متغیرهای تصادفی مرتبط منفی و برخی کاربردهای آنها

محمد امینی — مهد یس آزاد بخش [1]
دانشگاه سیستان و بلوچستان
$amini@math.usb.ac.ir$
$ma\_azadbakhsh@yahoo.com$

چکیده :
در این مقاله ، متغیر های تصادفی مرتبط منفی را معرفی می نماییم . ویژگی ها و برخی از کاربرد های آن را با ارائه مثالهای متنوع بیان می کنیم . علاوه بر این ، ارتباط بین متغیر های تصادفی NA و مفاهیم دیگر وابستگی ، به ویژه وابستگی در دنباله ، بطور شرطی نزولی در دنباله و ویژگی منظم معکوس نیز مورد توجه قرار گرفته اند .
در ادامه نیز ، توزیع های جایگشتی و برخی توزیع های چند متغیره کلاسیک که دارای ویژگی NA هستند را معرفی می نماییم .
واژه های کلیدی : مرتبط منفی ، وابسته منفی ، توزیع های چند متغیره ، وابستگی در دنباله ، منظم معکوس ، به طور شرطی نزولی در دنباله .

۱ — مقدمه :

مفهوم اولیه ارتباط(در حالت مثبت) که دستور العمل قویتری از وابستگی است ، ابتدا در سال (۱۹٦٦)توسط لهمن معرفی شد و سپس در سال (۱۹٦۷) ازاری ، پروشن و والکپ ، این مفهوم از وابستگی رامطالعه و برخی از کاربرد های آن را به ویژه، در نظریه اعتماد بدست آوردند. بعداً در سال (۱۹۸۳) ، کومار و همکاران ، ارتباط منفی را معرفی وبرخی از ویژگیهای متغیر های مرتبط منفی را بررسی نمودند. در سال های اخیر نیز مطالعات زیادی روی متغیر های تصادفی PA و NA انجام شده است و بویژه زمینه های کاربردی متغیر های تصادفی PA و NA در تحلیل رگرسیون و آزمون های ناپارامتری مورد توجه بسیاری از آماردانان قرار گرفته است . در این مقاله به مطالعه و بررسی بیشتر متغیر های تصادفی NA ، ویژگی ها و رابطه این مفهوم با مفاهیم دیگر وابستگی می پردازیم ، با ارائه مثال های متنوع ، برخی کاربردهای این متغیر های تصادفی را بیان نموده ، برخی توزیع های چند متغیره کلاسیک را که دارای ویژگی NA هستند را معرفی می کنیم.

تعریف ۱ : متغیرهای تصادفی $X_1, X_2, ..., X_n$ را به طورمنفی مرتبط (NA) گویند، اگربرای هر زوج از زیرمجموعه های مجزای A_1 و A_2 از مجموعه $\{1, 2, ..., n\}$وهر دو تابع غیر نزولی f وg نامساوی زیربرقرار باشد:

$$Cov(f(X_j; j \in A_1), g(X_k; k \in A_2)) \leq 0 \qquad (1)$$

که در آن ، $A_1 \cap A_2 = \phi$. بدیهی است که هر گاه f وg ،هر دو نزولی باشند نیز نامساوی فوق برقرار است .

[1] گروه ریاضی دانشگاه سیستان و بلوچستان

تعریف ۲: متغیر های تصادفی X و Y را به طور منفی وابسته مربعی (NQD) نامند ، اگر:

$$P(X \leq x, Y \leq y) \leq P(X \leq x).P(Y \leq y) \qquad \forall x, y \in R$$

تعریف ۳: متغیرهای تصادفی $X_۱, X_۲, ..., X_n$ را از بالا وابسته منفی (UND) نامند ، اگر:

$$P(\bigcap_{j=۱}^{n}(X_j > x_j)) \leq \prod_{j=۱}^{n} P(X_j > x_j) \qquad \forall x_۱, x_۲, ..., x_n \in R$$

و از پایین وابسته منفی (LND) نامند ، اگر:

$$P(\bigcap_{j=۱}^{n}(X_j \leq x_j)) \leq \prod_{j=۱}^{n} P(X_j \leq x_j) \qquad \forall x_۱, x_۲, ..., x_n \in R$$

متغیرهای تصادفی $X_۱, X_۲, ..., X_n$ را وابسته منفی (ND) نامند ، اگر LND و UND باشند.

قضیه ۱ ([۴]) : فرض کنید (X, Y) یک زوج از متغیر های تصادفی و Z یک متغیر یا یک بردار تصادفی باشد ، آنگاه :

$$Cov(X, Y) = E_Z[Cov(X, Y)|Z)] + Cov[E(X|Z), E(Y|Z)]$$

قضیه ۲ ([۴]) : فرض کنید $X_۱, X_۲, ..., X_n$ متغیرهای تصادفی مستقل و $E[f(X_i, i \in A)|\sum_{i \in A} X_i]$ به ازای هر تابع صعودی f و زیر مجموعه $A \subset \{۱, ..., n\}$ ، بر حسب $\sum_{i \in A} X_i$ صعودی باشد، آنگاه: $X_۱, X_۲, ..., X_n$ به شرط $\sum_{i \in A} X_i$ ، NA هستند.$a.e.$

اثبات : فرض کنید $A_۱, A_۲$ افرازی دلخواه از $\{۱, ..., n\}$ باشند و $S_۱ = \sum_{i \in A_۱} X_i$ و $S_۲ = \sum_{i \in A_۲} X_i$ ، $S = S_۱ + S_۲$ و $f_۱$ و $f_۲$ توابعی صعودی ، به ترتیب روی $A_۱$ و $A_۲$ باشند ، داریم :

$$Cov(f_۱(X_i, i \in A_۱), f_۲(X_j, j \in A_۲)|S) = Cov(E(f_۱|S_۱, S_۲), E(f_۲|S_۱, S_۲)|S)$$

$$+E(Cov((f_۱, f_۲)|S_۱, S_۲)|S)$$

$$E(Cov(f_۱, f_۲|S_۱, S_۲)|S) = E(E(f_۱.f_۲|S_۱, S_۲)|S) - E[E(f_۱|S_۱, S_۲)|S).E(f_۲|S_۱, S_۲)]$$

$$E(f_۱.f_۲|S_۱, S_۲) - E(f_۱|S_۱, S_۲).E(f_۲|S_۱, S_۲) = E(f_۱|S_۱).E(f_۲|S_۲) - E(f_۱|S_۱).E(f_۲|S_۲) = ۰$$

تساوی خط آخر از استقلال و مجزا بودن $f_۱$ و $f_۲$ نتیجه می شود. پس :

$$Cov(f_۱(X_i, i \in A_۱), f_۲(X_j, j \in A_۲) = Cov(E(f_۱|S_۱, S_۲), E(f_۲|S_۱, S_۲)|S)$$

هرگاه $S = s$ ، $E(f_۱|S_۱, S_۲)$ و $E(f_۲|S_۱, S_۲)$ دو تابع ناهماهنگ برحسب $S_۱$ هستند ، لذا : $Cov(E(f_۱|S_۱, S_۲), E(f_۲|S_۱, S_۲)|S) \leq ۰$ و در نتیجه $X_۱, X_۲, ...X_n|S = s$ ،NA هستند.

تعریف ۴ ([۳]) : تابع چگالی احتمال $r(x)$ ، تابع فراوانی پولیا از مرتبه دو $(PF_۲)$ نامیده می شود ،هر گاه :

$$\begin{vmatrix} r(x_۱ - z_۱) & r(x_۱ - z_۲) \\ r(x_۲ - z_۱) & r(x_۱ - z_۲) \end{vmatrix} \geq ۰$$

که در آن: $x_۱ \leq x_۲, z_۱ \leq z_۲$

قضیه ۳ ([۳]) : فرض کنید $X_n, ..., X_۲, X_۱$ متغیرهای تصادفی مستقل با چگالیهای $f_n, ..., f_۱$ باشند که $PF_۲$ نیز هستند و $S = \sum_{i=۱}^{n} X_i$ و $\phi(x_۱, ...x_n)$ در هر مولفه صعودی باشد، آنگاه با احتمال یک $E[\phi(X_۱, ..., X_n)|S = s]$ بر حسب s صعودی است.

قضیه ۴: فرض کنید $X_n, ..., X_۲, X_۱$ متغیرهای تصادفی مستقل با چگالیهای $PF_۲$ باشند ، آنگاه با احتمال یک $X_۱, ..., X_n)|S = \sum_{i=۱}^{n} X_i$ متغیر های تصادفی NA هستند.

لم ۱ (هافدینگ(۱۹۴۰)) :اگر (X, Y) بردار تصادفی دو متغیره باشد و $E(X^۲) < \infty$ و $E(Y^۲) < \infty$ ، انگاه:

$$Cov(X, Y) = \int_{R^۲} [P(X \leq x, Y \leq y) - P(X \leq x).P(Y \leq y)]dx.dy$$

۲– ویژگی های متغیر های تصادفی NA :

ویژگی ۱– برای هر زوج تصادفی (X, Y) ، $NQD \Leftrightarrow NA$

ویژگی ۲–اگر $A_m, ..., A_۲, A_۱$ زیر مجموعه های مجزا از $\{۱, ۲, ..., n\}$ و $f_m, ..., f_۱$ توابع مثبت صعودی باشند ،انگاه اگر $X_n, ..., X_۲, X_۱$ متغیر های تصادفی NA باشند ،

$$E\prod_{i=۱}^{m} f_i(X_j, j \in A_i) \leq \prod_{i=۱}^{m} Ef_i(X_j, j \in A_i)$$

ویژگی ۳– اگر $X_n, ..., X_۲, X_۱$ متغیرهای تصادفی NA باشند ، آنگاه ، ND نیز هستند .$(NA \Rightarrow ND)$

ویژگی ۴– هر زیر مجموعه از دو یا بیشتر از دو متغیر تصادفی NA ، NA است.

ویژگی ۵– یک مجموعه از متغیر های تصادفی مستقل ،NA هستند.

ویژگی ٦– توابع صعودی تعریف شده روی زیر مجموعه های مجزایی از یک مجموعه از متغیر های تصادفی NA ،NA هستند.

ویژگی ۷– اجتماع مجموعه های مستقل بامتغیر های تصادفی NA ،NA است.

ویژگی ۸– هر گاه برای تمام توابع غیر نزولی دو مقداری γ, δ ،

$Cov(\gamma(X_j; j \in A_۱), \delta(X_i; i \in A_۲)) \leq ۰$ ، که در آن $A_۱, A_۲$ دو مجموعهٔ مجزا از $\{۱, ۲, ..., n\}$ هستند، آنگاه $X_۱, ..., X_n$، NA هستند.

اثبات : فرض کنید f, g ،دو تابع صعودی باشند و $A_۱ \cap A_۲ = \phi$ ، اگر :

$$X_f(x) = \begin{cases} ۱ & f(X_j; j \in A_۱) > x \\ ۰ & o.w. \end{cases}$$

$$X_g(y) = \begin{cases} ۱ & g(X_i; i \in A_۲) > y \\ ۰ & o.w. \end{cases}$$

آنگاه بنابه لم (۱) ، داریم :

$$Cov(f(X_j; j \in A_۱), g(X_i; i \in A_۲) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} Cov(X_f(x), X_g(y))dxdy$$

که این حکم را ثابت می کند .

ویژگی ۹– اگر $X_۱, X_۲, ..., X_n$ متغیر های تصادفی دو تایی NA باشند، آنگاه متغیرهای تصادفی $۱ - X_۱, ..., ۱ - X_n$ نیز NA هستند.

تبصره ۱ :ویژگی های شش و هفت به طور قابل ملاحظه ای کاربرد NA را گسترش می دهند.برای مثال برای بررسی خاصیت NA در توزیع هایی که تلفیقی از توزیع های ساده هستند ،تنها لازم است این ویژگی را در توزیع های ساده بررسی کنیم .

نتیجه ۱ : ویژگی ND ، NA را نتیجه نمی دهد.

مثال ۱ : فرض کنید $X_۱, X_۲, X_۳, X_۴$ متغیرهای تصادفی دو مقداری با تابع احتمال توام ، مطابق جدول زیر باشند :

				$(X_۱, X_۲)$		
		$(۰, ۰)$	$(۰, ۱)$	$(۱, ۰)$	$(۱, ۱)$	$marginal$
	$(۰, ۰)$	.۰۵۷۷	.۰٦۲۳	.۰٦۲۳	.۰۵۷۷	.۲۴
	$(۰, ۱)$	.۰٦۲۳	.۰٦۷۷	.۰٦۷۷	.۰٦۲۳	.۲٦
$(X_۳, X_۴)$	$(۱, ۰)$	.۰٦۲۳	.۰٦۷۷	.۰٦۷۷	.۰٦۲۳	.۲٦
	$(۱, ۱)$	.۰۵۷۷	.۰٦۲۳	.۰٦۲۳	.۰۵۷۷	.۲۴
	$marginal$	.۲۴	.۲٦	.۲٦	.۲۴	

متغیرهای تصادفی $X_۱, X_۲, X_۳$ و $X_۴$ ، ND هستند([۴]) ولی NA، نیستند،زیرا اگر $A_۱ = \{۱, ۲\}$ و $A_۲ = \{۳, ۴\}$ و

$$f(X_j; j \in A_۱) = \begin{cases} ۱ & X_j > ۰; j \in A_۱ \\ ۰ & o.w. \end{cases}$$

$$g(X_i; i \in A_2) = \begin{cases} 1 & X_i > 0; i \in A_2 \\ 0 & o.w. \end{cases}$$

آنگاه :

$$COV(f,g) = Efg - Ef.Eg = P[X_i = 1; i = 1, ..., 4] - P[X_1 = X_2 = 1].P[X_3 = X_4 = 1]$$

$$= 0.0577 - 0.24 * 0.24 > 0$$

در نتیجه $X = (X_1, X_2, X_3, X_4)$ ،NA نیست.

تعریف ۵ : فرض کنید $X = (x_1, ..., x_n)$ مجموعه ای از n عدد حقیقی باشد. یک توزیع جایگشتی ، توزیع توام بردار $X = (x_1, ..., x_n)$ است که همه $n!$ مقدار جایگشت های $\underline{X}$ را با احتمال مساوی $1/n!$ می گیرد.$(n \geq 1)$

قضیه ۵ ([۴]) : یک توزیع جایگشتی ،NA است.

قضیه ٦ : اگر $X_1, ..., X_n$ ، n متغیر تصادفی مستقل با توزیع پیوسته باشند، آنگاه توزیع شرطی توام $(X_1, X_2, ..., X_n)$ به شرط اماره ترتیبی $(X_{n_1} = s_1, ..., X_{n_r} = s_r)$، برای هر $1 < n_1 < n_2 < ... < n_r < n$ ،NA است.

—در ادامه می خواهیم برخی دیگر از مفاهیم وابستگی منفی راکه توسط کارلین و رینوت (۱۹۸۰) تعمیم داده شد و توسط بلاک(۱۹۸۲) اصلاح گردید، بیان و مقایسه کنیم .

تعریف ٦ : الف)فرض کنید μ یک اندازه احتمال روی R^2 باشد. گوییم μ منظم معکوس از مرتبه ۲(RR_2) است ، اگر:

$$\mu(I_1, I_2).\mu(I'_1, I'_2) \leq \mu(I_1, I'_2).\mu(I'_1, I'_2) \qquad (2)$$

برای همه فاصله های $I_1 < I'_1$ و $I_2 < I'_2$ در R^1 .

که به طور کلی μ ، اندازه احتمال روی مجموعه بورل در R^n است و برای فاصله های $I_1, ..., I_n$ در R^1 : $\mu(I_1, ..., I_n) = \mu(I_1 * ... * I_n)$. علاوه بر این برای هر دو فاصله I و J از R^1 ،$I < J$ ، یعنی برای هر $y \in J$ و $x \in I$: $x < y$

ب) فرض کنید μ یک اندازه احتمال روی R^n،$(k \geq 2)$ باشد. گوییم μ در I_i و I_j $(\forall 1 \leq i < j \leq n)$ ،RR_2 است ، هنگامی که رابطه (۲) برای این زوج فاصله وقتی بقیه متغیرها ثابت در نظر گرفته می شوند، برقرار باشد. همچنین متغیرهای تصادفی $X_1, X_2, ..., X_n$ (یا تابع توزیع توام F)،RR_2 در زوج است ، اگر اندازه احتمال متناظر آن روی R^n ، در هر زوج RR_2 باشد.

تعریف ۷ : متغیرهای تصادفی $X_{1},...,X_{n}$ الف) به طورشرطی نزولی در دنباله (CDS) گفته میشوند ،اگر برای هر $i=1,...,n-1$ ، $[X_{i+1}|X_{1}=x_{1},...,X_{i}=x_{i}]$ به طور تصادفی در $x_{1},...,x_{i}$ نزولی باشد.
ب)وابسته منفی در دنباله (NDS) نامیده میشوند،اگربرای هر $i=1,...,n$ ، $[X_{1},...,X_{i-1}|X_{i}=x_{i}]$ در x_{i} به طور تصادفی نزولی باشد.

قضیه ۷ ([۱]) : فرض کنید $X_{0},X_{1},...,X_{n}$ متغیرهای تصادفی مستقل باشند و هر کدام یک تابع احتمال یا چگالی PF_{2} داشته باشند، آنگاه برای s ثابت ،متغیرهای تصادفی شرطی $(X_{1},...,X_{n}|X_{0}+X_{1}+...+X_{n}=s)$ ، RR_{2} در زوج هستند و در نتیجه CDS و NOD هستند.

نتیجه ۲ : هیچکدام از روابط RR_{2} در زوج ، CDS وNOD تحت اثر توابع صعودی روی مجموعه های مجزا بسته نیستند،اما رابطه NA بسته است.

مثال ۲ : فرض کنید $(X_{1},X_{2},X_{3})\sim Mult(3,p_{1},p_{2},p_{3})$ باتابع احتمال $P_{1}(i,j,k)$ ، که در آن $p_{i}>0$ ، $i=1,2,3$ و $X_{1}+X_{2}+X_{3}=3$ و $Y=(Y_{1},Y_{2})$ که در آن $Y_{1}=X_{1}.X_{2}$ و $Y_{2}=X_{3}$ دارای تابع احتمال توام $P_{2}(i,j)$ هستند . تابع احتمال $P_{1}(i,j,k)$ ،دارای ویژگیهای RR_{2} در زوج و در نتیجه CDS وNDS است ، ولی $P_{2}(i,j)$ دارای هیچکدام از این ویژگی ها نمی باشد .

از اینکه ، اگر $X_{i}\sim P(\lambda_{i})$ ، $i=1,...,n$ انگاه : $X_{1},...,X_{n}|\sum_{i=1}^{n}X_{i}=k\sim Mult(k,p_{1},...,p_{n})$ و چون چگالی پواسون PF_{2} است ،طبق قضیه (۷) متغیر های X_{1},X_{2},X_{3} ، RR_{2} در زوج، CDS ،NOD هستند.بنابراین تابع احتمال $P_{1}(i,j,k)$ دارای ویژگی های RR_{2} در زوج، CDS و NOD است . تابع چگالی توام $P_{2}(i,j)$ عبارتست از :

		Y_{1}		
		۰	۱	۲
	۰	$P_{1}(0,3,0)+P_{1}(3,0,0)$	۰	$P_{1}(1,2,0)+P_{1}(2,1,0)$
Y_{2}	۱	$P_{1}(0,2,1)+P_{1}(2,0,1)$	$P_{1}(1,1,1)$	۰
	۲	$P_{1}(0,1,2)+P_{1}(1,0,2)$	۰	۰
	۳	$P_{1}(0,0,3)$	۰	۰

به آسانی می توان نشان داد که، دترمینان مقابل بزرگتر از صفر است .

$$\begin{vmatrix} P_{2}(0,0) & P_{2}(0,1) \\ P_{2}(1,0) & P_{2}(1,1) \end{vmatrix} \geq 0$$

پس $P_{2}(i,j)$ نمی تواند RR_{2} باشد.
از اینکه $P(Y_{2}>0|Y_{1}=0)<P(Y_{2}>0|Y_{1}=1)$ نتیجه می گیریم که (Y_{1},Y_{2}) ، CDS نیست .
برای یک بردار دو متغیره CDS و NDS معادل هستند، در نتیجه Y ، NDS هم نیست.با استفاده از قضیه (۸) نتیجه می گیریم که $\underline{X}$ ، NA است و از ویژگی ششم ،NA بودن Y نتیجه می شود.

نتیجه ۳ : ویژگی NA ، هیچ یک از ویژگیهای RR_{2} در زوج ، CDS وNOD را نتیجه نمی دهد.

۲– کاربردها :

برخی از توزیع های چند متغیره مشهور که دارای ویژگی NA هستند، عبارتند از: الف) چند جمله ای، ب)پیچش (تلفیق)چند جمله ایهای نامشابه، ج)فوق هندسی چند متغیره، د)دیریکله، ه)چند جمله ای مرکب دیریکله و و)چند متغیره نرمال که ماتریس کوواریانس ویژه‌ای را داراست. برقرار بودن ویژگی NA را در موارد زیر بررسی می نمائیم :

توزیع چند جمله ای : فرض کنید $Z = (Z_1, ..., Z_k)$ برداری با توزیع چند جمله ای بر پایه تنها یک مشاهده باشد،بنابرین فقط یک Z_i برابر یک و بقیه برابر صفر هستند.

چون $Z_1 + ... + Z_k = N$ و $N = 1$ در نتیجه وجود دارد یک i که $Z_i = 1$ و $Z_j = 0 \quad \forall j \neq i$

NA بودن Z از تعریف (۱) و اینکه مجموع ، تابعی صعودی از متغیر هاست ، نتیجه می شود . از انجا که توزیع چند جمله ای تلفیق Z های مستقل است ،با استفاده از ویژگی ششم ،دارای ویژگی NA نیز می باشد.

با استفاده از این مطلب که متغیر های تصادفی چند جمله ای با متغیر های مستقل پواسون به شرط مجموع ثابتی از انها، هم توزیع هستند و قضیه (۴) می توان NA بودن توزیع چند جمله ای را نتیجه گرفت .

توزیع فوق هندسی چند متغیره :گلدانی حاوی N توپ است که هر کدام رنگهای متفاوتی دارند. فرض کنید نمونه ای تصادفی از n توپ بدون جایگذاری انتخاب شود و X_i ، $i = 1, ..., N$ متغیر تصادفی باشد که وجود یک توپ با i رنگ را در نمونه نشان می دهد . بدیهی است که متغیر های تصادفی X_i، $i = 1, ..., N$ توزیعی جایگشتی دارند. هر جایگشتی از N دارای توزیع مقابل است :

$$P(X_1 = x_1, ..., X_N = x_N) = \frac{1}{N}.\frac{1}{N-1}...1$$

و بنابرین NA هستند.

به طور کلی تر، فرض کنید N_i توپ رنگ i را دارند$i = 1, .., k$ ، که $\sum_{i=1}^{k} N_i = N$.فرض کنید Y_i تعداد توپ ها با رنگ i ام در نمونه باشد.Y_i را می توان به صورت مجموع N_i حالت ساده ،مانند بالا در نظر گرفت که در Y_i تای انها $X_i = 1$ است .

چون Y_i حاصلجمع مجموعه های مجزا از متغیر های تصادفی است که دارای خاصیت NA هستند،بنا به خاصیت ششم ، NA است.

توزیع نرمال چند متغیره با ماتریس همبستگی که عناصر آن منفی هستند :

با استدلالی مشابه اثبات اینکه متغیر های تصادفی نرمال همبسته مثبت ،مرتبط مثبت هستند، ثابت می شود که متغیر های تصادفی نرمال همبسته منفی ،NA هستند ، البته به دلیل اینکه در این حالت توابع روی مجموعه های مجزا تعریف می شوند ،اثبات ساده تر است .

نمونه گیری تصادفی بدون جایگذاری :

توزیع فوق هندسی چند متغیره حالت خاصی از این مدل است.

$$P(X_1 = x_1, ..., X_N = x_N) = \begin{pmatrix} M \\ N \end{pmatrix}^{-1} .[\prod_{i=1}^{n} \begin{pmatrix} M_i \\ t_i \end{pmatrix}]. \begin{pmatrix} M - \sum_{i=1}^{n} M_i \\ N - \sum_{i=1}^{n} t_i \end{pmatrix}.$$

برای مقادیر مثبت حقیقی مقدار

$$t_i \geq 0, \qquad \sum_{i=1}^{n} t_i \leq N, \qquad \sum_{i=1}^{n} M_i \leq M.$$

جامعه ای متناهی شامل N مقدار $x_1, ..., x_N$ را در نظر بگیرید.اگر $X_1, ..., X_n$ نمایانگر مقادیر نمونه بدست امده با نمونه گیری بدون جایگذاری باشند،انگاه $X_1, ..., X_n$ می توانند به عنوان زیر مجموعه ای از $X_1, ..., X_N$ در نظر گرفته شوند که توزیعی جایگشتی دارند و از خاصیت چهار نتیجه می شود که NA هستند.

توزیع توام رتبه ها :

فرض کنید $X_1, X_2, ..., X_n$ نمونه ای تصادفی از یک جامعه باشد.اگر R_i رتبه X_i، $i = 1, ..., n$ باشد،انگاه به وضوح $(R_1, ..., R_n)$ توزیع جایگشتی مانند زیر دارد ،

$$P(R_i = r) = (n-1)!/n! = 1/n$$

$$P(R_i = r, R_j = s) = P(R_i = r).P(R_j = s | R_i = r) = 1/n(n-1) \quad ...$$

$$P(R_1 = r_1, ..., R_k = r_k) = 1/n(n-1)...(n-k+1)$$

بنابرین NA است.

روشهای گزینش بر پایه رتبه بندی چند متغیره :

فرض کنید n شی توسط m مشاهده گر رتبه بندی می شوند. R_{ik} رتبه داده شده به i امین شی توسط k امین فرد باشد. $R_i = \sum_{k=1}^{m} R_{ik}$. فرض بر این استکه هیچ برتری بین اشیا وجود ندارد.

برای k ثابت ،بردار $(R_{1k}, ..., R_{nk})$، دارای توزیع جایگشتی است و بنابرین NA است .

$(R_{11}, ..., R_{n1}), ..., (R_{1m}, ..., R_{nm})$ دو به دو مستقل هستند ،در نتیجه $(R_1, ..., R_n)$، NA است.

نا مساوی های گشتاوری :

از ویژگی دوم می توان در بدست اوردن نامساوی های گشتاوری بهره گرفت.

فرض کنید $Y_1, ..., Y_k$ متغیرهای تصادفی مثبت NA باشند، $\alpha_i \geq 0$ ، $i = 1, ..., m$ که $m \leq k$ انگاه، چون Y_α تابعی صعودی است :

$$\mu_{\alpha_1, \alpha_2, ..., \alpha_m} \leq \mu_{\alpha_1} . \mu_{\alpha_2} ... \mu_{\alpha_m}$$

که :

$$\mu_{\alpha_i} = E(Y_i^{\alpha_i})$$

$$\mu_{\alpha_1,\alpha_2,...,\alpha_m} = E(Y_1{}^{\alpha_1}.Y_2{}^{\alpha_2}...Y_m{}^{\alpha_m})$$

به ویژه :

$$E(Y_1.Y_2...Y_m) \leq \prod_{i=1}^{m} E(Y_i)$$

در نامساوی های بالا $Y_1, ..., Y_m$ نشاندهنده هر زیر مجموعه m تایی از $Y_1, ..., Y_k$ است.

وابستگی میان فراوانی حجره ها در داده های رسته ای :

فرض کنید نمونه مورد نظر می تواند در یکی از دسته های A_i، $i = 1, ..., r$،متناظر با مشخصه A و در یکی از دسته های B_j ، $j = 1, ..., k$ ، متناظر با مشخصه B ، طبقه بندی شود.آنالیز رسته ای به طور معمول ، استقلال بین دو مشخصه A و B رابررسی می کند. برای ازمون چنین فرضیه ای مدل اماری معمول بر طبق زیر فرض می شود:
فرض کنید $n_1, ..., n_k$ اندازه های نمونه های تصادفی گرفته شده از زیر جامعه های شکل گرفته بر اساس افراز جامعه به رسته های $B_1, ...B_k$ باشند و X_{ij} فراوانی حجره متناظر با B_i وA_j همانند جدول زیر باشند:

	A_1	A_2	...	A_r	$totals$
B_1	X_{11}	X_{12}	...	X_{1r}	n_1
B_2	X_{21}	X_{22}	...	X_{2r}	n_2
...	...	...	...	...	...
B_k	X_{k1}	X_{k2}	...	X_{kr}	n_k
$totals$	T_1	T_2	...	T_r	$n = \sum n_i$

تحت فرض استقلال، مدل بالا نتیجه می دهدکه عناصر هر سطر توزیع چند جمله ای با پارامتر های $(P_1, ..., P_k)$ دارند.

در ادامه نشان می دهیم ، هنگامی که مجموع های حاشیه ای ثابت هستند ، وجود ارتباط مثبت نیز همانند ارتباط منفی ،آشکار می شود. تحت فرض بردار مجموع های ستونی $(T_1, ..., T_r)$ اماره بسنده است ،زیرا که توزیع توام $\{X_{ij}\}$ با شرط $T_j = t_j$ ، $j = 1, ..., r$ برابر است با:

$$P(X_{ij} = x_{ij}, i = 1, ..., k, j = 1, ..., r) = \frac{\prod_{j=1}^{r} t_j{}^r . \prod_{i=1}^{k} n_i!}{n!. \prod_{j=1}^{r} \prod_{i=1}^{k} x_{ij}!}$$

که مستقل از $P_1, .., P_k$ است.

در واقع دو ادعای زیر اثبات می شود ([۴]) :

۱—توزیع های حاشیه ای بردار های سطری (ستونی) ، دارای خاصیت NA هستند.

۲—توزیع حاشیه ای یک مجموعه از فراوانی حجره ها ، به طوریکه هیچ جفت از آنها در یک سطر یا ستون نباشند(برای مثال حجره های قطری)دارای خاصیت ارتباط مثبت ، هستند.

۳— نتیجه گیری :

در مسائل تئوری ، قضایا بیشتر در حالت استقلال متغیر ها مورد بحث قرار گرفته اند.اما در مسائل کاربردی و عملی ،

در بیشتر موارد با متغیر های غیر مستقل سر و کار داریم و این موضوع باعث اهمیت مطالعه و بررسی متغیر های غیر مستقل شده است . بنابر این شرط ضعیف تری نسبت به شرط استقلال در نظر گرفته شده و روی این خانواده بزرگترشامل متغیر های مرتبط ،مطالعات بسیاری انجام گرفته است . تا کنون دامنه کاربرد متغیر های مرتبط‌به مسائلی مانند تحلیل رگرسیون ، ازمون های ناپارامتری ، قضایای حدی ، قانون اعداد بزرگ و ... تعمیم داده شده است . مسائل بسیاری در مورد متغیر های تصادفی مرتبط به ویژه NA هنوز حل نشده است ،که بسیاری از امار دانان در این زمینه در حال مطالعه و تحقیق هستند.

مراجع :

[۱]Block,H.W.,Savits,T. and Shaked,M.(1982) Some concepts of negative dependence. Ann.Probab. 10 765-772.

[۲]Ebrahimi,N. and Ghosh,M.(1981).Multivariate negative dependence .Comm.Statist. A10(4) 307-337.

[۳]Efron,B.(1965)Increasing properties of Polya frequency functions .Ann.Math.Statist. 36 272-279.

[۴]Kumar,J.D. and Proschan,F.(1982).Negative association of random variables ,with applications.Ann.Statist.11,286-295.

[۵]Lehman,E.L.(1966)Some concepts of dependence.Ann.Math.Statist. 43 1137-1153.