انواع پیوندهای منفی و روابط بین آنها

محمد امینی دهک*

ابوالقاسم بزرگ‌نیا†

چکیده

عموماً تنها معیاری که برای تعیین نوع پیوند (منفی یا مثبت) بین متغیرهای تصادفی به کار می‌رود، ضریب همبستگی است. در این مقاله معیارهای متنوع دیگری معرفی می‌شوند که قوی‌تر از ضریب همبستگی‌اند و نوع پیوند بین متغیرها با استفاده از این معیارها تعیین می‌شود. چون هدف بررسی پیوندهای منفی است لذا این معیارها را تحت عنوان پیوندهای منفی بیان می‌کنیم و به بررسی خواص و رابطه‌های بین آنها در حالت دو متغیره می‌پردازیم. سپس کاربرد پیوندهای منفی در مسائل خطایابی را بررسی می‌کنیم.

مقدمه

لیمن [۴] در سال ۱۹۶۶ میلادی مفاهیم اولیه وابستگی منفی و مثبت را در حالت دو متغیره ارائه داد و قواعد قوی‌تری برای تعیین وابستگی مثبت بعداً در سال ۱۹۷۲ به وسیله اساری و پروشان [۳] بسط یافت. همچنین در سال ۱۹۶۷ اساری، پروشان و واکاپ [۲] مفهوم ارتباط را مطرح کردند که مفهومی قوی‌تر از وابستگی مثبت است. اما با آنکه در بارهٔ وابستگی مثبت مقاله‌های زیادی نوشته شده است، در بارهٔ وابستگی منفی کمتر مقاله‌ای ارائه شده است. پس از آنکه مفاهیم اولیه وابستگی منفی توسط لیمن ارائه شد، تعمیم چند متغیره مفاهیم وابستگی منفی بعداً به وسیله آماردانان آغاز شد تا اینکه در سال ۱۹۸۱ ابراهیمی و قوش تعمیم چند متغیره وابستگی منفی را توسعه دادند و در سال ۱۹۹۳ بزرگنیا و تایلور [۱] ارتباط منفی بین متغیرها را مطرح و قضایای حدی برای متغیرهای تصادفی به طور منفی وابسته را بررسی کردند. در این مقاله، مطالب بیشتری دربارهٔ انواع وابستگی‌های منفی، ارتباط بین آنها، و خواص آنها در حالت دو متغیره گردآوری شده است.

۱ وابستگی مربعی

۱.۱ تعریف

زوج تصادفی (x,y) را به طور منفی وابستهٔ مربعی (NQD)[۱] گوییم هرگاه

(۱)

$$P(X\leq x, Y\leq y)\leq P(X\leq x)\cdot P(Y\leq y);\qquad \forall x,y\in\mathbb{R}$$

خانواده تمام توزیعهایی را که در نامساوی (۱) صدق می‌کنند با $\mathcal{G}_1$ نشان می‌دهیم.

*) محمد امینی دهک، بخش آمار، دانشگاه فردوسی مشهد

†) دکتر ابوالقاسم بزرگ‌نیا، بخش آمار، دانشگاه فردوسی مشهد

1) Negatively Quadratic Dependent

۲.۱ تعریف

زوج تصادفی (X,Y) را به طور مثبت وابستهٔ مربعی (PQD)[۲] گوییم هرگاه

$$P(X\leq x,Y\leq y)\geq P(X\leq x)\cdot P(Y\leq y)\qquad \forall x,y\in\mathbb{R} \tag{۲}$$

خانوادهٔ تمام توزیعهایی راکه درنامساوی صدق می‌کنند با $\mathcal{F}_۱$ نشان می‌دهیم.

۳.۱ لم

(الف) در نامساوی (۱) هرکدام از نامساویهای $Y\leq y$ یا $X\leq x$ را می‌توان نظیر به نظیر با $Y<x$ یا $X<x$ جایگزین کرد.

(ب) هرگاه نامساوی (۱) برقرار باشد هر یک از نامساویهای زیر را می‌توان نتیجه گرفت:

$$P(X\leq x,Y\geq y)\geq P(X\leq x)\cdot P(Y\geq y) \tag{۳}$$

$$P(X\geq x,Y\leq y)\geq P(X\geq x)\cdot P(Y\leq y) \tag{۴}$$

$$P(X\geq x,Y\geq y)\leq P(X\geq x)P(Y\geq y) \tag{۵}$$

(ج)

$$(X,Y)\in\mathcal{F}_۱\Longleftrightarrow(X,-Y)\in\mathcal{G}_۱$$

برهان قسمتهای (الف) و (ب) به سادگی اثبات می‌شود. ما تنها قسمت (ج) را ثابت می‌کنیم.

(ج) فرض کنید $(X,Y)\in\mathcal{F}_۱$ در این صورت

$$P(X\leq x,Y\leq y)\geq P(X\leq x)\cdot P(Y\leq y),\ \forall x,y\in\mathbb{R}.$$

در نتیجه

$$\begin{aligned}P(X\leq x,-Y\leq y)&=P(X\leq x,\quad Y\geq -y)\\&=P(X\leq x)-P(X\leq x,Y<-y)\\&\leq P(X\leq x)-P(X\leq x)\cdot P(Y<-y)\\&=P(X\leq x)\cdot P(-Y\leq y)\end{aligned}$$

بنابراین

$$(X,-Y)\in\mathcal{G}_۱$$

عکس حکم نیز به همین صورت ثابت می‌شود.

۴.۱ تعریف

تابعهای حقیقی $f(x_۱,x_۲,\ldots,x_n)$ و $g(y_۱,y_۲,\ldots,y_n)$ را در iامین مختص هماهنگ گوییم اگر نسبت به آن مختص هر دو غیر نزولی یا هر دو غیر صعودی باشند و آنها را در iامین مختص ناهماهنگ گوییم اگر نسبت به آن مختص یکی غیر نزولی و دیگری غیر صعودی (یکی غیرصعودی و دیگری غیر نزولی) باشد.

۵.۱ قضیه

فرض کنید زوجهای $(X_۱,Y_۱),\ldots,(X_n,Y_n)$ مستقل باشند و توزیعهای توأم آنها را به ترتیب با $F_۱,\ldots,F_n$ نشان دهیم. اگر f و g تابعهایی از n متغیر باشند و

$$\mathbf{X}=f(X_۱,X_۲,\ldots,X_n),\ \mathbf{Y}=g(Y_۱,Y_۲,\ldots,Y_n)$$

آنگاه

(الف) $(\mathbf{X},\mathbf{Y})\in\mathcal{G}_۱$ هرگاه برای هر i، $i=۱,۲,\ldots,n$ یکی از شرطهای زیر صادق باشد.

۱) $F_i\in\mathcal{G}_۱$ و f و g در iامین مختص هماهنگ باشند.

۲) $F_i\in\mathcal{F}_۱$ و f و g در iامین مختص ناهماهنگ باشند.

(ب) متشابهاً $(\mathbf{X},\mathbf{Y})\in\mathcal{F}_۱$ اگر برای هر i، $F_i\in\mathcal{F}_۱$ و f و g در iامین مختص هماهنگ باشند یا $F_i\in\mathcal{G}_۱$ و f و g در iامین مختص ناهماهنگ باشند.

(ج) اگر U و V متغیرهای تصادفی مستقل و همچنین مستقل از $(X_۱,Y_۱),\ldots,(X_n,Y_n)$ باشند و

$$X=f(U,\mathbf{X}),\ Y=g(V,\mathbf{Y}),$$

$$\mathbf{X}=(X_۱,\ldots,X_n),\ \mathbf{Y}=(Y_۱,\ldots,Y_n)$$

آنگاه نتایج (الف) و (ب) تحت مفروضات فوق و بدون هیچ فرضی در رابطه با رفتار f و g نسبت به U و V برقرارند. (به دلیل طولانی بودن اثبات از بیان آن صرف‌نظر می‌کنیم).

2) Positively Quadratic Dependent

۶.۱ مثال

(الف) به ازای هر متغیر تصادفی X داریم

$$(X,-X)\in \mathcal{G}_1$$

حل: به راحتی ثابت می‌شود که $(X,X)\in\mathcal{F}_1$ و بنا به لم ۱-۳ (ج) نتیجهٔ مطلوب حاصل می‌شود.

(ب) هرگاه f تابعی غیر صعودی باشد آنگاه به ازای هر متغیر تصادفی X داریم

$$(X,f(X))\in\mathcal{G}_1$$

حل: فرض‌کنید $g(x)=x$ بنا به قسمت (الف) قضیهٔ ۱-۵ حکم ثابت می‌شود.

(ج) اگر

$$Y=V+bZ \quad , \quad X=U+aZ$$

برای متغیرهای تصادفی مستقل دلخواه U، V، Z و مقادیر حقیقی مختلف‌العلامهٔ a و b آنگاه

$$(X,Y)\in\mathcal{G}_1$$

(د) اگر $Y=V-X$ برای متغیرهای تصادفی مستقل دلخواه X و V،

$$(X,Y)\in\mathcal{G}_1$$

(ه) اگر $X=f(U,Z)$ و $Y=g(V,Z)$ که در آن U،Z ، V مستقل و f و g توابعی ناهماهنگ در z اند. آنگاه

$$(X,Y)\in\mathcal{G}_1$$

مثال فوق با استفاده از قضیهٔ ۱-۵ به راحتی اثبات می‌شود.

۷.۱ لم

(هافدینگ)[3] اگر F، F_X، F_Y به ترتیب تابعهای توزیع توأم و حاشیه‌ای X و Y باشند، آنگاه با شرط وجود $E|XY|$، $E|X|$، و $E|Y|$،

$$\mathrm{Cov}(X,Y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left[F(x,y)-F_X(x)F_Y(y)\right]dxdy$$

(به دلیل طولانی بودن اثبات از آن صرف‌نظر می‌شود)

۸.۱ قضیه

$(X,Y)\in\mathcal{G}_1$ اگر و تنها اگر برای تمام تابعهای غیرنزولی f و g

$$\mathrm{Cov}(f(X),g(Y))\leq 0$$

برهان فرض کنید $(X,Y)\in\mathcal{G}_1$ بنابر قضیهٔ ۱-۵ در حالت $m=1$، چون f و g هماهنگ هستند

$$(f(X),g(Y))\in\mathcal{G}_1$$

و حکم قضیه بنا بر لم ۱-۷ ثابت می‌شود.

برای اثبات عکس حکم، تابعهای غیرنزولی f و g را به صورت زیر تعریف می‌کنیم:

$$f_x(X)=\begin{cases}1 & , \quad X>x \quad \text{اگر} \\ 0 & , \quad X\leq x \quad \text{اگر}\end{cases}$$

$$g_y(Y)=\begin{cases}1 & , \quad Y>y \quad \text{اگر} \\ 0 & , \quad Y\leq y \quad \forall x,y\in R \quad \text{اگر}\end{cases}$$

بنا به فرض داریم:

$$\begin{aligned}0 &\geq \mathrm{Cov}(f_x(X),g_y(Y))=\mathrm{Cov}(1-f,1-g)\\ &=P(X\leq x,Y\leq y)-P(X\leq x)\cdot P(Y\leq y)\end{aligned}$$

پس $(X,Y)\in\mathcal{G}_1$.

۹.۱ نتیجه

فرض کنید $(X,Y)\in\mathcal{G}_1$ و $(X,Y)\rightsquigarrow F$ آنگاه τی کندال[4] و ρ_s اسپیرمن[5] منفی‌اند.

برهان فرض‌کنید (X_1,Y_1) و (X_2,Y_2) مستقل و با (X,Y) همتوزیع باشند. تعریف می‌کنیم

$$X=Sgn(X_2-X_1)=\begin{cases}1 & X_2-X_1>0 \quad \text{اگر}\\ 0 & X_2-X_1=0 \quad \text{اگر}\\ -1 & X_2-X_1<0 \quad \text{اگر}\end{cases}$$

3) Hoefdding 4) Kendall 5) Spearman

$$Y = Sgn(Y_۲ - Y_۱) = \begin{cases} ۱ & \text{اگر } Y_۲ - Y_۱ > ۰ \\ ۰ & \text{اگر } Y_۲ - Y_۱ = ۰ \\ -۱ & \text{اگر } Y_۲ - Y_۱ < ۰ \end{cases}$$

چون $\tau = EXY$ و $EX = EY = ۰$ پس بنابر قضیهٔ ۱-۵

$$\tau = \mathrm{Cov}(X,Y) \le ۰.$$

با توجه به رابطه‌ای که بین τی کندال و ρ_s اسپیرمن به صورت زیر برقرار است، داریم:

$$\frac{\rho_s}{۳} = \frac{\tau}{(n+۱)} + \frac{n-۲}{n+۱}(۲\pi_{C_r} - ۱)$$

که در آن

$$\pi_{C_۲} = P(X_j - X_i)(Y_k - Y_i) > ۰ \quad i < j,\ k \neq j$$

پس هرگاه $(n = ۳)$، $(X_۱, Y_۱), (X_۲, Y_۲), (X_۳, Y_۳)$ همتوزیع با F و مستقل باشند داریم،

$$\begin{aligned} \pi_{C_۲} &= \mathrm{Cov}(Sgn(X_۲ - X_۱), Sgn(Y_۲ - Y_۱)) \\ &= ۲\iint F(x,y)f(x)f(y)dxdy \\ &\le ۲\iint F_x(x)F_y(Y)f(x)f(y)dxdy \\ &= \frac{۱}{۲} \\ &\Longrightarrow ۲\pi_{C_۲} - ۱ \le ۰, \qquad \tau \le ۰ \end{aligned}$$

پس

$$\rho_s \le ۰$$

۲ وابستگی رگرسیونی

۱.۲ تعریف

زوج تصادفی (X,Y) را به طور منفی وابستهٔ رگرسیونی (NRD)[6] گوییم هرگاه $P(Y \le y|X = x)$ به ازای کلیهٔ مقادیر حقیقی x و y تابعی غیرنزولی نسبت به x باشد یا می‌گوییم Y در X به طور منفی وابسته رگرسیونی است.

بعداً نشان می‌دهیم که رابطهٔ $P(Y \le y|X = x)$ نسبت به X و Y متقارن نیست ولی در لم ۱-۳ ثابت کرده‌ایم که رابطهٔ (۱) متقارن است.

خانوادهٔ تمام توزیعهایی را که در تعریف فوق صدق می‌کنند با $\mathcal{G}_۲$ نشان می‌دهیم.

هرگاه به ازای کلیهٔ مقادیر حقیقی x و y، $P(Y \le y|X = x)$ تابعی غیرصعودی نسبت به x باشد، آنگاه (X,Y) را به طور مثبت وابستهٔ رگرسیونی[7] (PRD) گوییم. خانوادهٔ تمام توزیعهایی را که در شرط فوق صدق کنند با $\mathcal{F}_۲$ نشان می‌دهیم.

۲.۲ قضیه

هرگاه $Y = \alpha + \beta X + U$ و $\beta \le ۰$ و $(X,U) \in \mathcal{G}_۲$ آنگاه

$$(X,Y) \in \mathcal{G}_۲$$

۳.۲ لم

(الف) زوج (X,Y)،NRD است اگر و تنها اگر $(-X,Y)$،PRD باشد.

(ب) زوج (X,Y)،NRD است اگر و تنها اگر $(X,-Y)$،PRD باشد.

۴.۲ مثال

(الف) اگر $(X,Y) \rightsquigarrow N(\mu_۱, \mu_۲, \sigma_۱^۲, \sigma_۲^۲, \rho)$ آنگاه

$$(X,Y) \in \mathcal{G}_۲ \Longleftrightarrow \rho \le ۰$$

(ب) اگر $(U_۱, U_۲, \dots, U_s)$ دارای توزیع چند جمله‌ای متناظر با n آزمایش و احتمالهای موفقیت $p_۱$، $p_۲$، ...، p_s باشد، آنگاه

$$(U_i, U_j) \in \mathcal{G}_۲ \qquad ;\forall i \neq j$$

(ج) فرض کنید $X_۱$، $X_۲$، ...، X_n یک نمونهٔ تصادفی از توزیع F باشد و $X_{u_۱} < X_{u_۲} < \cdots < X_{u_m}$، $(m < n)$ رکوردهای بالای نمونه فوق باشند. اگر $\frac{۱-F(x+a)}{۱-F(x)}$ به ازای هر $\alpha > ۰$ تابعی صعودی نسبت به x باشد، آنگاه

$$(X_{u_m}, X_{u_n} - X_{u_m}) \in \mathcal{G}_۲, \qquad m \le n.$$

6) Negalively Regressional Dependent 7) Positively Regressional Dependent

۳ وابستگی نسبت درستنمایی

۱.۳ تعریف

فرض کنید زوج (X,Y) دارای تابع چگالی احتمال توأم $f(x,y)$ باشد در این صورت (X,Y) را به طور منفی وابستهٔ نسبت درستنمایی[8] (NLRD) گوییم هرگاه

$$f(x,y')f(x',y) \geq f(x,y)f(x',y') \qquad \forall x < x', y < y'. \tag{۶}$$

شرط (۱) همراه با شرط $f(x) \neq 0$، $\forall x$ با شرط زیر هم ارز است:

$$\frac{f(y|x')}{f(y|x)} \geq \frac{f(y'|x')}{f(y|x)} \qquad \forall x < x', y < y'. \tag{۷}$$

پس (X,Y) را به طور منفی وابسته درستنمایی گوییم هرگاه $\frac{f(y|x')}{f(y|x)}$ $(x < x')$ تابعی غیر صعودی بر حسب y باشد.

خانواده تمام توزیعهای F را که در شرط (۱) یا (۲) صدق کنند به $\mathcal{G}_3$ نشان می‌دهیم.

۲.۳ تعریف

اگر $\frac{f(y|x')}{f(y|x)}$، $x < x'$، تابعی غیرنزولی بر حسب y باشد (X,Y) را دارای پیوند (PLDR) گوییم و خانوادهٔ تمام توزیعهای F را که در تعریف فوق صدق می‌کنند با $\mathcal{F}_3$ نشان می‌دهیم.

۳.۳ مثال

(الف) توزیع نرمال دو متغیره به $\mathcal{G}_3$ تعلق دارد اگر و تنها اگر $\rho \leq 0$.

(ب) هر دو مؤلفه از توزیع چند جمله‌ای مثال ۲-۴ (ب) عضو $\mathcal{G}_3$اند.

(ج) توزیع دیریکله با چگالی زیر در $\mathcal{G}_3$ است.

$$d(x,y) = Cx^{f_1}y^{f_2}(1-x-y)^{f_3}$$

$$0 \leq x+y \leq 1, x,y \geq 0,$$

$$f_1 = f_2 = \frac{n-1}{2} - 1$$

$$f_3 = \frac{(n-1)(s-2)}{2} - 1$$

۴ ارتباط بین متغیرها

۱.۴ تعریف

(بزرگنیا و تایلور (۱۹۹۳)) زوج (X,Y) را به طور منفی مرتبط[9] گوییم هرگاه به ازای تمام تابعهای غیرنزولی f و g داشته باشیم،

$$\mathrm{Cov}(f(X), g(Y)) \leq 0$$

خانوادهٔ تمام توزیعهای F مربوط به (X,Y) را که در تعریف فوق صدق می‌کنند با $\mathcal{G}_4$ نشان می‌دهیم.

۲.۴ قضیه

تابعهای غیر نزولی از متغیرهای تصادفی به طور منفی مرتبط به طور منفی مرتبط‌اند.

۳.۴ قضیه

اگر به ازای تمام تابعهای دوتایی غیرنزولی λ و δ داشته باشیم

$$\mathrm{Cov}(\lambda(X), \delta(Y)) \leq 0$$

آنگاه

$$(X,Y) \in \mathcal{G}_4.$$

برهان برای تابعهای غیرنزولی f و g و تابعهای دوتایی غیرنزولی $X_f(x)$ و $X_g(y)$ را به صورت زیر تعریف می‌کنیم.

$$X_f(x) = \begin{cases} 1 & f(X) > x \quad \text{اگر} \\ 0 & f(X) \leq x \quad \text{اگر} \end{cases}$$

$$X_g(y) = \begin{cases} 1 & g(Y) > y \quad \text{اگر} \\ 0 & g(Y) \leq y \quad \text{اگر} \quad \forall x, y \in R \end{cases}$$

با استفاده از لم ۱-۷ ثابت می‌شود که

$$\mathrm{Cov}(f(X) \cdot g(Y)) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{Cov}(X_f(x), X_g(y))dxdy$$

با توجه به فرض و تعریف ۴-۱ حکم ثابت است.

---

8) Negatively Likelihood Ratio Dependent 9) Negatively Associated

۴.۴ قضیه

اگر X و Y متغیرهای تصادفی دوتایی باشند و $(X,Y) \in \mathcal{G}_۲$ آنگاه

$$(۱-X, ۱-Y) \in \mathcal{G}_۲$$

۵.۴ نتیجه

اگر X و Y متغیرهای تصادفی دوتایی باشند، آنگاه

$$(X,Y) \in \mathcal{G}_۲ \Longleftrightarrow \mathrm{Cov}(X,Y) \leq ۰$$

۶.۴ قضیه

برای هر زوج تصادفی (X,Y) داریم

$$(X,Y) \in \mathcal{G}_۲ \Longleftrightarrow (X,Y) \in \mathcal{G}_۱$$

و یا

$$\mathcal{G}_۱ = \mathcal{G}_۲.$$

برهان بنا بر قضیهٔ ۱-۸ و تعریف، حکم ثابت است.

تذکر

با وجود اینکه در حالت دو متغیره، $\mathcal{G}_۱ = \mathcal{G}_۲$ ولی به ازای مقادیر n بزرگتر از ۲ ثابت می‌شود که برای دنباله‌ای از متغیرهای تصادفی $X_۱$، $X_۲$، ...، X_n،

$$\mathcal{G}_۲ \subset \mathcal{G}_۱$$

روابط بین پیوندها

۵ پیوندهای (ND) و (NRD)

۱.۵ قضیه

$$\mathcal{G}_۲ \subset \mathcal{G}_۱.$$

برهان فرض کنید $(X,Y) \in \mathcal{G}_۲$، بنابر تعریف،

$$P(Y > y | X = x) = ۱ - P(Y \leq y | X = x), \ \forall x, y \in R$$

تابعی غیرصعودی از X است. تابع $h(u)$ را به صورت زیر تعریف می‌کنیم.

$$h(u) = P(Y > y | X = u), \quad \forall u \in R$$

داریم

$$P(Y > y, X \leq x) = \int_{-\infty}^{x} P(Y > y | X = u) dF_X(u)$$
$$= \int_{-\infty}^{x} h(u) dF_x(u) \geq \int_{-\infty}^{x'} h(u) dF_X(u) = P(Y > y, X \leq x')$$

در نتیجه

$$P(Y \leq y | X \leq x) \leq P(Y \leq y | X \leq x'), \ \forall x < x'$$

حال اگر $x' \longrightarrow \infty$، آنگاه

$$P(Y \leq y | X \leq x) \leq P(Y \leq y)$$
$$\Longrightarrow P(X \leq x, Y \leq y) \leq P(X \leq x) P(Y \leq y), \ \forall x, y \in R$$

پس

$$(X,Y) \in \mathcal{G}_۱ \Longrightarrow \mathcal{G}_۲ \subset \mathcal{G}_۱.$$

۲.۵ نتیجه

$$\mathcal{G}_۱ \not\subset \mathcal{G}_۲$$

۳.۵ مثال

فرض کنیم (X,Y) دارای توزیع احتمال توأمی مطابق با جدول زیر باشد.

$y \backslash x$	۱	۲	۳	
۱	۰	$\frac{۱}{۹}$	$\frac{۲}{۹}$	$\frac{۳}{۹}$
۲	$\frac{۱}{۹}$	$\frac{۱}{۹}$	۰	$\frac{۲}{۹}$
۳	$\frac{۲}{۹}$	$\frac{۱}{۹}$	$\frac{۱}{۹}$	$\frac{۴}{۹}$
	$\frac{۳}{۹}$	$\frac{۳}{۹}$	$\frac{۳}{۹}$	

۱ـ توزیع احتمال توأم (X,Y)

برای این توزیع بسهولت ثابت می‌شود که $(X,Y) \in \mathcal{G}_۱$. اما

$$P(X \leq ۲ | Y = ۲) = ۱ > P(X \leq ۲ | Y = ۳) = \frac{۳}{۴}$$

در نتیجه،

$$(Y,X) \notin \mathcal{G}_۲$$

از طرفی جدول ۲ در زیر نشان می‌دهد که $(X,Y) \in \mathcal{G}_۲$. پس این مثال

نشان می‌دهد که رابطه و تعریف ۲-۱ نسبت به X و Y نامتقارن است.

$P(Y=y\mid X=x)$ \ y	۱	۲	۳
$X=1$	0	$\frac{1}{3}$	$\frac{2}{3}$
$X=2$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$X=3$	$\frac{2}{3}$	0	$\frac{1}{3}$

۲ـ توزیع احتمال شرطی $Y|X$

۶ پیوندهای NRD و NLRD

۱.۶ قضیه

$$\mathcal{G}_3 \subset \mathcal{G}_2 \subset \mathcal{G}_1.$$

برهان فرض کنید $(X,Y)\in\mathcal{G}_3$ در نتیجه $\frac{f(y|x')}{f(y|x)}$، $x<x'$، تابعی غیرصعودی بر حسب y است تابع سادهٔ $\psi(y)$ را برای هر y به صورت زیر تعریف می‌کنیم.

$$\psi(y)=\begin{cases} 1 & y\le y_\circ \quad \text{اگر} \\ & \forall y_\circ\in R \\ 0 & y>y_\circ \quad \text{اگر}\end{cases}$$

حال فرض کنیم Y_1 و Y_2 مستقل و همتوزیع با Y باشند تعریف می‌کنیم:

$$E_x\phi(Y)=\int\phi(y)f(y|x)dy$$

داریم

$$I=\frac{1}{2}E_x\left[(\psi(Y_1)-\psi(Y_2))\left(\frac{f(Y_1|x')}{f(Y_1|x)}-\frac{f(Y_2|x')}{f(Y_2|x)}\right)\right]\ge 0$$

پس از انجام محاسبات نتیجه می‌شود که

$$E_{x'}\psi(Y)\ge E_x\psi(Y)$$

در نتیجه

$$\int_{-\infty}^{y_\circ}f(y|x)dy\le\int_{-\infty}^{y_\circ}f(y|x')dy,\qquad \forall x<x'$$

پس

$$P(Y\le y_\circ|X=x)\le P(Y\le y_\circ|X=x'),\ \forall x<x',\ y_\circ\in R$$

بنابراین

$$(X,Y)\in\mathcal{G}_2.$$

۲.۶ نتیجه

$$\mathcal{G}_2\not\subset\mathcal{G}_3.$$

۳.۶ مثال

فرض کنید $Y=U-X$ و U,X مستقل باشند در این صورت بنابر قضیهٔ ۱-۲، $(X,Y)\in\mathcal{G}_2$. اما همواره نمی‌توان نتیجه گرفت که $(X,Y)\in\mathcal{G}_3$.

حل برای اثبات این حکم فرض کنید $h(u)=\frac{1}{\pi(1+u^2)},u\in R$ و X یک متغیر تصادفی مستقل از U باشد آنگاه رابطهٔ

$$f(x,y')f(x',y)\ge f(x,y)f(x',y'),\ \forall x<x',\forall y<y'$$

همواره برقرار نیست. زیرا

$$f(x,y)=g(x)h(x+y),\ g(x)\neq 0$$

پس کافی است رابطهٔ زیر را بررسی کنیم:

$$h(x+y')\cdot h(x'+y)\ge h(x+y)h(x'+y'),\ x<x',y<y'$$

که رابطه فوق نیز همواره برقرار نیست.

با توجه به مطالب بیان شده در این قسمت مقاله می‌توان جدول زیر را تنظیم کرد:

	NLRD	NRD	ND	NA
NLRD		$\uparrow$	$\uparrow$	$\uparrow$
NRD	$\not\uparrow$		$\uparrow$	$\uparrow$
ND	$\not\uparrow$	$\not\uparrow$		$\uparrow$
NA	$\not\uparrow$	$\not\uparrow$	$\not\uparrow$	

۷ کاربرد در مسائل خطایابی

مسئلهٔ آزمون فرض H را در برابر مجموعهٔ فرضهای مقابل $K_i, i=1,\dots,s$ در نظر بگیرید و فرض کنید فرض H به نفع K_i در سطح α_i رد می‌شود اگر و تنها اگر

$$T_i\ge C_i,\ i=1,\dots,s,\ C_i\in R,\ \alpha_i=P_H[T_i\ge C_i]$$

فرض کنید s پارامتر θ_1، ...، θ_s تحت فرض H مساوی باشند و یا دقیقاً یکی از آنها با دیگران اختلاف داشته باشند.

$$\begin{cases} H: \theta_1 = \theta_2 = \cdots = \theta_s \\ K_i: \theta_i \neq \theta_j, \quad j = 1, \ldots, s, \quad i \neq j \end{cases}$$

در چنین مواقعی هدف کنترل احتمال Pی رد کردن H به خطا به نفع هر یک از فرضهای مقابل K_i است. با به کار بردن برابری بونفرونی و

$$\alpha = \sum_{i=1}^{s} \alpha_i, (T_i, T_j) \in \mathcal{G}_1 \; i \neq j, \; i, j = 1, \ldots, s$$

و انجام محاسبات لازم، داریم

$$\sum_{j=1}^{s} P_H[T_j \geq C_j] - \sum_{i<j}\sum P_H[T_i \geq C_i, T_j \geq C_j] \leq P$$
$$= P_H[\bigcup_{j=1}^{s}(T_j \geq C_j)] \leq \sum_{j=1}^{s} P_H[T_j \geq C_j]$$
$$\alpha - \frac{1}{2}(1 - \frac{1}{s})\alpha^2 \leq P \leq \alpha.$$

نابرابری فوق نشان می‌دهد که برای مقادیر کوچک α، α تقریب خوبی برای P است. نابرابری فوق در صورتی صادق است که هر زوج از آماره‌های آزمون در $\mathcal{G}_1$ باشند:

$$\forall i \neq j, \quad (T_i, T_j) \in \mathcal{G}_1$$

۱.۷ مثال

فرض کنیم n موضوع به وسیله m فرد به طور مستقل رتبه‌بندی می‌شوند و H: بین موضوعات تفاوتی وجود ندارند، بنابراین هر یک از افراد، آنها را به تصادف رتبه‌بندی می‌کند.

K_i: iـامین موضوع برتری دارد و دیگر موضوعات تفاوتی ندارند.

فرض کنید: رتبه iامین موضوع توسط Kامین فرد $= R_{ik}$

$$R_i = \sum_{k=1}^{m} R_{ik} =$$ مجموعه رتبه‌های iـامین موضوع

فرض H به نفع K_i رد می‌شود اگر و تنها اگر R_i به اندازه کافی بزرگ باشد.

$$\alpha_i = P_H[RH] = P_H[R_i \geq C_i]$$

پس از انجام محاسبات ثابت می‌شود که

$$P(R_{ik} \leq x, R_{jk} \leq y) = \frac{xy}{n(n-1)} - \frac{\min(x,y)}{n(n-1)}$$
$$P(R_{ik} \leq x) = \frac{x}{n}, \quad \forall k, \forall x, y \in R$$
$$P(R_{ik} \leq x, R_{jk} \leq y) = \frac{xy}{n(n-1)} - \frac{\min(x,y)}{n(n-1)}$$
$$\leq P(R_{ik} \leq x)P(R_{jk} \leq y) = \frac{xy}{n^2}, \quad \forall x, y \in R$$

و از آنجا بنا بر قضیهٔ ۱-۱،

$$(R_i, R_j) \in \mathcal{G}_1, \quad \forall i \neq j.$$

پس با فرض $\alpha = \sum_{i=1}^{m} \alpha_i$، خواهیم داشت:

$$\alpha - \frac{1}{2}(1 - \frac{1}{m})\alpha^2 \leq P = P_H[\bigcup_{j=1}^{m}(R_j \geq C_j)] \leq \alpha.$$

می‌توان کاربردهای دیگری از متغیرهای تصادفی دارای پیوند منفی را مورد بحث قرار داد که آن را به فرصتهای دیگر موکول می‌کنیم.

مراجع

[1] Bozorgnia, A. Patterson. R. F. and Tyloor. R. L. (1993) Limit Theorems for Negatively Dependent Random Variables. *University of Georgia Technical Report.*

[2] Esary. J. D. Proschan, F. and walkup, D. W. (1967) Association of r. v, s. with Applications. *Ann. Math. Stat.* 38, 144-147.

[3] Esary J. D. and Proschan, F. (1972) Relationships Among Some Concepts of Bivariate Dependence Positive. *Ann. Math. Stat* 43, 661-665.

[4] Lehmann, E. (1966). Some Concepts of Dependence. *Ann. Math. Stat* 37-1137-1153.