $\alpha_2 < \alpha_1$.

ب) اگر $\beta_2 < \beta_1$ و $\alpha_2 > \alpha_1$.

بنابراین اگر یکی از حالات فوق رخ دهد آن گاه $V_1 \leq_{lr} V_2$ برقرار می شود و نتیجه قضیه ۵.۲ از آن بدست می آید.

مراجع

Aalen, O.O. (1988), Heterogeneity in survival analysis, *Statist. Medicine*, **7**, 1121-1137.

Aalen, O.O. (1992), Modelling heterogeneity in survival analysis by the compound Poisson distributhin, *Ann. Appl. Prpbab.*, **2**, 951-972.

Barlow, R.E., Proschan, F. (1981), Statistical Theory of Reliability and Life Testing, To Begin With, Silver Spring, MD.

Gupta, P.L., Gupta, R.C. (1996), Ageing classes of Weibull mixtures, *Probab. Eng. Inform. Sci.*, **10**, 591-600.

Gupta, R.C., Kirmani, S.N.U.A (2006), Stochastic comparisons in frailty models , *J.statist. Plann. Inference*, **136**, 3647-3658.

Hougaard, p., (1984), Life table methods for heterogeneous population: distribution describing the heterogeneity, *Biometrika*, **71**, 75-83.

Hougaard, p., (1991), Modelling heterogeneity in survival analysis, *J. Appl. Prpbab.*, **28**, 695-701.

Hougaard, p., (1995), Frailty models for survival analysis, *Lifetime Data Anal.*, **1**, 255-273.

Muller, A., Stoyan, D. (2002), *Comparison Methods for Stochastic Models and Risks*, Wiley, NewYork.

Nelsen, R.B. (1999), *An Introduction to Copulas*, Lecture Notes in Statistics, Springer, NewYork.

Shaked, M., Shanthikumar, J.G. (1994), *Stochastic Orders and their Applications*, Academic Press, San Diego, CA.

Vaupel, J.W., Manton, K.G., Stallard, E. (1979), The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography*, **16**, 439-454.

Xu, M., Li, X. (2008), Negative dependence in frailty models,*J.statist. Plann. Inference* , **138**, 1433-1441.