



# گشتاورهای متغیر همراه رکورد کلاسیک و کران‌های ناپارامتری میانگین تحت مدل فارلی-گامبل-مورگنشترن

مرتضی امینی\* و جعفر احمدی

دانشگاه فردوسی مشهد

چکیده. در یک دنباله از متغیرهای تصادفی جفت شده، زمانی که فقط بررسی دنباله‌ای از آماره‌های یکی از دو مؤلفه مدنظر باشد، مؤلفه‌ی دوم متناظر با هر آماره تحت عنوان متغیر همراه (concomitant) آن آماره مورد بررسی قرار می‌گیرد. از جمله‌ی این آماره‌ها می‌توان آماره‌های تربیی، رکوردها، سانسورها و غیره را نام برد. در این مقاله هدف بررسی خواص دنباله‌ی متغیرهای همراه رکوردها است. در یک دنباله از متغیرهای تصادفی، رکوردها مقادیری هستند که از مشاهدات قبلی بزرگ‌تر و یا کوچک‌تر هستند. بسیاری از پدیده‌های طبیعی، داده‌های مربوط به مسابقات ورزشی و یا آزمایه‌های بالینی در قالب این چنین الگوهای احتمال، قابل بررسی هستند. بدین منظور برای دنباله‌ای از جفت مشاهدات مستقل و هم‌توزیع با متغیرهای تصادفی  $(X, Y)$  خواص گشتاورهای متغیرهای تصادفی  $Y$  متناظر با رکوردهای متغیر تصادفی  $X$  در حالت کلی و تحت مدل فارلی-گامبل-مورگنشترن (Farlie-Gumbel-Morgenstern) مورد بررسی قرار گرفته است. در ادامه کران‌های ناپارامتری میانگین تحت مدل بالا و بهشیوه‌های متفاوت به دست آمده و با هم مقایسه شده‌اند.

واژگان کلیدی. رکورد بالا؛ متغیر همراه؛ گشتاور؛ مدل فارلی-گامبل-مورگنشترن؛ ناپارامتری کوشی-شوارتس؛ کران ناپارامتری.

\* نویسنده‌ی عهده‌دار مکاتبات.

## ۱ مقدمه

در بسیاری از موارد در دنباله‌ای از متغیرهای تصادفی حفت شده، فقط بررسی دنباله‌ای از آماره‌های یکی از دو مؤلفه مد نظر است. این مسئله می‌تواند به سبب عدم دسترسی آسان به مؤلفه دوم، اهمیت نخستین مؤلفه، یا پیچیدگی مدل‌های ریاضی تأم باشد. برای مثال، در گزینش‌های مرحله‌ای، افراد در مرحله‌ای اول بر اساس مهم‌ترین مؤلفه انتخاب می‌شوند. بنا بر این، مؤلفه دوم هر فرد انتخاب شده تابعی از گزینش مرحله‌ای اول است. مطالعه در خصوص چنین مسائلی گوشی از کاربردهای متغیر همراه آماره‌های تربیتی را نشان می‌دهد. از جمله کارهای انجام شده در این خصوص می‌توان به دیوید و ناگاراجا (۱۹۸۸) اشاره کرد. در خصوص رکوردها نیز با مسائلی مشابه رو به رو هستیم. برای مثال در یک مسابقه‌ی ورزشی وزنه‌برداری، رکوردهای وزنه‌برداران در رقابت دو ضرب ثبت شده است و ما علاقه‌مند به کسب اطلاعات بیش‌تر درباره‌ی مقدار رکوردهای ورزشکاران در رقابت یک ضرب هستیم. تا کنون مطالعه‌ی قابل توجهی در مورد متغیر همراه رکوردها انجام نشده است. هوچنز (۱۹۸۴) فقط نگاهی کوتاه به بعضی خواص متغیر همراه رکورد بالا دارد. وی همچنین به بررسی مثال‌هایی کوتاه در این خصوص پرداخته است. از جمله کارهای اخیر انجام شده در این مورد می‌توان به احسان ... (۱۹۹۴)، (۲۰۰۰)، خالدی و کوچار (۲۰۰۲)، و رکب و احسان ... (۲۰۰۲) اشاره کرد. برای ارائه‌ی تعاریف دقیق‌تر، نخست، آماره‌های رکورد را تعریف می‌کنیم.

در بسیاری از موارد در بررسی دنباله‌ای از پدیده‌های تصادفی، فقط مشاهداتی ثبت می‌شوند که از مقادیر پیشین خود بزرگ‌تر یا کوچک‌تر باشند. این دنباله از متغیرها به آماره‌ی رکورد معروف‌اند. بدین ترتیب که اگر مشاهده‌ی جدید از مشاهدات قبل از خودش بزرگ‌تر باشد آن را رکورد بالا و اگر کوچک‌تر باشد آن را رکورد پایین می‌نامیم. برای مثال می‌توان از ثبت بزرگ‌ترین زلزله در مقیاس ریشتر تا زمان حال، بیش‌ترین رکورد ثبت شده توسط یک ورزشکار، یا سرددترین روز یک ماه خاص از سال در یک شهر نام برد. این نظریه به دلیل اهمیت کاربردی در زمینه‌های مختلف به تازگی نظر سیاری از پژوهشگران را به خود جلب کرده است. همچنین، مقاله‌ها و کتاب‌های زیادی در این زمینه به چاپ رسیده است. در اینجا ما فقط به بیان تعریف‌های ابتدایی مربوط به رکوردهای بالا بسنده می‌کنیم. برای آگاهی بیش‌تر از این پژوهش‌ها می‌توان به آرنولد و همکاران (۱۹۹۸) مراجعه کرد. فرض کنید  $X_1, X_2, \dots, X_n$  دنباله‌ای نامتناهی از متغیرهای تصادفی باشد. زمان رخداد رکوردها با دنباله‌ی  $\{T_n\}_{n \geq 0}$  نشان داده شده و به صورت زیر تعریف می‌شود:

$$T_0 = 1, \quad \text{با احتمال یک}$$

و برای هر  $n \geq 1$

$$T_n = \min\{j : X_j > X_{T_{n-1}}\}.$$

در دنباله‌ی نامتناهی  $\{X_i\}$  از متغیرهای تصادفی، مشاهده‌ی  $X_j$  را یک رکورد بالا می‌گوییم اگر برای هر  $j < i$  داشته باشیم  $X_j < X_i$ . به عبارت دیگر با استفاده از تعریف ۱،

$$R_n = X_{T_n}, \quad n = 0, 1, 2, \dots$$

با فرض اینکه  $X_i$ ‌ها متغیرهای تصادفی پیوسته‌ی مستقل و هم‌توزیع با تابع چگالی احتمال  $f_X$  و تابع توزیع تجمعی  $F_X$  باشند، تابع چگالی رکورد  $n$ ام بالا به صورت زیر است:

$$(1) \quad f_{R_n}(x) = \frac{1}{n!} f_X(x) \{-\log(S_X(x))\}^n,$$

که در آن

$$S_X(x) = 1 - F_X(x).$$

همچنین چگالی توأم نخستین  $n+1$  رکورد بالا برابر است با:

$$(2) \quad f_{R_{n-1}, R_n}(x_0, \dots, x_n) = f_X(x_n) \prod_{i=0}^{n-1} \frac{f_X(x_i)}{S_X(x_i)},$$

و برای  $n < m$  تابع چگالی توأم دو رکورد به صورت زیر است:

$$(3) \quad f_{R_m, R_n}(x_0, x_1) = \frac{\{-\log(S_X(x_0))\}^m \left\{ -\log\left(\frac{S_X(x_1)}{S_X(x_0)}\right) \right\}^{n-m} f_X(x_0) f_X(x_1)}{m!(n-m-1)! S_X(x_0)}.$$

برای جزئیات اثبات (۱)، (۲)، و (۳) می‌توان به آرنولد و همکاران (۱۹۹۸) مراجعه کرد. دو تعریف ۱ و ۲ را می‌توان به صورت توأم نیز مطرح نمود. همچنین می‌توان تعریف متغیر همراه را در متن آن گنجاند.

تعریف ۳ به طور هم‌زمان مفاهیم زمان رکورد، دنباله‌ی رکوردها و متغیر همراه رکوردها را معرفی می‌کند.

فرض کنید  $(X_i, Y_i)$ ،  $i \geq 1$  دنباله‌ای از متغیرهای تصادفی جفت‌شده باشد. دنباله‌ی متغیرهای تصادفی جفت‌شده  $(R_i, R_{[i]})$ ،  $i \geq 0$  را به صورت زیر تعریف می‌کنیم:

$$(R_0, R_{[0]}) = (X_0, Y_0), \quad \text{با احتمال یک}$$

و برای هر  $n \geq 1$

$$(R_n, R_{[n]}) = (X_{T_n}, Y_{T_n}),$$

که در آن  $T_n$  زمان رخداد  $n$ امین رکورد در مؤلفه اول است.

در این صورت به  $R_n$  رکورد  $n$ ام بالا و به  $R_{[n]}$  متغیر همراه رکورد  $n$ ام بالا گفته می‌شود. در ادامه مقاله، منظور از متغیر همراه  $n$ ام همان متغیر همراه رکورد  $n$ ام بالا است. بخش ۲ شامل تابع چگالی توأم و حاشیه‌ای متغیر همراه و فرمول‌هایی برای گشتاورهای این متغیر تحت مدل فارلی-گامبل-مورگشنترن (FGM) است. در بخش ۳ کران‌های ناپارامتری برای میانگین متغیر همراه  $n$ ام تحت مدل FGM با فرضیات متفاوت به دست آمده و با هم مقایسه شده‌اند. مشابه این کران‌ها برای رکوردها و آماره‌های ترتیبی را می‌توان در بالاکریشنان (۱۹۹۰) و ناگاراجا (۱۹۷۸) ملاحظه کرد.

## ۲ متغیر همراه تحت مدل FGM

فرض کنید  $(X_1, Y_1), (X_2, Y_2), \dots$  دنباله‌ای مستقل و هم‌توزیع از متغیرهای تصادفی جفت‌شده‌ی پیوسته با تابع توزیع  $F(x, y)$  و تابع چگالی  $f(x, y)$  باشد. چگالی شرطی  $Y = x$  به شرط  $X = x$  را با  $f_{Y|X}(y|x)$  و چگالی‌های حاشیه‌ای را با  $f_X(x)$  و  $f_Y(y)$  نشان می‌دهیم. همچنین فرض کنید  $f_{R_k}(x) = f_{R_{r_1}, \dots, R_{r_k}}(x_1, \dots, x_k)$  نشان دهنده تابع چگالی احتمال توأم  $R_{r_1}, \dots, R_{r_k}$  باشد که در آن

$$1 \leq r_1 \leq \dots \leq r_k \leq n, \quad 1 \leq k \leq n.$$

در این صورت با توجه به استقلال زوج‌های  $(X_i, Y_i)$ ،  $i \geq 1$ ، همان‌طور که در یانگ (۱۹۷۷) بحث شده است داریم:

$$f_{R_{[r_1]}, \dots, R_{[r_k]} | R_{r_1}, \dots, R_{r_k}}(y_1, \dots, y_k | x_1, \dots, x_k) = \prod_{i=1}^k f_{Y|X}(y_i | x_i).$$

بنابراین،  $f_{R_{[K]}}(y) = f_{R_{[r_1]}, \dots, R_{[r_k]}}(y_1, \dots, y_k)$  تابع چگالی توأم  $k$  متغیر همراهه به صورت زیر است:

$$(4) \quad f_{R_{[k]}}(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{x_k} \dots \int_{-\infty}^{x_1} \prod_{i=1}^k f_{Y|X}(y_i | x_i) f_{R_k}(x) \prod_{i=1}^k dx_i$$

تحت مفروضات بالا،  $f_{R_{[n+1]}}(y) = f_{R_{[1]}, \dots, R_{[n]}}(y_1, \dots, y_n)$  تابع چگالی توان نخستین  $n + 1$  متغیر همراه رکوردهای بالا برابر است با:

$$f_{R_{[n+1]}}(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f_{X,Y}(x_n, y_n) \prod_{i=1}^{n-1} \frac{f_{X,Y}(x_i, y_i)}{S_X(x_i)} \prod_{i=1}^n dx_i.$$

بدین ترتیب، تابع چگالی حاشیه‌ای متغیر همراه  $n$  ام به صورت زیر است:

$$(5) \quad f_{R_{[n]}}(y) = \frac{1}{n!} \int_{-\infty}^{\infty} f_{Y|X}(y|x) f_X(x) \{-\log(S_X(x))\}^n dx.$$

رابطه‌ی (5) نشان‌دهنده‌ی یک رابطه بین تابع چگالی متغیر همراه و تابع چگالی شرطی  $Y$  به شرط مقادیر مختلف  $X$  است. در واقع با معلوم بودن توزیع متغیر تصادفی  $X$  و گشتاورهای شرطی  $Y$  به شرط مقادیر مختلف  $X$  و با استفاده از مطالب گفته شده، می‌توان فرمول‌هایی به صورت زیر برای محاسبه‌ی گشتاورهای متغیر همراه به دست آورد:

$$(6) \quad \begin{aligned} \alpha_{[n]}^{(k)} &= E\{E(Y^k | X = R_n)\} \\ \alpha_{[m,n]}^{(t,s)} &= E\{E(Y^t | X = R_m) E(Y^s | X = R_n)\}. \end{aligned}$$

که در آن  $\alpha_{[n]}^{(k)}$  و  $\alpha_{[m,n]}^{(t,s)}$  به ترتیب گشتاور مرتبه‌ی  $k$  ام و گشتاور حاصل‌ضرب متغیر همراه هستند. برای جزئیات بیش‌تر در خصوص اثبات موارد (4)، (5) و (6) می‌توان به رکب و احسان ا... (۲۰۰۲) یا خالدی و کوچار (۱۹۹۰) مراجعه کرد.

حال خواص متغیر همراه و گشتاورهای آن را تحت مدل FGM بررسی می‌کنیم. مدل FGM در حقیقت مدلی از چگالی توان دو متغیر تصادفی بر حسب چگالی حاشیه‌ای آن دو متغیر است و به صورت زیر تعریف می‌شود:

برای یک بردار تصادفی  $(X, Y)$  با تابع چگالی حاشیه‌ای  $f_Y(y)$  و  $f_X(x)$ ، تابع چگالی توان مدل FGM به صورت زیر است:

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) \{1 + \alpha(1 - 2F_X(x))(1 - 2F_Y(y))\},$$

که در آن  $\alpha \in (-1, 1)$ .

حال تابع چگالی  $R_{[n]}$  را به دست می‌آوریم.

قضیه‌ی ۱ تحت مدل FGM تابع چگالی احتمال متغیر همراه  $n$  ام به صورت زیر است:

$$(7) \quad f_{R_{[n]}}(y) = f_Y(y) \{1 + \alpha(1 - 2^{-n})(2F_Y(y) - 1)\}.$$

برهان. با استفاده از (۵) و تحت مدل FGM داده شده درتابع چگالی متغیر همراه رکورد  $n$  بالا برابر است با:

$$\begin{aligned} f_{R_{[n]}}(y) &= \int_{-\infty}^{\infty} f_{Y|X}(y|x) f_X(x) \frac{\{-\log(S_X(x))\}^n}{n!} dx \\ &= \int_{-\infty}^{\infty} f_{X,Y}(x,y) \frac{\{-\log(S_X(x))\}^n}{n!} dx \\ &= \frac{1}{n!} \int_{-\infty}^{\infty} f_X(x) f_Y(y) \{1 + \alpha(1 - 2F_X(x))(1 - 2F_Y(y))\} \\ &\quad \times \{-\log(S_X(x))\}^n dx. \end{aligned}$$

با تغییر متغیر  $(z = -\log(S_X(x)))$  مقدار انتگرال بالا برابر است با:

$$f_Y(y) \left\{ 1 + \frac{\alpha}{n!} (2F_Y(y) - 1) \int_0^{\infty} (e^{-z} - 2e^{-z}) z^n dz \right\},$$

که با محاسبه های جبری لازم برهان کامل می شود.

چگالی داده شده در (۷) یک رابطه‌ی کاملاً مستقل از رکوردها را برای متغیرهای همراه آن‌ها فراهم می‌کند. بدین ترتیب می‌توان گشته اورهای مربوط به این متغیر را فقط با داشتن اطلاعاتی از توزیع متغیر تصادفی  $Y$  بررسی کرد. قضیه‌ی زیر رابطه‌ای بین گشته اور مرتبه‌ی  $k$  ام متغیر همراه و دو نوع گشته اور مربوط به توزیع  $Y$  را بیان می‌کند.

**قضیه‌ی ۲** تحت مدل FGM، گشته اور  $k$  ام متغیر همراه  $n$  ام را می‌توان به صورت زیر نوشت:

$$\alpha_{[n]}^{(k)} = \alpha^{(k)} + \alpha(1 - 2^{-n})(\alpha_{\frac{n}{2}; \frac{n}{2}}^{(k)} - \alpha^{(k)}),$$

که در آن  $\alpha^{(k)}$  گشته اور  $k$  ام توزیع  $Y$  و  $\alpha_{\frac{n}{2}; \frac{n}{2}}^{(k)}$  گشته اور  $k$  بزرگ‌ترین آماره‌ی مرتب در یک نمونه‌ی دوتایی از این توزیع است.

برهان. با استفاده از (۷) و تابع چگالی احتمال دومین آماره‌ی ترتیبی در یک نمونه‌ی دوتایی، داریم:

$$\begin{aligned} \alpha_{[n]}^{(k)} &= \int_{-\infty}^{\infty} y^k f_{R_{[n]}}(y) dy \\ &= \alpha^{(k)} + \alpha(1 - 2^{-n}) \left( \int_{-\infty}^{\infty} 2y^k f_Y(y) F_Y(y) dy - \alpha^{(k)} \right) \\ &= \alpha^{(k)} + \alpha(1 - 2^{-n})(\alpha_{\frac{n}{2}; \frac{n}{2}}^{(k)} - \alpha^{(k)}). \end{aligned}$$

**نتیجه‌ی ۱** با توجه به قضیه‌ی ۳ میانگین و واریانس متغیر همراه  $n$  ام برابر است با:

$$\alpha_{[n]} = \mu + \alpha(1 - 2^{-n})(\alpha_{2:2} - \mu),$$

و

$$\begin{aligned}\sigma_{[n]}^2 &= \sigma^2 + \alpha(1 - 2^{-n})(\mu^2 - \sigma^2 + \alpha_{2:2}^{(2)} - 2\mu\alpha_{2:2}) \\ &\quad - \alpha^2(1 - 2^{-n})^2(\alpha_{2:2} - \mu)^2,\end{aligned}$$

که در آن‌ها  $\mu$  و  $\sigma^2$  میانگین و واریانس متغیر تصادفی  $Y$  هستند.

**مثال ۱** فرض کنید  $Y_i$  ها دارای توزیع نرمال استاندارد باشند. در این صورت می‌توان نشان داد  $\alpha_{2:2}^{(2)} = \alpha_{2:2}^{(1)} = \alpha_{2:2} = 1$  و  $\alpha_{2:2} = \frac{\alpha}{\sqrt{\pi}}$ . بنا بر این، میانگین و واریانس متغیر همراه  $n$  ام به ترتیب برابرند با:

$$\alpha_{[n]} = \alpha(1 - 2^{-n}) \quad \alpha_{2:2} = \frac{\alpha}{\sqrt{\pi}}(1 - 2^{-n}),$$

و

$$\sigma_{[n]}^2 = 1 - \frac{\alpha^2}{\pi}(1 - 2^{-n})^2.$$

اما برای به دست آوردن گشتوهای حاصل ضرب به توزیع توان دو متغیر همراه نیاز داریم. این تابع چگالی توان در قضیه‌ی ۳ به دست آمده است. اما قبل از بیان قضیه نیاز به محاسبه‌ی یک انتگرال در لم زیر داریم.

**لم ۱** فرض کنید برای  $m < n$

$$I(a, b) = \int_{-\infty}^{\infty} \int_{x_1}^{\infty} F_X(x_1)^a F_X(x_2)^b f_{R_m, R_n}(x_1, x_2) dx_2 dx_1,$$

که در آن  $1, a = 0, b = 0$  و در این صورت داریم:

$$I(a, b) = \begin{cases} 1, & a = 0, b = 0 \\ (1 - 2^{-n-1}), & a = 0, b = 1 \\ (1 - 2^{-m-1}), & a = 1, b = 0 \\ 1 - 2^{-m-1} - 2^{-n-1} + 2^{-m-1}2^{m-n}, & a = 1, b = 1 \end{cases}$$

برهان. با تغییر متغیر  $z = -\log \left( \frac{S_X(x_2)}{S_X(x_1)} \right)$  برای انتگرال داخلی و تغییر متغیر  $w = -\log(S_X(x_1))$  برای انتگرال بیرونی مقدار انتگرال به راحتی محاسبه می‌شود.

قضیه ۳ تحت مفروضات مورد بحث، برای  $n < m$  چگالی توأم دو متغیر همراه برابر است با:

$$f_{R_{[m]}, R_{[n]}}(y_1, y_2) = f_Y(y_1) f_Y(y_2) \left\{ 1 + \alpha(2F_Y(y_1) - 1)(1 - 2^{-m}) \right. \\ \left. + \alpha(2F_Y(y_2) - 1)(1 - 2^{-n}) + \alpha^2(2F_Y(y_1) - 1) \right. \\ \left. \times (2F_Y(y_2) - 1)(3^{-m-1}2^{m+2-n} - 2^{-m} - 2^{-n} + 1) \right\}.$$

برهان. با استفاده از (۴) می‌توان نوشت:

$$f_{R_{[m]}, R_{[n]}}(y_1, y_2) = \int_{-\infty}^{\infty} \int_{x_1}^{\infty} f_{Y|X}(y_1|x_1) f_{Y|X}(y_2|x_2) f_{R_m, R_n}(x_1, x_2) dx_2 dx_1.$$

بنا بر این، تحت مدل FGM داریم:

$$f_{R_{[m]}, R_{[n]}}(y_1, y_2) = f_Y(y_1) f_Y(y_2) \int_{-\infty}^{\infty} \int_{x_1}^{\infty} g(\mathbf{x}, \mathbf{y}) f_{R_m, R_n}(x_1, x_2) dx_2 dx_1,$$

که در آن

$$g(\mathbf{x}, \mathbf{y}) = \prod_{i=1,2} [1 + \alpha(2F_X(x_i) - 1)(2F_Y(y_i) - 1)].$$

پس با استفاده از لم ۱ اثبات کامل می‌شود.

با استفاده از قضیه ۳ نتایج زیر را داریم.

نتیجه ۲ تحت مدل FGM، برای  $t$  و  $s$  دلخواه و  $m < n$  داریم:

$$(A) \quad \alpha_{[m,n]}^{(t,s)} = \alpha^{(s)} \alpha^{(t)} + \alpha(1 - 2^{-n}) \alpha^{(t)} (\alpha_{[1,2]}^{(s)} - \alpha^{(s)}) \\ + \alpha(1 - 2^{-m}) \alpha^{(s)} (\alpha_{[1,2]}^{(t)} - \alpha^{(t)}) + (\alpha_{[1,2]}^{(t)} - \alpha^{(t)}) (\alpha_{[1,2]}^{(s)} - \alpha^{(s)}) \\ \times \alpha^2 (3^{-m-1}2^{m-n+2} - 2^{-m} - 2^{-n} + 1).$$

نتیجه ۳ تحت مدل FGM، کواریانس دو متغیر همراه برای  $n < m$  برابر است با:

$$(B) \quad \text{Cov}(R_{[m]}, R_{[n]}) = (3^{-m-1}2^{m-n+2} - 2^{-m-n}) \alpha^2 (\alpha_{[1,2]} - \mu)^2.$$

برهان. با جایگذاری  $s = t$  در نتیجه‌ی قبل داریم:

$$\begin{aligned} E(R_{[m]}R_{[n]}) &= \mu^{\frac{1}{2}} + (2 - 2^{-m} - 2^{-n}) \alpha \mu (\alpha_{2:2} - \mu) \\ &\quad + (3^{-m-1} 2^{m-n+2} - 2^{-m} - 2^{-n} + 1) \alpha^{\frac{1}{2}} (\alpha_{2:2} - \mu)^{\frac{1}{2}}. \end{aligned}$$

همچنین با استفاده از نتیجه‌ی ۱

$$\begin{aligned} E(R_{[m]})E(R_{[n]}) &= \mu^{\frac{1}{2}} + (2 - 2^{-m} - 2^{-n}) \alpha \mu (\alpha_{2:2} - \mu) \\ &\quad + (1 - 2^{-n})(1 - 2^{-m}) \alpha^{\frac{1}{2}} (\alpha_{2:2} - \mu)^{\frac{1}{2}}. \end{aligned}$$

بنا بر این (۶) حاصل می‌شود.

تذکر ۱ نتیجه‌ی ۳ را می‌توان با استفاده از (۶) و با توجه به این که در مدل FGM داریم:

$$E(Y|X=x) = \mu + \alpha(2F_X(x) - 1)(\alpha_{2:2} - \mu),$$

به دست آورده.

مثال ۲ (ادامه‌ی مثال ۱) بنا بر (۶) کوواریانس دو متغیر همراه  $m$  و  $n$  برای  $n < m$  برابر است با:

$$\text{Cov}(R_{[m]}, R_{[n]}) = \frac{\alpha^{\frac{1}{2}}}{\pi} (3^{-m-1} 2^{m-n+2} - 2^{-m-n}),$$

با توجه به واریانس به دست آمده در مثال ۱ نمی‌توان ضریب همبستگی را به صورت مضربی از  $\alpha^{\frac{1}{2}}$  نوشت. بنا بر این، مقادیر ضریب همبستگی به منظور مقایسه به ازای  $\frac{1}{\sqrt{\alpha}}$  برای بعضی مقادیر مختلف  $m$  و  $n$  در جدول ۱ داده شده است.

جدول ۱. جدول مقادیر ضریب همبستگی دو متغیر همراه در توزیع نرمال استاندارد

$m$				$n$		
۱۰	۵	۳	۱	۲	۴	۶
				۰,۰۰۰۷۹۹		
				۰,۰۰۱۴۴	۰,۰۰۲۰۲	
				۰,۰۰۰۱۹	۰,۰۰۰۳۶	۰,۰۰۰۵۱
۰,۰۰۰۰۰	۰,۰۰۰۰۰	۰,۰۰۰۰۱	۰,۰۰۰۰۲		۱۱	
۰,۰۰۰۰۰	۰,۰۰۰۰۰	۰,۰۰۰۰۰	۰,۰۰۰۰۰	۰,۰۰۰۰۰		۱۵

### ۳ کران‌های ناپارامتری میانگین تحت مدل FGM

در مواردی که توزیع دقیق  $Y_i$  ها معلوم نباشد با استفاده از شیوه‌های مختلف، می‌توان یک کران بالا برای میانگین  $R_{[n]}$  یافت. یکی از متدالول ترین این شیوه‌ها استفاده از نابرابری کوشی-شوارتس است. در این بخش با استفاده از این نابرابری و تحت شرایط متفاوت، کران‌های ناپارامتری بالای متفاوت برای میانگین متغیر همراه به دست می‌آید.

#### ۳.۱ کران عمومی

برای به دست آوردن کران‌های ناپارامتری میانگین، بدون کاستن از کلیت می‌توان فرض کرد که  $E(Y_i) = 0$ ،  $Var(Y_i) = 1$  می‌توان نوشت:

$$(10) \quad \alpha_{[n]} = \alpha(1 - 2^{-n}) \alpha_{\frac{1}{2}, \frac{1}{2}},$$

قضیه‌ی زیر ساده‌ترین کران ناپارامتری را با استفاده از نابرابری کوشی-شوارتس ارائه می‌دهد.

قضیه‌ی ۴ یک کران ناپارامتری برای میانگین  $R_{[n]}$  به صورت زیر است:

$$(11) \quad \alpha_{[n]} \leq \frac{\alpha}{\sqrt{3}}(1 - 2^{-n}).$$

برهان. با تغییر متغیر  $y = F_Y(u)$  در  $\alpha_{\frac{1}{2}, \frac{1}{2}} = \int_{-\infty}^{\infty} y f_Y(y) F_Y(y) dy$  داریم:

$$\begin{aligned} \alpha_{[n]} &= \alpha(1 - 2^{-n}) \alpha_{\frac{1}{2}, \frac{1}{2}} \\ &= 2\alpha(1 - 2^{-n}) \int_0^1 F_Y^{-1}(u) u du \\ &= \int_0^1 F_Y^{-1}(u) \{2\alpha(1 - 2^{-n})u - \lambda\} du. \end{aligned}$$

با استفاده از نابرابری کوشی-شوارتس می‌توان نوشت:

$$\alpha_{[n]} \leq \left[ \int_0^1 \{2\alpha(1 - 2^{-n})u - \lambda\}^2 du \right]^{\frac{1}{2}}.$$

که پس از ساده کردن سمت راست نابرابری بالا، عبارت زیر نتیجه می‌شود:

$$\alpha_{[n]} \leq \left\{ \frac{4}{3}\alpha^2(1 - 2^{-n})^2 - 2\alpha(1 - 2^{-n})\lambda + \lambda^2 \right\}^{\frac{1}{2}}.$$

سمت راست نابرابری بالا به ازای  $\lambda = \alpha(1 - 2^{-n})$  مینیمیم می‌شود، در نتیجه اثبات کامل است.

همین طور که ملاحظه می‌کنید توزیع نرمال استاندارد (مثال ۱) در این قاعده صدق می‌کند.

**نتیجه‌ی ۴** با توجه به شرط برقراری تساوی در نابرابری کوشی-شووارتس، کران بالا زمانی به دست می‌آید که  $Y_i$ ‌ها از جامعه‌ای با تابع توزیع تجمعی معکوس زیر به دست آمده باشند:

$$\begin{aligned} F_Y^{-1}(u) &= \frac{1}{\frac{\alpha}{\sqrt{3}}(1 - 2^{-n})} \{1 + \alpha(1 - 2^{-n})(2u - 1) - 1\} \\ &= \sqrt{3}(2u - 1). \end{aligned}$$

بنا بر این، تابع توزیع تجمعی به صورت زیر خواهد بود:

$$F_Y(y) = \frac{1 + \sqrt{3}y}{2}.$$

در نتیجه، کران بالا زمانی به دست می‌آید که  $Y_i$ ‌ها دارای توزیع یکنواخت در بازه‌ی  $(-\sqrt{3}, \sqrt{3})$  باشند.

**نتیجه‌ی ۵** زمانی که اندازه‌ی نمونه افزایش یابد، حد کران (۱۱) به صورت زیر به دست می‌آید:

$$\lim_{n \rightarrow \infty} \frac{\alpha}{\sqrt{3}}(1 - 2^{-n}) = \frac{\alpha}{\sqrt{3}}.$$

به عبارت دیگر

$$\lim_{n \rightarrow \infty} \alpha_{[n]} \leqslant \frac{\alpha}{\sqrt{3}}.$$

با توجه به این که  $(1 - 2^{-n})$  نسبت به  $n$  صعودی است می‌توان کران بالا را به عنوان یک کران عمومی در نظر گرفت. همچنین با حدگیری از (۷) و اجرای مراحل برهان قضیه‌ی ۴، می‌توان این کران را برای توزیع حدی زیر به دست آورد:

$$\lim_{n \rightarrow \infty} f_{R_{[n]}}(y) = f_Y(y)\{1 + \alpha(2F_Y(y) - 1)\}.$$

**تذکر ۲** با استفاده از نتایج مشابه برای  $\alpha_{2:2}$  می‌توان به طور غیر مستقیم، قضیه‌ی ۴ و نتیجه‌ی ۴ را به دست آورد. زیرا همان‌طور که هارتی و دیوید (۱۹۵۴) و گامبل (۱۹۵۴) و موریگوتوی (۱۹۵۱) نشان داده‌اند:

$$\alpha_{n:n} \leqslant \frac{n - 1}{(2n - 1)^{\frac{1}{r}}}.$$

برابری زمانی به دست می‌آید که نمونه‌ی تصادفی از توزیع زیر تولید شده باشد:

$$F(y) = \left[ \frac{1}{n} \left\{ \frac{n-1}{(2n-1)^{\frac{1}{n}}} x + 1 \right\} \right]^{\frac{1}{n-1}}, \quad -\frac{(2n-1)^{\frac{1}{n}}}{n-1} \leq x \leq (2n-1)^{\frac{1}{n}}.$$

که با جایگذاری مقدار  $n = 2$  قضیه‌ی ۴ و نتیجه‌ی ۴ به دست خواهد آمد.

### ۳،۲ کران برای توزیع‌های متقارن

معمولًاً وقتی که توزیع جامعه متقارن باشد، کران‌های ناپaramتری بهبود پیدا می‌کنند، اما این مسئله برای متغیر همراه در مدل FGM برقرار نیست. این موضوع در قضیه‌ی زیر اثبات می‌شود.

قضیه‌ی ۵ در حالتی که توزیع  $Y_i$  ها متقارن و استاندارد باشد، نابرابری کوشی-شوارتس کران (۱۱) را بهبود نمی‌دهد.

برهان. با توجه به تقارن توزیع می‌توان نوشت:

$$\alpha_{[n]} = \alpha(1 - 2^{-n}) \int_0^1 F_Y^{-1}(u) \{u - (1-u) - \lambda\} du.$$

بنا بر نابرابری کوشی-شوارتس داریم:

$$\alpha_{[n]} \leq \alpha(1 - 2^{-n}) \left( \frac{1}{3} + \lambda^2 \right)^{\frac{1}{2}},$$

که سمت راست نابرابری به ازای  $\lambda = 0$  مینیمم می‌شود. بدین ترتیب، حکم ثابت می‌شود.

### ۳،۳ بهبود کران‌ها با دو مقدار میانگین معلوم

در این بخش، حالتی را در نظر می‌گیریم که دو میانگین متغیرهای همراه مثل  $\alpha_{[m]}$  و  $\alpha_{[p]}$  ( $m < p$ ) معلوم باشند. کران‌هایی که در این حالت به دست می‌آیند در اصطلاح کران‌های برونویسی (extrapolation-type bounds) نامیده می‌شوند.

قضیه‌ی ۶ وقتی دو میانگین متغیرهای همراه مثل  $\alpha_{[m]}$  و  $\alpha_{[p]}$  ( $m < p$ ) معلوم باشند

$$\alpha_{[n]} \leq (L(n, n) - A)^{\frac{1}{2}},$$

که در آن

$$L(m, n) = \frac{\alpha^r}{r} (1 - 2^{-n})(1 - 2^{-m}),$$

$$A = \frac{\left( L(m, n) - \frac{\alpha[m]}{\alpha[p]} L(p, n) \right)^r}{L(m, n) - r \frac{\alpha[m]}{\alpha[p]} L(m, p) + \left( \frac{\alpha[m]}{\alpha[p]} \right)^r L(p, p)}.$$

همچنین، کران بالا نسبت به (۱۱) بهبود یافته است، مگر در حالت  
برهان. برای آسان‌تر شدن کار، تابع کمکی زیر را تعریف می‌کنیم:

$$g_l(u) = 1 + \alpha(1 - 2^{-l})(2u - 1).$$

پس می‌توان نوشت:

$$\begin{aligned} \alpha[n] &= \int_0^1 F_Y^{-1}(u) [g_n(u) - c\{g_m(u) - \alpha[m]F_Y^{-1}(u)\}] \\ &\quad - d\{g_p(u) - \alpha[p]F_Y^{-1}(u)\} - \lambda \, du. \end{aligned}$$

به راحتی می‌توان دید که

$$\int_0^1 g_n(u) \, du = 1.$$

به علاوه، برای خلاصه‌سازی نتایج قرار می‌دهیم:

$$K(m, n) = \int_0^1 g_n(u) g_m(u) \, du = 1 + \frac{\alpha^r}{r} (1 - 2^{-n})(1 - 2^{-m}).$$

بدین ترتیب با اعمال نابرابری کوشی-شوارتس بر (۱۳) داریم:

$$\begin{aligned} \alpha[n] &\leq \left\{ K(n, n) + c^r (K(m, m) - \alpha[m]^r) + d^r (K(p, p) - \alpha[p]^r) + \lambda^r \right. \\ &\quad - r c (K(m, n) - \alpha[m]\alpha[n]) - r d (K(p, n) - \alpha[p]\alpha[n]) - r \lambda (1 - c - d) \\ &\quad \left. + r c d (K(m, p) - \alpha[m]\alpha[p]) \right\}^{\frac{1}{r}}. \end{aligned}$$

با انتقال مقادیر نامعلوم به طرف چپ، نابرابری به صورت زیر ساده می‌شود:

$$\alpha_{[n]}(1 - 2c\alpha_{[m]} - 2d\alpha_{[p]}) \leq \left\{ K(n, n) + c^{\frac{1}{r}}(K(m, m) - \alpha_{[m]}^{\frac{1}{r}}) + d^{\frac{1}{r}}(K(p, p) - \alpha_{[p]}^{\frac{1}{r}}) + \lambda^{\frac{1}{r}} - 2cK(m, n) - 2dK(p, n) - 2\lambda(1 - c - d) + 2cd(K(m, p) - \alpha_{[m]}\alpha_{[p]}) \right\}^{\frac{1}{r}}.$$

$d$  را طوری انتخاب می‌کنیم که ضریب  $\alpha_{[n]}$  در سمت چپ نابرابر یک شود، یعنی قرار می‌دهیم  $d = -\frac{\alpha_{[m]}}{\alpha_{[p]}}c$ . همچنین، مقدار بهینه‌ی  $\lambda$  یعنی  $\lambda = \left\{ 1 - c\left(1 - \frac{\alpha_{[m]}}{\alpha_{[p]}}\right) \right\}^{\frac{1}{r}}$  را جایگزین می‌کنیم. بدین ترتیب مقدار بهینه‌ی  $c$  برابر است با:

$$c_{\text{opt}} = \frac{L(m, n) - \frac{\alpha_{[m]}}{\alpha_{[p]}}L(p, n)}{L(m, n) - 2\frac{\alpha_{[m]}}{\alpha_{[p]}}L(m, p) + \left(\frac{\alpha_{[m]}}{\alpha_{[p]}}\right)^{\frac{1}{r}}L(p, p)},$$

که در آن

$$L(m, n) = K(m, n) - 1 = \frac{\alpha^{\frac{1}{r}}}{3}(1 - 2^{-n})(1 - 2^{-m}).$$

با جایگذاری مقادیر بهینه در سمت راست نابرابری حکم ثابت می‌شود. بهبود این کران نسبت به (۱۱) با توجه به این که  $A \geq 0$  نتیجه می‌شود. اما در حالت  $A = 0$  و کران نسبت به (۱۱) تغییری نمی‌کند.

برای بهبود دادن بیشتر حالت بالا می‌توان توزیع‌های متقارن استاندارد با دو مقدار میانگین معلوم را در نظر گرفت.

قضیه‌ی ۷ با توجه به فرضیات قضیه‌ی ۶ برای توزیع‌های متقارن استاندارد داریم:

$$\alpha_{[n]} \leq (L^*(n, n) - A^*)^{\frac{1}{r}},$$

که در آن

$$L^*(i, j) = L(i, j) - 1,$$

$$A^* = \frac{\left( L^*(m, n) - \frac{\alpha_{[m]}}{\alpha_{[p]}} L^*(p, n) \right)^r}{L^*(m, n) - r \frac{\alpha_{[m]}}{\alpha_{[p]}} L^*(m, p) + \left( \frac{\alpha_{[m]}}{\alpha_{[p]}} \right)^r L^*(p, p)}.$$

برهان. در این حالت می‌توان نوشت:

$$\begin{aligned} \alpha_{[n]} &= \frac{1}{\gamma} \int_0^\gamma F_Y^{-1}(u) \left[ g_n(u) - g_n(\gamma - u) - c \{g_m(u) - g_m(\gamma - u) \right. \\ &\quad \left. - r \alpha_{[m]} F_Y^{-1}(u)\} - d \{g_p(u) - g_p(\gamma - u) - r \alpha_{[p]} F_Y^{-1}(u)\} - \lambda \right] du. \end{aligned}$$

بنا بر این،

$$\begin{aligned} \alpha_{[n]} &\leq \frac{1}{\gamma} \left\{ r K(n, n) - r \int_0^\gamma g_n(u) g_n(\gamma - u) du \right. \\ &\quad + c^r \left( r K(m, m) - r \alpha_{[m]}^r - r \int_0^\gamma g_m(u) g_m(\gamma - u) du \right) \\ &\quad + d^r \left( r K(p, p) - r \alpha_{[p]}^r - r \int_0^\gamma g_p(u) g_p(\gamma - u) du \right) \\ &\quad - r c \left( r K(m, n) - r \int_0^\gamma g_n(u) g_m(\gamma - u) du - r \alpha_{[m]} \alpha_{[n]} \right) \\ &\quad - r d \left( r K(p, n) - r \int_0^\gamma g_n(u) g_p(\gamma - u) du - r \alpha_{[p]} \alpha_{[n]} \right) \\ &\quad + r c d \left( r K(m, p) - r \int_0^\gamma g_m(u) g_p(\gamma - u) du - r \alpha_{[m]} \alpha_{[p]} \right) \\ &\quad \left. - r \lambda (\gamma - c - d) + \lambda^r \right\}^{\frac{1}{r}}. \end{aligned}$$

اما برای خلاصه نمودن طرف راست نابرابری بالا، توجه کنید که

$$\begin{aligned} \int_0^1 g_i(u)g_j(1-u)du &= \int_0^1 \left\{ 1 - \alpha(1-2^{-i})(2u-1) \right\} \\ &\quad \times \left\{ 1 - \alpha(1-2^{-j})(2u-1) \right\} du \\ &= \int_0^1 1 - \alpha^2(1-2^{-i})(1-2^{-j})(2u-1)^2 du \\ &= 1 - \frac{\alpha^2}{4}(1-2^{-i})(1-2^{-j}). \end{aligned}$$

بدین ترتیب با توجه به این که

$$K(m, n) - \int_0^1 g_n(u)g_m(1-u)du = 2(K(m, n) - 1) = 2L(m, n),$$

نابرابری بالا به صورت زیر خلاصه می‌شود:

$$\begin{aligned} \alpha_{[n]}(1 - 2c\alpha_{[m]} - 2d\alpha_{[p]}) &\leq \left\{ L(n, n) + c^2(L(m, m) - \alpha_{[m]}^2) \right. \\ &\quad + d^2(L(p, p) - \alpha_{[p]}^2) - 2cL(m, n) \\ &\quad - 2dL(p, n) + 2cd(L(m, p) - \alpha_{[m]}\alpha_{[p]}) \\ &\quad \left. - \lambda(1 - c - d) + \frac{1}{4}\lambda^2 \right\}^{\frac{1}{2}}. \end{aligned}$$

درست مانند برهان قضیه ۶ با در نظر گرفتن  $d = -\frac{\alpha_{[m]}}{\alpha_{[p]}}c$  و مقدار بهینه  $\lambda = 2\left\{ 1 - c\left(1 - \frac{\alpha_{[m]}}{\alpha_{[p]}}\right) \right\}$  یعنی مقدار بهینه  $c$  به صورت زیر به دست می‌آید:

$$c_{\text{opt}} = \frac{L^*(m, n) - \frac{\alpha_{[m]}}{\alpha_{[p]}}L^*(p, n)}{L^*(m, n) - 2\frac{\alpha_{[m]}}{\alpha_{[p]}}L^*(m, p) + \left(\frac{\alpha_{[m]}}{\alpha_{[p]}}\right)^2 L^*(p, p)},$$

بدین ترتیب نابرابری به صورت مورد نظر خلاصه می‌شود.

## ۴ نتیجه‌گیری

برابری (۱۰) به ما اجازه می‌دهد تا خواص میانگین متغیر همراه رکوردهای بالا در مدل FGM را بدون نیاز به دانستن توزیع  $X$  و مستقیماً نسبت به خواص میانگین آماره‌ی ترتیبی دوم در یک نمونه‌ی دوتایی از توزیع  $Y$  بررسی کنیم. بدین ترتیب، می‌توان در این مدل خاص برآوردهای  $\alpha_{[n]}$  را بر حسب برآوردهای  $\alpha_{[2:2]}$  به راحتی محاسبه کرد. همچنین فواصل اطمینان، آزمون‌های فرض و خواص حدی مربوط به پارامترهای بالا به راحتی قابل مقایسه هستند. همچنین، می‌توان روش‌های بیشتری را برای به دست آوردن کران‌های تقریبی در این مدل بررسی کرد. نویسنده‌گان ترجیح می‌دهند در مقاله‌های بعدی مسائل بالا را مورد بررسی قرار دهند. البته باید توجه داشت که گرچه رده‌ی بالا می‌تواند شامل طیفی وسیع از توزیع‌های توان باشد، اما فقط بخش بسیار کوچکی از کل است و همچنان مسئله‌ی اصلی، یعنی بررسی متغیر همراه در یک مدل کلی ترو یا مدل‌های بیشتر و کارامدتر به عنوان هدف نهایی مورد توجه خواهد بود.

## ۵ قدردانی

نویسنده‌گان از داوران محترم به خاطر پیشنهادات سازنده‌ی ایشان در تهیه‌ی نسخه‌ی نهایی این مقاله قدردانی می‌نمایند.

### مرجع‌ها

- Arnold, B.C.; Balakrishnan, N.; Nagaraja, H.N. (1998). *Records*. Wiley, New York.
- Ahsanullah, M. (1994). Record values, random record models and concomitants, *J. Statist. Res.* **28**, 89-109.
- Ahsanullah, M. (2000). Concomitants of record values. *Pakistan J. Statist.* **16**, 207-215.
- Balakrishnan, N. (1990). Improving the Hartley-David-Gumbel bound for the mean of extreme order statistics, *Statis. Probab. Lett.* **9**, 291-294.
- David, H.A. (1981). *Order Statistics. Second Edition*, Wiley, New York.
- David, H.A.; Nagaraja H.N. (1988). Concomitants of order statistics, In *Handbook of Statist.* **16**, pp. 487-513.
- Gumbel, E.J. (1954). The maxima of the mean largest value and of the range, *Ann. Math. Statist.* **25**, 76-84.

- Hartley, H.O.; David, H.A. (1954). Universal bounds for mean rang and extreme observation. *Ann. Math. Statist.* **25**, 85-99.
- Houchens, R.L. (1984). Record Value Theory and Inference. *Ph.D. thesis*, University of California, Riverside, CA.
- Khaledi, B; Kocher, S.C. (2001). On dependence structure of multivariate mixture distributions. *Ann. Inst. Stat. Math.* **53**, 620-630.
- Moriguti, S. (1951). External properties of extreme value distributions. *Ann. Math. Statist.* **22**, 523-536.
- Moriguti, S. (1954). Bounds for second moments of the sample range. *Rep. Stat. Appl. Res. JUSE* **2**, 99-103.
- Nagaraja, H.N. (1978). On the expected values of record values. *Aust. J. Statist.*, **20**, 176-182.
- Raqab, M.Z. ; Ahsanullah, M. (2002). Concomitants of ordered random variables – A review, *J. Stat. Theory App.* **1**, 15-26.
- Yang, S.S. (1977). General distribution theory of the concomitants of order statistics. *Ann. Statist.* **5**, 996-1002.

دريافت:	۱۷ خرداد ۱۳۸۴
آخرین اصلاح:	۱۲ شهریور ۱۳۸۴

<b>جعفر احمدی</b> گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، مشهد، ایران. <i>ahmadi@math.um.ac.ir</i>	<b>مرتضی امینی</b> گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد، مشهد، ایران. <i>mort_amini@hotmail.com</i>
--	---