نشریه علوم دانشگاه تربیت معلم
جلد۱۰، شماره ۱
تاریخ چاپ: بهار ۱۳۹۱

مدل‌بندی درست‌نمایی و بیز تجربی خشک‌سالی با استفاده از تابع مفصل

مهدی امیدی، *محسن محمدزاده: دانشگاه تربیت مدرس، گروه آمار

چکیده

یکی از مراحل اساسی بررسی و مدیریت خشک‌سالی، شناخت و تحلیل بسامدی ویژگی‌های آن از جمله شدت و مدت خشک‌سالی است. با توجه به هم‌بستگی بالای این دو عامل باید از ابزاری استفاده شود که میزان ارتباط و تأثیراتی را که در تحلیل خشک‌سالی دارند، نمایان سازد. توابع مفصل ساختار وابستگی بین متغیرها را به‌صورت یك مدل نشان می‌دهند و علاوه بر این‌ها می‌توان ضرایب هم‌بستگی بین متغیرها را هم به‌دست آورد. در این مقاله خانواده‌های مناسب برای مدل‌بندی پدیده خشک‌سالی ارائه می‌شوند، سپس با روش‌های ماکسیمم درست‌نمایی و بیز تجربی پارامتر تابع مفصل برآورد می‌گردد. آن‌گاه برترین خانواده توابع مفصل برای به‌دست آوردن توزیع توأم مدت و شدت خشک‌سالی در ایستگاه هواشناسی تهران تعیین می‌گردد و با استفاده از آن پدیده خشک‌سالی در تهران را بر اساس داده‌های مدت و شدت خشک‌سالی در دورهٔ ۳۷ سالهٔ ۱۳۴۸ تا ۱۳۸۴ مدل‌بندی شده و نحوهٔ کاربرد آن در مدیریت ایمنی ذخایر آب مطرح می‌شود.

مقدمه

خشک‌سالی وضعیتی از کمبود بارندگی و افزایش دماست که در هر وضعیت اقلیمی ممکن است رخ دهد. پژوهش بر روی این پدیده با توجه به تأثیر منفی آن بر محیط زندگی و فعالیت انسان در زمینه‌های مختلف، همواره از موضوعات مهم علمی بوده است. چون پدیده خشک‌سالی طبیعتی تصادفی دارد [۱۷]، تعداد زیادی از محققان مشخصات احتمالی خشک‌سالی را بررسی کردند [۳]، [۷]، [۱۲]. اما در این بررسی‌ها عمدتاً به تحلیل تك متغیرهٔ مشخصات خشک‌سالی پرداخته شده است، در حالی‌که این پدیده ذاتاً با چندین متغیر تصادفی وابسته مشخص می‌شود. چون تحلیل چند متغیرهٔ پدیدهٔ خشک‌سالی نیاز به داده‌های زیاد و محاسبات ریاضی پیچیده دارد، در تحقیقات کمی، آن هم فقط با استفاده از توزیع‌های دو متغیره به آن پرداخته شده است. عده‌ای از محققان توزیع‌های دو متغیره را با فرض آن‌که توزیع‌های کناری آن‌ها متعلق به خانواده‌ای از توزیع‌ها هستند در مسائل مربوط به‌سیل به‌کار برده‌اند [۲]، [۱۶]، [۲۰]. تعدادی از محققان هم حاصل‌ضرب توزیع شرطی شدت خشک‌سالی برای مقدار داده شدهٔ مدت خشک‌سالی و توزیع کناری مدت خشک‌سالی را برای به‌دست آوردن

واژه‌های کلیدی: تابع مفصل، مدت و شدت خشک‌سالی، برآورد بیز تجربی.
دریافت ۸۸/۵/۲۱ پذیرش۹۰/۶/۱۶
*نویسنده مسئول

توزیع توأم مدت و شدت خشک‌سالی به‌کار گرفته‌اند [۷]، [۱۲]، [۱۵]. شیائو تعیین توزیع توأم دو متغیرهٔ مدت و شدت خشک‌سالی را به‌کمك توابع مفصل به‌دست آورد [۱۷]. مؤلفان با استفاده از مدل‌بندی درست‌نمایی و طیف وسیع‌تری از توابع مفصل به تحلیل خشک‌سالی در استان تهران پرداخته‌اند [۱].

تحلیل رفتار و ویژگی‌های دو متغیر تصادفی از طریق توزیع توأم آن‌ها میسر است. اما در عمل گاهی با مواردی مواجه می‌باشیم که توزیع توأم متغیرهای تصادفی نامعلوم است. چنان‌که دو متغیر تصادفی مستقل باشند توزیع توأم به‌صورت حاصل‌ضرب توزیع‌های کناری حاصل می‌شود. در صورت عدم استقلال متغیرهای تصادفی تعیین توزیع توأم سرراست نیست و برای تحلیل رفتار دو متغیر لازم است به نحوی توزیع توأم مشخص گردد. یکی از روش‌های تعیین توزیع توأم دو متغیر تصادفی استفاده از تابع مفصل است. به‌طورکلی توابع مفصل ابزاری قوی برای ساخت توابع توزیع چند متغیره بر اساس حاشیه‌ای‌های یك بعدی هستند که نوع و چگونگی ارتباط بین متغیرها را نیز نشان می‌دهند. اسکلار برای اولین بار توابع مفصل را در قضیه مربوط به توابعی که توزیع‌های یك متغیره را به توزیع چند متغیرهٔ آن‌ها پیوند می‌دهد مطرح کرده است [۱۸]. [۱٤] خصوصیات وابستگی توابع مفصل را بررسی کردده و [۹] از توابع مفصل به‌عنوان توابعی از توزیع یك‌نواخت و [٤] و[٦] با عنوان توابع وابسته استفاده کردند. [٤] و [۸] خانواده‌های توابع مفصل که در علوم مختلف کاربرد دارند را بررسی کردند. [۱۱] روش‌های مهم ساخت توابع مفصل و طیف وسیع‌تری از خانواده‌های آن‌ها را که محققان در زمینه‌های مختلف به‌کار گرفته‌اند، را جمع‌آوری کرده است. توابع مفصل می‌توانند ساختار وابستگی بین متغیرها را به‌صورت یك مدل نشان دهند و علاوه بر آن بستری فراهم نمایند که بر اساس آن می‌توان تمامی معیارهای ارتباط بین متغیرها را بررسی کرده و همچنین اندازه‌های هم‌بستگی را به‌دست آورد. در این مقاله رایج‌ترین توابع مفصل مناسب برای مدل‌بندی خشک‌سالی انتخاب شده و پارامتر آن‌ها برای داده‌های مدت و شدت خشک‌سالی ایستگاه تهران با دو روش ماکسیمم درست‌نمایی و بیز تجربی برآورد می‌گردد. سپس برای هر یك از این روش‌ها مناسب‌ترین تابع مفصل بر اساس معیار اطلاع بیزی[۱]

$$BIC = -2\log L(x, y; \theta) + k\log(n)$$

انتخاب می‌شود، که در آن k تعداد پارامترهای مدل است [۱۳]. از آن‌جا که برآورد ماکسیمم درست‌نمایی تابع مفصل همواره مقدار BIC کم‌تری نسبت به برآورد بیز تجربی دارد، برای انتخاب برترین تابع مفصل بین برآوردگرهای ماکسیمم درست‌نمایی و بیز تجربی از معیار جذر میانگین توان دوم خطاها[۲] به صورت

$$RMSD = [\frac{1}{n}\sum_{i=1}^{n}(F(d_i, s_i) - F_n(d_i, s_i))^2]^{\frac{1}{2}}$$

استفاده می‌شود، که در آن $F(d_i, s_i)$ توزیع توأم حاصل از به‌کارگیری هر یك از توابع مفصل و $F_n(d_i, s_i)$ تابع توزیع توأم تجربی داده‌ها است. در انتها کاربرد مدل حاصل در مدیریت ایمنی ذخایر آب ارائه می‌گردد.

۱. Bayesian Information Criterion ۲. Root of Mean Squared Errors

توابع مفصل

دو متغیر تصادفی پیوسته X و Y به‌ترتیب با توابع توزیع $F(x)$ و $G(y)$ و تابع توزیع توأم $H(x,y)$ را در نظر بگیرید، برای هر جفت عدد حقیقی (x,y) می‌توان سه مقدار $(F(x),G(y),H(x,y))$ را در نظر گرفت که هر کدام به يك نقطهٔ $(F(x),G(y))$ در مربع واحد $[0،1]\times[0،1]$ و این زوج نیز با مقدار $H(x,y)$ در بازهٔ $[0،1]$ مطابقت دارد، رابطه‌ای که هر مقدار تابع توزیع توأم $H(x,y)$ را به توابع توزیع يك‌بعدی $F(x)$ و $G(y)$ مرتبط می‌سازد، تابع مفصل نامیده می‌شود. برای بیان فرم ریاضی تابع مفصل، [۱۸] نشان داد اگر H تابع توزیع توأم دو متغیر تصادفی X و Y با توابع توزیع حاشیه‌ای F و G باشند، توابع مفصلی مانند C وجود دارند که برای هر x و y در $\bar{R}=R\cup\{-\infty,+\infty\}$ داریم

$$H(x,y)=C(F(x),G(y);\theta) \qquad (۱)$$

اگر F و G مطلقا پیوسته باشند، تابع مفصل C یکتاست، در غیر این صورت C به صورت يك تابع غیریکتا بر روی $R_F\times R_G$ تعریف می‌شود، که در آن R_F و R_G به‌ترتیب دامنهٔ F و G است. اگر $f(x)$ و $g(y)$ توابع چگالی X و Y باشند، آن‌گاه تابع چگالی توأم X و Y به صورت

$$f(x,y)=c(F(x),G(y);\theta)f(x)g(y) \qquad (۲)$$

است، که در آن c تابع مفصل تابع چگالی توأم است و از رابطهٔ $c(u,v)=\partial^2 C(u,v)/\partial u\partial v$ به‌دست می‌آید.

در این بخش توابع مفصل و تابع مفصل چگالی مربوط به هر کدام که با توجه به ساختار وابستگی آن‌ها برای مدل‌بندی خشکسالی مناسب هستند، معرفی می‌شوند.

۱. خانوادهٔ علی-میخاییل-حق:

$$C(u,v;\theta)=\frac{uv}{1-\theta(1-u)(1-v)},\theta\in[-1,1]$$

$$c(u,v;\theta)=\frac{[1-\theta(1-u)(1-v)](1-\theta)+2\theta uv}{[1-\theta(1-u)(1-v)]^3}$$

۲. خانوادهٔ پلاکت:

$$C(u,v;\theta)=\frac{1+(\theta-1)(u+v)-\{[1+(\theta-1)(u+v)]^2-4uv\theta(\theta-1)\}^{\frac{1}{2}}}{2(\theta-1)},\theta\geq 0$$

$$c(u,v;\theta)=[(1+(\theta-1)(u+v))^2-4uv\theta(\theta-1)\}^{\frac{-3}{2}}\theta[1+(\theta-1)(u+v)-2uv]$$

۳. خانوادهٔ کلایتون:

$$C(u,v;\theta)=(u^{-\theta}+v^{-\theta}-1)^{\frac{-1}{\theta}} \quad ,\theta\geq 0$$

$$c(u,v;\theta)=(\theta+1)(u^{-\theta}+v^{-\theta}-1)^{\frac{-1}{\theta}-2}(uv)^{-\theta-1}$$

۴. خانوادهٔ فرانك:

$$C(u,v;\theta)=\frac{-1}{\theta}\ln[1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}]\ \ ,\theta\neq 0$$

$$c(u,v;\theta)=\frac{\theta e^{-\theta(u+v)}(e^{-\theta}-1)}{[e^{-\theta(u+v)}-e^{-\theta u}-e^{-\theta v}-e^{-\theta}]^{2}}$$

۵. خانوادهٔ گالامبوس:

$$C(u,v;\theta)=uv\exp\{[(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{\frac{-1}{\theta}}\},\theta\geq 0$$

$$c(u,v;\theta)=\frac{C(u,v;\theta)}{uv}\{1-[(-\ln u)^{-\theta-1}+(-\ln v)^{-\theta-1}][(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{\frac{-1}{\theta}-1}$$
$$+[(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{\frac{-1}{\theta}-2}[(-\ln u)(-\ln v)]^{-\theta-1}\times[1+\theta+[(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{\frac{-1}{\theta}}]\}$$

۶. خانوادهٔ گامبل-بارنت:

$$C(u,v;\theta)=uv\exp(-\theta\ln u\ln v)\ ,\theta\in[0,1)$$

$$c(u,v;\theta)=\frac{C(u,v;\theta)}{uv}[1-\theta(\ln(u)+\ln(u)+1-\theta\ln(u)\ln(v))]$$

۷. خانوادهٔ گامبل-هوگارد:

$$C(u,v;\theta)=\exp\{-[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{\frac{1}{\theta}}\}\ ,\theta\geq 1$$

$$c(u,v;\theta)=C(u,v;\theta)\frac{[(-\ln u)(-\ln v)]^{\theta-1}}{uv}[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{\frac{2}{\theta}-2}\times$$
$$\{(\theta-1)[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{\frac{1}{\theta}}+1\}$$

۸. خانوادهٔ جو:

$$C(u,v;\theta)=1-[(1-u)^{\theta}+(1-v)^{\theta}-((1-u)(1-v))^{\theta}]^{\frac{1}{\theta}}\quad,\theta\geq 1$$

$$c(u,v;\theta)=[(1-u)^{\theta}+(1-v)^{\theta}-((1-u)(1-v))^{\theta}]^{\frac{1}{\theta}-2}[(1-u)(1-v)]^{\theta-1}$$
$$[(\theta+1)+(1-u)^{\theta}+(1-v)^{\theta}-((1-u)(1-v))^{\theta}]$$

خانواده‌های توابع مفصل بر اساس نوع وابستگی بین دو متغیر در تحلیل خانوادهٔ توزیع‌های توأم به‌کار برده می‌شود، از آن جمله می‌توان به مفصل گالامبوس در تحلیل توزیع توأم مقادیر کرانگین[۱] و مفصل گامبل بارنت‌در برای توزیع نمایی دومتغیره در تحلیل بقا[۲] اشاره کرد. فرض استقلال متغیرها که در اغلب مسائل و مدل‌های آماری درنظر گرفته می‌شود، موجب راحتی محاسبات نظری می‌شود، اما يك پديده ممكن است به‌وسیلهٔ عامل‌های مختلفی تفسیر شود که تحت شرایط محیطی دارای وابستگی قوی باشند، چنان‌که تحلیل جداگانه هر کدام از این عوامل بر روی پدیدهٔ خاص فاقد ارزش و اعتبار باشد. به‌کمك توابع مفصل می‌توان ساختار وابستگی بین متغیرها

۱. Exterm Value ۲. Survival Copula

را به‌صورت يك مدل نشان داد و بر اساس آن اندازه هم‌بستگی بین متغیرها را به‌دست آورد. از جمله اندازه‌های هم‌بستگی که می‌توان با استفاده از توابع مفصل تعیین کرد [۱۰]، ضرایب هم‌بستگی پیرسون، کندال و اسپیرمن به‌ترتیب بدین‌صورت هستند:

$$r_C = \int_0^1 \int_0^1 [C(u,v) - uv] dF^{-1}(u) dG^{-1}(v) / \sqrt{Var(X) Var(Y)}$$

$$\tau_C = 4 \int_0^1 \int_0^1 C(u,v) dC(u,v) - 1$$

$$\rho_C = 12 \int_0^1 \int_0^1 [C(u,v) - uv] du dv$$

که در ادامه از آن‌ها استفاده می‌شود.

برآورد پارامترهای تابع مفصل

فرض کنید $F(x)$ و $G(y)$ به‌ترتیب توزیع‌های کناری X و Y و $f(x)$ و $g(y)$ توابع چگالی متناظر آن‌ها باشند. بر اساس مشاهدات $(x_1, y_1), \ldots, (x_n, y_n)$ تابع درستنمایی به‌صورت $L(x,y) = \prod_{i=1}^{n} f(x_i, y_i)$ است. برای برآورد درستنمایی ماکسیمم پارامتر تابع مفصل و بر اساس رابطهٔ (۲.۲) این رابطه به‌صورت

$$\prod_{i=1}^{n} f(x_i, y_i) = \prod_{i=1}^{n} [c((F(x_i), G(y_i); \theta) f(x_i) g(y_i)]$$

به‌دست می‌آید. با گرفتن لگاریتم طبیعی از این رابطه لگاریتم تابع درستنمایی به‌صورت

$$\begin{aligned} LnL((x,y);\theta) &= Ln[\prod_{i=1}^{n} c((F(x_i), G(y_i); \theta) f(x_i) g(y_i)] \\ &= \sum_{i=1}^{n} Ln[c((F(x_i), G(y_i); \theta)) f(x_i) g(y_i)] \qquad (۳) \\ &= \sum_{i=1}^{n} [Lnl_c((F(x_i), G(y_i); \theta)) + Lnf(x_i) + Lng(y_i)] \end{aligned}$$

به دست می‌آید، که در آن $Lnl_c(\cdot)$ لگاریتم طبیعی تابع مفصل چگالی است.

برآورد ماکسیمم درستنمایی پارامتر θ بر اساس ماکسیمم کردن رابطهٔ (۳) مستلزم معلوم بودن توابع چگالی کناری دو متغیر تصادفی X و Y است. چنان‌چه این توابع نامعلوم باشند می‌توان برآورد آن‌ها را در رابطهٔ (۳) جای‌گذاری کرد. یکی از مشکلات روش ماکسیمم درستنمایی در صورت معلوم بودن توزیع توأم متغیرهای تصادفی، در نظر نگرفتن عدم حتمیت پارامترها و کم‌برآورد نمودن واریانس برآوردگرها است. برای لحاظ کردن عدم حتمیت پارامترها می‌توان از رهیافت بیزی استفاده کرد، که در آن پارامترهای نامعلوم به‌عنوان تحقق متغیرهایی تصادفی در نظر گرفته می‌شوند و از يك توزیع پیشینی معلوم پیروی می‌کنند.

فرض کنید θ، پارامتر تابع مفصل تحققی از يك متغیر تصادفی با توزیع احتمال پیشینی معلوم $\pi(\theta)$ باشد. اگر

$f(x,y\,|\,\theta)$ تابع چگالی توأم دو متغیر تصادفی (X,Y) باشد، آن‌گاه توزیع پسینی θ

$$\pi(\theta\,|\,x,y)=\frac{f(x,y\,|\,\theta)}{\int_\Theta f(x,y\,|\,\theta)d\theta},\quad \theta\in\Theta$$

و برآورد بیزی آن با تابع زیان درجه دوم میانگین توزیع پسینی، یعنی $\theta^B=E(\theta\,|\,x,y)$ است. چون با در نظر گرفتن تابع مفصل چگالی $c(\cdot,\cdot;\theta)$، تابع چگالی توأم به‌صورت

$$f(x,y;\theta)=\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)f(x_i)g(y_i)$$

است، برآوردگر بیزی θ از رابطهٔ

$$\theta^B=\frac{\int_\Theta \theta\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)f(x_i)g(y_i)\pi(\theta)d\theta}{\int_\Theta\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)f(x_i)g(y_i)\pi(\theta)d\theta}\tag{٤}$$

$$=\frac{\int_\Theta \theta\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)\pi(\theta)d\theta}{\int_\Theta\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)\pi(\theta)d\theta}$$

به‌دست می‌آید. همان‌طور که ملاحظه می‌شود انتگرال‌های صورت و مخرج رابطهٔ (٤) بر حسب آن‌که تابع مفصل چگالی $c(\cdot,\cdot;\theta)$ چه فرمی داشته باشد بسیار پیچیده هستند و در صورت وجود حل تحلیلی آن‌ها معمولاً مقدور نیست. لذا برای به دست آوردن برآورد بیزی θ می‌توان روش انتگرال‌گیری مونت کارلو را به‌کار گرفت. برای این منظور رابطهٔ (٤) را می‌توان به‌صورت

$$\theta^B=\frac{E(\theta\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta))}{E(\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta))}\tag{٥}$$

نوشت، که در آن امید ریاضی‌های صورت و مخرج بر حسب توزیع پیشینی $\pi(\theta)$ هستند. اگر m مقدار تصادفی $\theta_1,\dots,\theta_m$ از توزیع $\pi(\theta)$ تولید شده باشند، آن‌گاه برای عبارت صورت کسر رابطهٔ (٥) طبق قانون قوی اعداد بزرگ داریم:

$$\frac{1}{m}\sum_{j=1}^{m}\theta_j\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta_j))\xrightarrow{as}E(\theta\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)),\quad m\to\infty$$

و مشابهاً برای مخرج کسر نیز داریم:

$$\frac{1}{m}\sum_{j=1}^{m}\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta_j))\xrightarrow{as}E(\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)),\quad m\to\infty$$

با توجه به خاصیت همگرایی $a.s$ برآوردگر رابطهٔ (٥) به‌صورت

$$\hat{\theta}_m{}^B = \frac{\frac{1}{m}\sum_{j=1}^{m}\theta_j\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta_j))}{\frac{1}{m}\sum_{j=1}^{m}\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta_j))} \qquad (٦)$$

پیشنهاد می‌گردد. واضح است که این تابع برای مقادیر بزرگ m به (٥) میل می‌کند. بنا بر این برای محاسبه $\hat{\theta}_m{}^B$، از توزیع پیشینی $\pi(\theta)$، m مقدار $\theta_1,...,\theta_m$ را تولید کرده و با جای‌گذاری در رابطه‌ٔ (٦) برآورد بیزی پارامتر θ به‌دست می‌آید. چنان‌چه تولید مقادیر تصادفی از $\pi(\theta)$ دشوار باشد می‌توان از روش نمونه‌برداری از نقاط مهم[١] اقدام به برآورد صورت و مخرج کسر (٥) کرد[١٩].

از آن‌جا که در عمل توزیع پیشینی $\pi(\theta)$ نامعلوم است، در رهیافت بیز تجربی برآورد این توزیع در روابط مربوط به برآورد بیزی جای‌گزین می‌گردد. در مواردی که فرم پارامتری توزیع پیشینی به‌صورت $\pi(\theta\,|\,\lambda)$ باشد، که در آن λ مقداری ثابت و نامعلوم است، ابتدا بر اساس توزیع کناری

$$m(x,y\,|\,\lambda) = \int_{\Theta} f(x,y\,|\,\theta)\pi(\theta\,|\,\lambda)d\theta$$

مقدار λ برآورد می‌شود. سپس برآورد بیز تجربی با جای‌گذاری برآورد λ در توزیع پیشینی به‌صورت $\hat{\pi}(\theta)=\pi(\theta\,|\,\hat{\lambda})$ یا در روابط مربوط به برآورد بیزی حاصل می‌شود. بنا بر این با توجه به رابطهٔ (٤) برآورد بیز تجربی پارامتر تابع مفصل به‌صورت زیر تعیین می‌گردد:

$$\theta^{EB} = \frac{\int_{\Theta}\theta\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)\hat{\pi}(\theta)d\theta}{\int_{\Theta}\prod_{i=1}^{n}c((F(x_i),G(y_i);\theta)\hat{\pi}(\theta)d\theta} \qquad (٧)$$

مشابه برآورد بیزی در این‌جا می‌توان با تولید مقادیر تصادفی از توزیع $\hat{\pi}(\theta)$ انتگرال‌های صورت و مخرج کسر (٧) را به روش مونت کارلو یا نمونه‌برداری از نقاط مهم به‌صورت تقریبی محاسبه کرد.

توزیع توأم مدت و شدت خشکسالی

در این بخش پدیدهٔ خشکسالی در تهران بر حسب دو عامل مدت و شدت خشکسالی با استفاده از توابع مفصل بر اساس داده‌های مدت خشکسالی بر حسب ماه (D) و شاخص نسبی شدت خشکسالی (S) در دورهٔ ٣٧ سالهٔ ١٣٤٨ تا ١٣٨٤ گزارش شده در ایستگاه هواشناسی تهران مدل‌بندی می‌شوند. چون ضرایب هم‌بستگی $r_C=$ ٠/٩، $\tau_C=$ ٠/٧٨ و $\rho_C=$ ٠/٩١ بیان‌گر هم‌بستگی قوی بین دو عامل مدت و شدت خشکسالی است، تحلیل جداگانهٔ این دو عامل برای خشکسالی نمی‌تواند میزان ارتباط و تأثیراتی را که این دو عامل در تحلیل خشکسالی بر هم دارند نمایان سازد. بنا بر این لازم است به‌کمك توابع مفصل توزیع توأم آن‌ها را به‌دست آورد.

١. Importance Sampling

برای این منظور خانوادهٔ توابع مفصل معرفی شده در بخش ۲ را برای مدل‌بندی خشکسالی در نظر می‌گیریم و پارامتر هر یك از این خانواده‌ها را به روش‌های ماکسیمم درست‌نمایی و بیز تجربی برآورد می‌کنیم. لذا ابتدا بر اساس [۲۱] و با توجه به این‌که مدت خشکسالی بر اساس مدت زمان انتظار تا پایان خشکسالی تعریف شده به آن توزیع‌نمایی E (۳/۰۵۲۱) و برای داده‌های شدت خشکسالی با توجه به بررسی‌های [۱٤] و [۲۱] توزیع گاما (۳/۲٦٤۱ و ۰/۷۲۸٥) Γ برازش داده شده است. با توجه به p- مقدارهای ۰/۲٦٤٥ و ۰/۸۷۳٤ حاصل از آزمون‌های نکویی برازش، توزیع‌های نمایی و گاما به‌ترتیب برای مدت و شدت خشکسالی در سطح ۰/۰٥ قابل پذیرش هستند. برای هر کدام از خانواده‌های معرفی شده در بخش ۲ تابع مفصل تابع چگالی توأم و برآورد توزیع‌های نمایی و گاما را در (۳) قرار داده و آن‌را بر حسب θ ماکسیمم کرده و برآورد درست‌نمایی θ به‌دست آمده است.

برای برآورد بیزی θ، لازم است توزیع پیشینی مناسب برای θ در نظر گرفته شود. با توجه به مثبت بودن ضرایب هم‌بستگی بین داده‌های مدت و شدت خشکسالی که مستلزم مثبت بودن θ است، از آن‌جا که دامنهٔ θ نامحدود و مقادیر آن مثبت است، توزیع پیشینی آن نمایی در فاصلهٔ $(0,\infty)$ در نظر گرفته شده است. اما به‌دلیل نامعلوم بودن پارامتر توزیع پیشینی مقدار آن برای هر تابع مفصل طوری انتخاب شده است که میانگین پیشینی بر برآورد ماکسیمم درست‌نمایی θ منطبق شود. با این توضیح که برای خانوادهٔ جو و گامبل-هوگارد چون مقدار پارمتر این توابع در بازهٔ (۱،∞] است، توزیع نمایی دو پارامتری با پارامتر مکان۱ در نظر گرفته شده است. برای توابع مفصل با دامنهٔ محدود، توزیع یك‌نواخت بر بازهٔ [۰،۱] درنظر گرفته‌ایم. با تولید ۱۰۰۰ مقدار $\theta_1,...,\theta_{1000}$ از توزیع‌های پیشینی و جای‌گذاری در رابطهٔ (٦) برآورد بیز تجربی پارامتر θ برای هر تابع مفصل به‌دست آمده است. برای داده‌های ایستگاه تهران برآورد ماکسیمم درست‌نمایی، برآورد توزیع پیشینی، برآورد بیز تجربی پارامتر مفصل و مقدار معیار BIC برای هر یك از خانواده‌های توابع مفصل در جدول ۱ ارائه شده‌اند.

جدول ۱. برآورد ماکسیمم درست‌نمایی، بیز تجربی و معیار BIC

ردیف	تابع مفصل	برآورد ماکسیمم درست‌نمایی	معیار BIC	توزیع پیشین برآورد شده	برآورد بیز تجربی	معیار BIC
۱	علی- میخاییل- حق	۰/۹۴۵۵	۴۷۲/۶۶۲۶	U(۰،۱)	۰/۹۷۳۴	۴۷۲/۹۲۷۱
۲	کلایتون	۱/۵۳۹۱	۴۷۱/۲۴۷۷	E(۱/۵۳۹۱)	۱/۶۴۲۴	۴۷۱/۳۳۶۹
۳	فرانک	۱۰/۶۹۴۰	۴۲۱/۲۸۰۷	E(۱۰/۶۹۴۰)	۱۱/۲۰۴۰	۴۲۱/۴۰۷۵
۴	گالامبوس	۱/۹۷۹۲	۴۲۳/۳۷۲۱	E(۱/۹۷۹۲)	۲/۰۸۴۰	۴۲۳/۵۰۶۶
۵	گامبل-بارنت	۰/۳۳۸۸	۴۷۶/۳۰۴۴	U(۰،۱)	۰/۲۸۱۴	۴۷۶/۷۸۷۳
۶	گامبل-هوگارد	۲/۶۸۷۳	۲۳/۲۸۳۳	E(۱،۲/۶۸۷۳)	۲/۸۴۵۰	۲۳/۵۷۰۹
۷	پلاکت	۰/۷۸۳۸	۴۹۸/۷۲۴۹	E(۰/۷۸۳۸)	۰/۸۴۲۳	۴۹۸/۹۸۰۱
۸	جو	۳/۴۶۳۵	۴۲۵/۱۷۸۵	E(۱،۳/۴۶۳۵)	۳/۶۵۸۵	۴۲۵/۳۷۰۹

چنان‌که ملاحظه می‌شود تابع مفصل فرانك، با برآورد ماکسیمم درست‌نمایی $\hat{\theta}^{ML} =$ ۱۰/٦۹٤۰ و برآورد بیز تجربی $\hat{\theta}^{EB} =$ ۱۱/۲۰٤۰ برای هر دو روش برآورد دارای کم‌ترین BIC هستند. مقدار RMSE برای برآورد ماکسیمم درست‌نمایی و بیز تجربی به ترتیب ۰/۰۳۸ و ۰/۰۳٦ به‌دست آمد، که نشان می‌دهد تابع مفصل فرانك با

برآورد بیز تجربی دارای RMSE کم‌تر است. بنا بر این این تابع مفصل با برآورد بیز تجربی $\hat{\theta}^{EB} =$ ۱۱/۲۰۴۰ به‌عنوان تابع مفصل برتر برای مدل‌بندی خشکسالی در ایستگاه تهران برگزیده می‌شود که بر اساس آن تابع توزیع توأم مدت و شدت خشکسالی به‌صورت زیر حاصل می‌گردد:

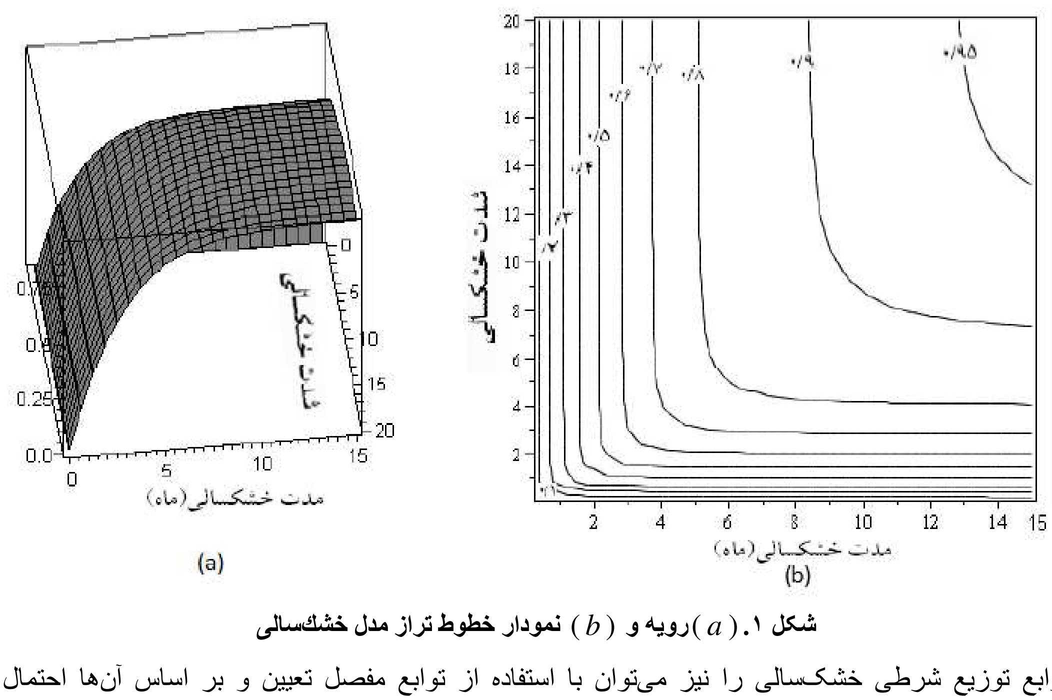
$$F(d,s)=\frac{-1}{11.2040}\ln\{1+\frac{(e^{-11.2040\int_0^d \frac{1}{3.0521}e^{\frac{-y}{3.0521}}dy}-1)(e^{-11.2040\int_0^s \frac{y^{-0.2715}}{3.2641^{0.2715}\Gamma(0.7285)}e^{\frac{-y}{3.2641}}dy}-1)}{e^{-11.2040}-1}\} \quad (۸)$$

بررسی توزیع توأم خشکسالی

با استفاده مدل (۸) می‌توان احتمال وقوع خشکسالی با مدت و شدت بیش‌تر از مقدار یك آستانه را به‌دست آورد، که برای مدیریت ذخیرهٔ منابع آب در شرایط بحرانی بسیار مفید است. این احتمال را می‌توان با استفاده از تابع مفصل و توابع توزیع حاشیه‌ای و جای‌گذاری در رابطهٔ

$$P(D\geq d,S\geq s)=1-F_D(d)-G_S(s)+C(F_D(d),G_S(s)) \quad (۹)$$

محاسبه کرد. به‌عنوان مثال در ایستگاه تهران یك خشکسالی با مدت بیش‌تر از ۳ ماه و شدت بیش‌تر از آستانه ۲ بر اساس رابطهٔ (۹) با احتمال ۰/۳۲۳۷ رخ می‌دهد. شکل۱ نشانگر رویهٔ و نمودار خطوط تراز مدل خشکسالی در ایستگاه تهران است، که بر اساس تابع مفصل فرانك به‌دست آمده است.



شکل ۱. (a) رویه و (b) نمودار خطوط تراز مدل خشك‌سالی

توابع توزیع شرطی خشکسالی را نیز می‌توان با استفاده از توابع مفصل تعیین و بر اساس آن‌ها احتمال چگونگی تغییر یك عامل در قبال تغییرات کنترل شده عامل دیگر را بررسی کرد. توزیع شرطی شدت

خشکسالی برای مدت بیش‌تر از آستانه d به‌صورت

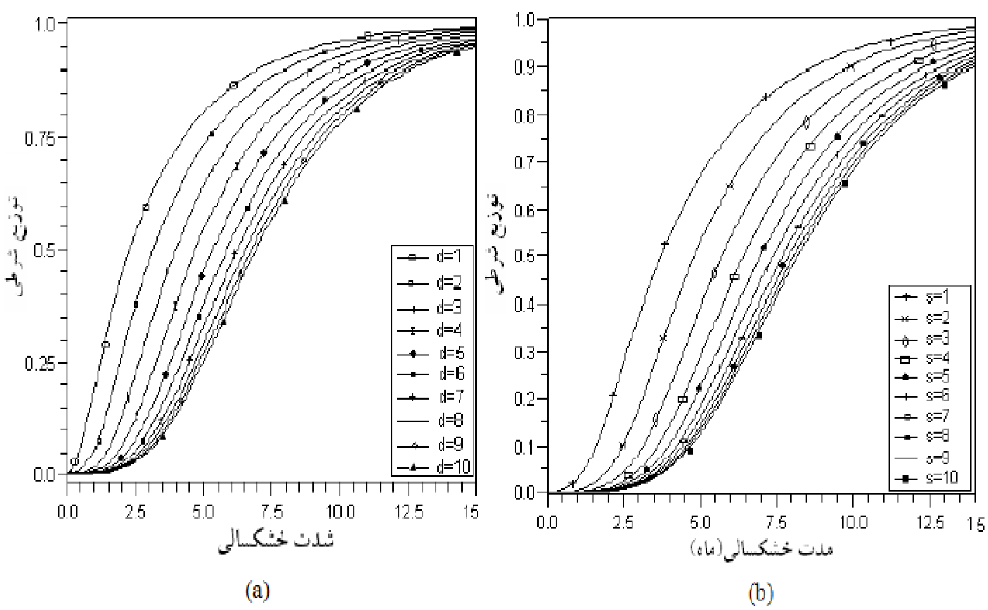
$$P(S \leq s \mid D \geq d) = \frac{G_S(s) - C(F_D(d), G_S(s))}{1 - F_D(d)}$$

و توزیع شرطی مدت خشکسالی برای شدت بیش‌تر از آستانه s به‌صورت

$$P(D \leq d \mid S \geq s) = \frac{F_D(d) - C(F_D(d), G_S(s))}{1 - G_S(s)}$$

است، که نمودار آن‌ها در شکل ۲ برای داده‌های ایستگاه تهران نشان داده شده‌است. با محاسبهٔ این احتمالات شرطی می‌توان ایمنی لازم برای ذخایر آب در زمان‌های معین را فراهم کرد.

تمامی محاسبات مربوط به برآوردها، شبیه‌سازی و رسم نمودارها با استفاده از نرم‌افزار MAPLE انجام شده است.



شکل ۲. (a) **نمودار توزیع شرطی شدت خشکسالی و** (b) **نمودار توزیع شرطی مدت خشکسالی**

بحث و نتیجه‌گیری

توابع مفصل را می‌توان برای تعیین ساختار وابستگی متغیرهای تصادفی هم‌بسته به‌کار گرفت. نتایجی که تاکنون به‌دست آمد، بیانگر آن استکه توابع مفصل ابزار مفیدی برای تعیین توزیع توأم و ساختار وابستگی متغیرهای هم‌بسته خشکسالی هستند. هر چند در این پژوهش توزیع‌های دو متغیره و توابع مفصل يك پارامتری بررسی شدند، اما به‌سادگی می‌توان تکنیك‌های مطرح شده را به توزیع‌های چندمتغیره تعمیم داد و مسئلهٔ خشکسالی را با متغیرهای بیش‌تر مانند شدت-مدت-سطح نیز بررسی کرد. توابع مفصلی که در این مقاله

مطالعه و استفاده شدند دارای يك پارامتر بودند. نحوهٔ تعريف خانوادهٔ توابع مفصل با پارامترهای بیش‌تر نیازمند انجام بررسی‌های بیش‌تر است. هر چند تعمیم روش ارائه شده برای برآورد بیزی و سلسله مراتبی پارامتر تابع مفصل برای خانواده‌های توابع مفصل چند پارامتری از نظر تئوری به‌سادگی امکان‌پذیر است، همواره با مسئلهٔ انتخاب توزیع پیشینی پارامتر یا ابرپارامترها مواجه است.

قدردانی و تشکر

نویسندگان از داوران محترم مجله که نظرات ارزنده آن‌ها موجب بهبود مقاله گردید و از حمایت قطب علمی داده‌های ترتیبی و فضایی دانشگاه فردوسی مشهد قدردانی می‌کنند.

منابع

۱. امیدی، م.، محمدزاده، م. و مرید، س.، *تحلیل احتمالاتی شدت مدت خشکسالی در استان تهران با استفاده از توابع مفصل*، مجله تحقیقات آب و خاك ایران، ج ٤١، ۱ (۱۳۸۹) ۹۵-۱۰۱.

2. B. Bacchi, G. Becciu, N. T. Kottegoda, "Bivariate Exponential Model Applied to Intensities and Durations of Extreme Rainfall", Journal of Hydrology 155 (1994) 225-236.
3. A. Cancelliere, J. D. Salas, "Drought Length Properties for Periodic-Stochastic Hydrologic Data, Water Resources Research 40", W02503, DOI: 10.1029/2002WR001750 (2004).
4. U. Cherubini, E. Luciano, W. Vecchiato, "Copula Methods in Finance", John Wiley and Sons, Chichester (2004).
5. P. Deheuvels, "Caracterisation Complet des Lois Extremes Multivareets et de la Convergence des Types Extremes", Publications de l'Institut de Statistique de l'Universite de Paris, 23 (1978) 1-37.
6. J. Galambos, "The Asymptotic Theory of Extreme Order Statistics", Wiley, New York (1978).
7. J. Gonzalez, J. B. Valdes, "Bivariate Drought Recurrence Analysis Using Tree Ring Reconstructions", Journal of Hydrologic Engineering ASCE, 8 (2003) 247-258.
8. H. Joe, "Multivariate Models and Dependence Concepts", Chapman and Hall (1997).
9. G. Kimeldorf, A. Sampson, "Uniform Representations of Bivariate Distributions", Comm Statist A Theory Methods, 4 (1975) 617-627.
10. P. Kumar, "Probability Distributions and Estimation of Ali-Mikhail-Haq Copula", Applied Mathematical Sciences, 4 (2010) 657-666.

11. R. B. Nelsen, "An Introduction to Copulas", Springer (2006).

12. J. D. Salas, C. Fu, A. Cancelliere, D. Dustin, D. Bode, A. Pineda, E. Vincent, "Characterizing the Severity and Risk of Drought in the Poudre River", Colorado, Journal of Water Resources Planning and Management", ASCE, 131 (2005) 383-393.

13. G. Schwarz, "Estimating the Dimension of a Model", Annals of Statistics, 6 (1978) 461-464.

14. B. Schweizer, A. Sklar, "Associative Functions and Statistical Triangle Inequalities", Publ, Math. Debrecen, 8 (1961) 169-186.

15. J. T. Shiau, H. W. Shen, "Recurrence Analysis of Hydrologic Droughts of Differing Severity", Journal of Water Resources Planning and Management, 127 (2001) 30-40.

16. J. T. Shiau, "Return Period of Bivariate Distributed Hydrological Events", Stochastic Environmental Research and Risk Assessment, 17 (2003) 42-57.

17. J. T. Shiau, "Fitting Drought Duration and Severity with Two-Dimensional Copulas", Water Resources Management, 20 (2006) 795-815.

18. A. Sklar, "Fonctions de Repartition an Dimensions et Leurs Marges", Publications de l'Institut de Statistique de l'Universite de Paris, 8 (1959) 229-231.

19. G. A. Youange, R. L. Smith, "Essentials of Statistical Inference, Cambridge University Press", UK (2005).

20. S. Yue, T. B. M. J. Ouarda, B. Bobee, P. Legendre, P. Bruneau, "The Gumbel Mixed Model for Flood Frequency Analysis", Journal of Hydrology, 226 (1999) 88-100.

21. E. Zelenhastic, A. Salvai, "A Method of Streamflow Drought Analysis", Water Resources Research, 23 (1987) 156-168.