ریاضی و جامعه
شاپا (چاپی): ۲۳۴۵-۶۴۹۳، شاپا(الکترونیکی): ۶۵۰۷ – ۲۳۴۵
جلد ۷، شماره ۲ (۱۴۰۱)، صص. ۹-۱۸
http://math-sci.ui.ac.ir
http://www.ui.ac.ir

برازش بهترین توزیع احتمالاتی به بیشینهٔ بارش روزانه در سال (مطالعهٔ موردی ایستگاه‌های همدیدی اصفهان و کاشان)

حمید قربانی

چکیده. هدف این مقاله، یافتن بهترین توزیع احتمالاتی برای بیشینهٔ بارش روزانه در سال با استفاده از داده‌های ایستگاه‌های همدیدی اصفهان با دورهٔ (۲۰۱۰ – ۱۹۵۱) و کاشان برای دورهٔ (۲۰۱۰ – ۱۹۶۶) است. توزیع‌های متداول در متون از جمله گاما، لگ–نرمال، وایبل، پیرسون نوع سه، توزیع‌های معکوس گاما، معکوس گامبل و معکوس نرمال، همچنین توزیع‌های مقدار فرین، با استفاده از روش بیشینه‌ی درستنمایی به داده‌ها برازش شدند. نیکوئی برازش توزیع‌ها با استفاده از آزمون کلموگروف–اسمرنوف ارزیابی گردید و در ادامه بهترین توزیع برازش شده به داده‌ها با استفاده از معیار اطلاع آکائیک انتخاب گردید. در مجموع، توزیع‌های نرمال معکوس و لگ–نرمال به‌عنوان بهترین توزیع به‌ترتیب برای ایستگاه اصفهان و کاشان انتخاب شدند. برای محاسبات مربوط به برازش مدل‌ها از قابلیت‌های بسته‌های مختلف نرم افزار R استفاده شده است.

۱. مقدمه

بارندگی یکی از مهمترین منابع طبیعی برای تولید محصولات زراعی در مناطق نیمه خشک است. از دیرباز، یافتن توزیع احتمال مناسب برای میزان بارندگی، موضوع مورد علاقه پژوهشگران آمار و آب‌شناسی بوده است. فیشر [۵] تأثیر بارندگی را بر عملکرد گندم مطالعه و نشان داد که توزیع بارندگی در طول یک فصل است که بر عملکرد محصول تأثیر می‌گذارد، نه مقدار کل بارش. برای بررسی توزیع آماری بارش، بسته به هدف پژوهش و چگونگی ثبت داده‌ها، می‌توان بازه‌های زمانی مختلفی مانند بارش روزانه، هفتگی، ماهانه، فصلی، سالانه، بیشینهٔ بارش چند روز پشت سرهم در مناطق پربارش و یا بیشینه (و یا کمینهٔ) بارش در بازه‌های مذکور را در نظر گرفت، که همگی در پیش‌بینی الگوی بارش مهم هستند. از آنجا که بارش‌های شدید یکی از دلایل اولیه و طبیعی سیل می‌باشند، از نظر کاربردی داشتن اطلاعات کافی در مورد پیشامدهای فرین با دورهٔ بازگشت طولانی، می‌تواند برای طراحی سد، پل و سامانه‌های آبیاری، جلوگیری از آسیب‌های مرتبط با سیل و پروژه‌های مختلف مهندسی مرتبط با آب‌شناسی مفید باشد. از این بین بیشینهٔ بارش سالانه توسط مهندسین طراح و کارشناسان آب‌شناس برای برنامه‌ریزی اقتصادی، طراحی سازه‌های مرتبط به

عبارات و کلمات کلیدی: روش‌های ناپارامتری، آزمون تصادفی بودن، آزمون نیکوئی برازش، بیشینهٔ بارش روزانه در سال، توزیع‌های احتمالاتی، کمینه‌کردن غیر خطی، روش بیشینه‌ی درستنمایی.
دبیرتخصصی رابط: افشین پرورده
نوع مقاله: پژوهشی
تاریخ دریافت: ۱۴۰۱/۰۳/۳۱ تاریخ پذیرش: ۱۴۰۱/۰۷/۲۰
. http://dx.doi.org/10.22108/MSCI.2022.134147.1512

آب با اندازهٔ کوچک و متوسط، همچنین تعیین پارامترهای زهکشی برای زمین های کشاورزی در آب و هوای نیمه خشک استفاده می‌شود. از طرفی در مناطق خشک و نیمه‌خشک، بارش‌های فرین پیشامدهای نادر آب و هوایی محسوب می‌شوند. ابزار آماری برای تحلیل چنین پیشامدهایی قضیهٔ مقدار فرین است. توزیع‌های تعمیم یافتهٔ مقدار فرین در کاربرد این قضیه، نقش مهمی ایفا می‌کنند [۱۲]. با استفاده از نظریهٔ توزیع فرین می‌توان نشان داد که فراوانی پیشامدهای فرین بیش از آنچه متأثر از تغییر میانگین بارش (پارامتر مکان توزیع) باشد به تغییر در واریانس بارش (پارامتر مقیاس توزیع) وابسته است [۹]. از منظر تاریخی توزیع‌های گامبل، فریشه و وایبل به عنوان سه خانوادهٔ مهم از توزیع‌های مقدار فرین عملکرد خوبی در برازش به مقادیر بارش فرین در طول بازه‌های زمانی مختلف، نشان داده‌اند [۷].

با تعریف بیشینهٔ بارش روزانه در سال (برای اختصار آن را بیشینهٔ بارش سالانه می‌نامیم) به‌عنوان یک مقدار فرین، در [۱۰] نشان داده است که در بنگلادش توزیع مقدار فرین تعمیم یافته (و توزیع گامای تعمیم یافتهٔ چهار پارامتری) برای نیمی از ایستگاه‌ها، بهترین برازش را فراهم می‌آورند. توزیع‌های دیگر نیز ممکن است برای برازش به بیشینهٔ بارش سالانه مورد استفاده قرار گیرد. برای مثال نشان داده شده است که توزیع لگ–نرمال دو پارامتری در هند [۱۱] و [۱۸] و توزیع لگ–پیرسون نوع ۳ در پاکستان [۱]، بهترین برازش برای داده‌های بیشینهٔ بارش سالانه را فراهم آورده‌اند. به‌طور معمول، برای برآورد پارامترهای توزیع‌های احتمال، از روش بیشینه‌ی درستنمایی، گشتاورهای خطی و روش بیزی استفاده می‌شود [۱۴]، [۱۷]، [۱۹]. این پژوهش با هدف یافتن بهترین توزیع احتمال برای بیشینهٔ بارش روزانه در سال، با استفاده از داده‌های دو ایستگاه همدیدی در سطح استان اصفهان طراحی شده است. برای تحلیل داده‌های مربوط به ایستگاه‌های دیگر، در صورت دسترسی به اطلاعات آنها، می‌توان از روش مطالعه در این مقاله استفاده کرد.

۲. منطقهٔ مورد مطالعه و داده‌ها

میانگین بارش سالیانه در جهان، ۷۵۰ میلیمتر و در ایران، میانگین سالانهٔ حدود ۲۵۰ میلیمتر است، یعنی میانگین بارندگی در ایران در طول سال، از یک سوم میزان بارندگی در جهان نیز کم‌تر است. در حدود ۸۳ درصد از مناطق ایران را می‌توان در زمره مناطق کم‌بارش قرار داد که میزان بارندگی در آن‌ها در حدود ۱۸۸ میلیمتر اندازه گرفته شده است. بخش پربارش کشور بیشتر در محدود کرانه‌های دریای خزر با بارش در حدود ۹۰۰ الی ۱۰۰۰ میلیمتر است. همچنین ۱/۱ درصد از وسعت خشکی‌های جهان در ایران واقع شده به‌طوری‌که ۲۵ درصد خاک ایران در مناطق فراخشک، ۴۰ درصد در مناطق خشک و ۲۵ درصد در مناطق نیمه‌خشک قرار دارد [۲۰]. استان اصفهان با مساحت حدود ۱۰۷۱۴۵ کیلومتر مربع (معادل ۶/۵۷ درصد از مساحت کشور) دارای میانگین بارندگی برابر ۱۳۰ میلیمتر در سال است. این مقدار بارندگی نزدیک به نصف میانگین سالانه بارندگی در کشور و یک‌ششم میانگین جهانی است. این استان به دلیل موقعیت جغرافیایی، همچنین وجود مناطق خشک و نیمه‌خشک، مستعد روبرو شدن با خشکسالی‌های شدید می‌باشد. بر اساس گزارش ادارهٔ تحقیقات هواشناسی استان اصفهان، تحلیل شاخص استاندارد بارش تبخیر و تعرق دورهٔ ده‌ساله تا پایان اسفند ۱۴۰۰، بیانگر آن است که درجات خفیف تا بسیار شدید خشکسالی بلند مدت بسیاری از مناطق استان اصفهان را فرا گرفته است به گونه ای که ۸۶/۸ درصد از مساحت استان درگیر خشکسالی بوده است (۱۷/۹ درصد خشکسالی خفیف، ۲۰/۳ درصد خشکسالی متوسط، ۱۶/۵ درصد خشکسالی شدید و ۱۳/۲ درصد خشکسالی بسیار شدید) و تنها ۱۳/۲ درصد از مساحت استان شرایط عادی تا ترسالی متوسط را داشته است.

http://dx.doi.org/10.22108/MSCI.2022.134147.1512

در این مطالعه، آمار مربوط به بیشینهٔ بارش روزانه در سال مربوط به دو ایستگاه اصفهان و کاشان از سازمان هواشناسی ایران گرفته شده است. البته واضح است که برای انجام یک بررسی آماری با دقت قابل اطمینان، طول آماری طولانی‌تر ارجح است ولی از دیرباز مشکل ساختاری در ثبت داده و فراهم آوردن آن‌ها از سوی سازمان‌ها برای پژوهشگران وجود داشته است. به‌تازگی چند درگاه اینترنتی که امکان دانلود داده‌های هواشناسی از آن‌ها فراهم بود بسته‌شده است و پژوهشگر برای دستیابی به داده‌ها، ناچار به ارائه درخواست اداری و پرداخت هزینه مالی می‌شود. به‌هرحال در مورد این دو ایستگاه جدول ۱، آمار خلاصهٔ مقدار بارش و طول آماری داده‌ها را نشان می‌دهد.

جدول ۱. ایستگاه‌های مورد مطالعه و آمار بارش

نام ایستگاه	ارتفاع از سطح دریا (m)	طول جغرافیایی	عرض جغرافیایی	میانگین بارش ماهانه (mm)	انحراف م. بارش ماهانه (mm)	دورهٔ آماری
اصفهان	۱۵۵۵٫۴	$۵۱^\circ\ ۴۰' E$	$۳۲^\circ\ ۳۷' N$	۱۰٫۴	۱۵٫۵	(۲۰۱۰ − ۱۹۵۱)
کاشان	۹۸۲٫۳	$۵۱^\circ\ ۲۷' E$	$۳۳^\circ\ ۵۹' N$	۱۱٫۲	۱۶٫۸	(۲۰۱۰ − ۱۹۶۶)

۳. مدل‌های احتمالاتی

برای مدل‌سازی داده‌های بارندگی در مقیاس‌های زمانی مختلف می‌توان از توزیع‌های احتمالی استفاده کرد. با توسعهٔ روش‌های تحلیلی در آمار کاربردی و افزایش نقش نرم‌افزارها در آسان‌سازی محاسبات پیچیدهٔ ریاضی، یکی از کاربردهای آمار در رشته‌های مرتبط با آب و هوا، یافتن بهترین توزیع احتمال مناسب برای مجموعه داده‌های مشاهده شده و برازش آن است، که در قالب روش‌های تحلیل فراوانی مورد توجه برخی از پژوهشگران قرار گرفته است [۱۳]. به‌طور خلاصه، هدف از تحلیل فراوانی، برای نمونه تحلیل فراوانی بیشینهٔ بارش، یافتن برآوردی از احتمال این پیشامد است که بیشینهٔ بارش، به‌عنوان یک متغیر تصادفی، بیشتر از یک مقدار فرین باشد. توجه کنید که در تحلیل فراوانی دو نوع رویکرد وجود دارد: نخست تحلیل بر اساس داده‌های هر ایستگاه (محلی) به‌طور جداگانه و رویکرد دیگر، تحلیل بر اساس داده‌های همه ایستگاه‌های یک منطقه جغرافیایی. اگر فقط داده‌های بارش ثبت‌شده در یک ایستگاه در دسترس باشد، تجزیه و تحلیل فراوانی فقط روی داده‌های همان ایستگاه انجام می‌شود که روش تحلیل در ایستگاه (یا محلی) نامیده می‌شود. حال اگر مشاهدات بارش در ایستگاه‌های مجزا در یک منطقه ثبت شده باشد و این داده‌ها به‌طور مشترک برای استنباط آماری استفاده شوند، به آن تحلیل فراوانی منطقه‌ای گفته می‌شود [۱۵].
در این تحقیق برای تحلیل داده‌ها، از روش تحلیل در ایستگاه استفاده شده است. برای برازش توزیع به داده‌ها، نیاز به فرض استقلال سری زمانی مشاهدات داریم. همچنین فرض می‌کنیم مشاهدات هم توزیع‌اند، یعنی از یک توزیع آماری پیروی می‌کنند. حجم نمونه نیز باید به اندازه کافی بزرگ باشد تا نتایج برازش قابل اطمینان باشند. برای بررسی استقلال یک سری زمانی، بیشتر از تابع خودهمبستگی با محاسبهٔ خود همبستگی‌ها در تأخیرهای زمانی متفاوت، استفاده می‌شود. اگر سری زمانی مستقل باشد، انتظار داریم فرض صفر مبنی بر صفر بودن همهٔ خودهمبستگی‌ها به ازای هر تأخیر زمانی، پذیرفته شود. به‌عبارت دیگر چنانچه سری زمانی غیر تصادفی باشد، فرض مقابل مبنی بر اختلاف معنی‌دار یک یا چند مقدار خودهمبستگی با صفر پذیرفته خواهد شد. آزمون

http://dx.doi.org/10.22108/MSCI.2022.134147.1512

لیونگ–باکس نوعی آزمون آماری است که صفر بودن همزمان گروهی از خودهمبستگی‌ها به ازای تأخیرهای زمانی متفاوت را آزمون می‌کند. این آزمون به‌طور وسیع در اقتصاد سنجی و دیگر کاربردهای تحلیل سری‌های زمانی استفاده می‌شود [۳]. همچنین بستهٔ randtests در نرم افزار R، چندین آزمون ناپارامتری برای تصادفی بودن یک دنباله از مشاهدات، ارائه می‌دهد که از آن جمله می‌توان به آزمون دوها و آزمون رتبه‌ای بارتلت اشاره کرد [۲].

جدول ۲. توزیع‌های احتمالاتی و پارامترهای آنها

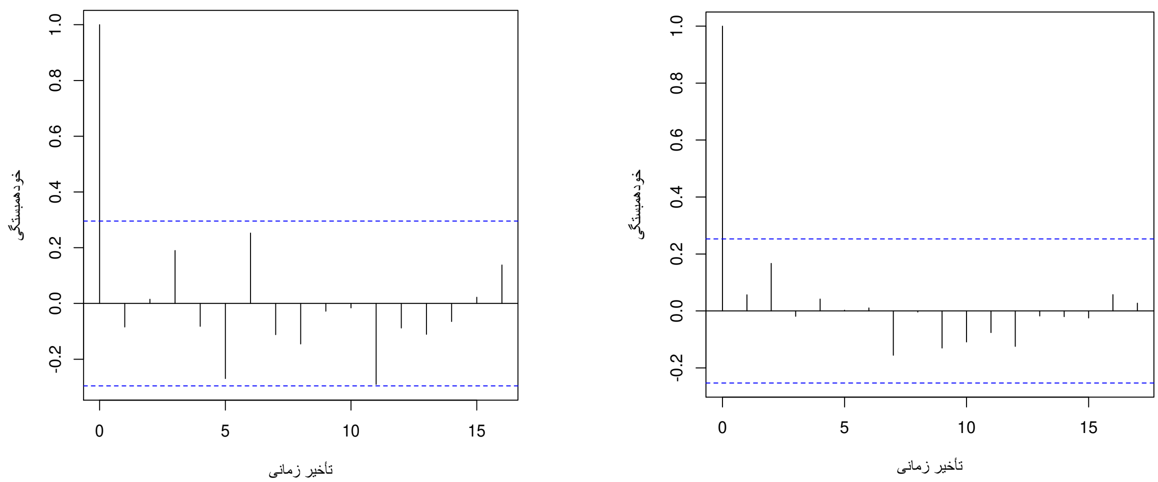
توزیع	تابع چگالی	میانگین	واریانس
گاما	$\frac{1}{\Gamma(\alpha)\lambda^{\alpha}}x^{\alpha-1}e^{-x/\lambda},\ x\in(0,\infty)$	$\alpha\lambda$	$\alpha\lambda^{2}$
گامبل معکوس	$\frac{1}{\sigma}e^{-\frac{x-\mu}{\sigma}-e^{-\frac{x-\mu}{\sigma}}},\ x\in(-\infty,\infty)$	$0.57721\sigma+\mu$	$\frac{\pi^{2}\sigma^{2}}{6}$
نرمال معکوس	$\sqrt{\frac{1}{2\pi\lambda x^{3}}}e^{-\frac{(x-\mu)^{2}}{2x\lambda\mu^{2}}},\ x\in(0,\infty)$	μ	$\lambda\mu^{3}$
گامای معکوس	$\frac{\mu^{\alpha}(1+\alpha)^{\alpha}x^{-(\alpha+1)}e^{-\mu(1+\alpha)/x}}{\Gamma(\alpha)},\ \alpha=\frac{1}{\sigma^{2}};x\in(0,\infty),$	$\frac{(1+\sigma^{2})\mu}{(1-\sigma^{2})},\sigma^{2}<1$	$\left(\frac{(1+\sigma^{2})\mu}{(1-\sigma^{2})}\right)^{2}.\frac{1}{1-2\sigma^{2}},\sigma^{2}<0.5$
لگ–نرمال	$\frac{1}{\sqrt{2\pi}\sigma x}\exp\left(\frac{-(\log(x)-\mu)^{2}}{2\sigma^{2}}\right),\ x\in(0,\infty)$	$exp(\mu+1/2\sigma^{2})$	$exp(2\mu+\sigma^{2})(exp(\sigma^{2})-1)$
وایبل	$\alpha\beta^{-\alpha}x^{\alpha-1}\exp\left[-(\frac{x}{\beta})^{\alpha}\right],\ x\in(0,\infty)$	$\beta\Gamma(1+\frac{1}{\alpha})$	$\beta^{2}\left(\Gamma(1+\frac{2}{\alpha})-\Gamma^{2}(1+\frac{1}{\alpha})\right)$
پیرسون نوع سه	$\frac{1}{\Gamma(\alpha)\lvert s\rvert^{\alpha}}\lvert x-\lambda\rvert^{\alpha-1}e^{-\frac{\lvert x-\lambda\rvert}{s}},\ s\neq 0,\alpha>0,\frac{x-\lambda}{s}\geq 0$	$\alpha s+\lambda$, for $s>0$	αs^{2}, for $s>0$

۴. برازش مدل‌ها

قبل از اقدام برای یافتن بهترین توزیع احتمالاتی برای داده‌های بیشینهٔ بارش، نخست فرض استقلال داده‌ها را بررسی می‌کنیم. شکل ۱ نمودار توابع خودهمبستگی برای تأخیرهای مختلف زمانی دو مجموعه از داده‌ها را نشان می‌دهد. از این نمودارها چنین برداشت می‌شود که فرض صفر توابع خودهمبستگی به ازای هر لگ زمانی درج شده روی محور افقی تفاوت معنی‌داری با صفر ندارد.

http://dx.doi.org/10.22108/MSCI.2022.134147.1512

همچنین بر اساس آزمون رتبه‌ای بارتلت فرض استقلال به ترتیب برای بیشینهٔ بارش در ایستگاه اصفهان (p-value = ۰/۱۲) و کاشان (p-value = ۰/۷۲) در سطح ۰/۰۵ پذیرفته می‌شود.



(ب) ایستگاه کاشان

(آ) ایستگاه اصفهان

شکل ۱. توابع خود همبستگی نمونه‌ای بیشینهٔ بارش روزانه در سال

در ادامه، جداول ۳ و ۴ به ترتیب نشان‌دهندهٔ نتایج برازش توزیع‌های موجود در جدول ۳ به داده‌های بیشینهٔ بارش در ایستگاه‌های اصفهان و کاشان هستند. در این جداول، پارامترهای برآورد شدهٔ توزیع‌های مختلف به روش بیشینه‌ی درستنمایی و انحراف معیار نظیر آن‌ها، به همراه AIC و آمارهٔ آزمون نیکوئی برازش کلموگروف–اسمرنوف آمده است. برای برازش توزیع‌ها از بستهٔ نرم‌افزاری R به نام `fitdistrplus` استفاده شده است [۴]. چنانچه توزیعی در هستهٔ R نبوده است ولی در یکی از بسته‌های نرم‌افزار تعریف شده باشد (بستهٔ مبدأ)، مانند توزیع پیرسون نوع سه، ابتدا تابع چگالی از بستهٔ مبدأ به بستهٔ مقصد `fitdistrplus` فراخوانی شده است، و سپس از امکانات بستهٔ مقصد برای برازش آن استفاده شده است. با توجه به پی–مقدار آزمون نیکوئی برازش، همهٔ توزیع‌ها به خوبی به داده‌ها برازش می‌شوند ولی با توجه به مقادیر AIC نظیر توزیع‌های برازش شده، از بین توزیع‌های ذکر شده در جدول ۳، توزیع‌های نرمال معکوس و لگ–نرمال، به ترتیب، بهترین برازش به داده‌های اصفهان و کاشان را فراهم می‌آورند. شکل‌های ۲ و ۳ نیز نمودارهایی برای ارزیابی نیکوئی برازش این دو توزیع را به‌دست می‌دهند. لازم به ذکر است که افزون بر توزیع‌های جدول ۳، یک سیاهه کامل از توزیع‌های موجود در بستهٔ نرم‌افزاری R به نام `gamlss` [۱۶] به داده‌ها برازش شد که با توجه به معیار AIC هیچکدام نتوانستند عملکرد بهتری از دو توزیع معرفی شده بالا داشته باشند. از معروف‌ترین توزیع‌های این سیاهه می‌توان به توزیع‌های پارتوی نوع یک و دو، پارتوی تعمیم یافته، گامای معکوس تعمیم یافته، نرمال معکوس تعمیم یافته، بتای تعمیم یافتهٔ نوع دو و وایبل معکوس اشاره کرد. همچنین همهٔ توزیع‌های خانواده لیندلی، که در مقالهٔ [۶] ذکر شده‌اند، به‌کار گرفته شدند که شاید برازش بهتری از دو مدل بالا فراهم آورند که چنین نشد. همچنین توزیع مقدار فرین تعمیم یافته با استفاده از بستهٔ `extRemes` [۸] نیز به داده‌ها برازش شد که برازش بهتری حاصل نشد.

http://dx.doi.org/10.22108/MSCI.2022.134147.1512

جدول ۳. پارامترهای برآورد شده و خطای معیار نظیر آنها، معیار آکائیک و آماره آزمون کلموگروف–اسمیرنوف برای داده‌های اصفهان

نام توزیع	برآورد (خطای معیار)	AIC	KS (پی–مقدار)
گاما	$\hat{\alpha} = ۵/۳۸\ (۰/۹۵)$	۴۲۴/۲۸	$۰/۰۹۷\ (p = ۰/۶۳)$
	$\hat{\lambda} = ۳/۷۰\ (۰/۶۹)$		
لگ–نرمال	$\hat{\mu} = ۲/۸۹\ (۰/۰۶)$	۴۲۱/۷۰	$۰/۰۶۱\ (p = ۰/۹۸)$
	$\hat{\sigma} = ۰/۴۳\ (۰/۰۴)$		
وایبل	$\hat{\alpha} = ۲/۳۳\ (۰/۲۲)$	۴۳۱/۴۰	$۰/۱۳\ (p = ۰/۲۵)$
	$\hat{\beta} = ۲۲/۵۱\ (۱/۳۲)$		
گامبل معکوس	$\hat{\mu} = ۱۵/۸۶\ (۰/۸۹)$	۴۲۳/۲۴	$۰/۰۸۵\ (p = ۰/۷۸)$
	$\hat{\sigma} = ۶/۵۸\ (۰/۶۹)$		
نرمال معکوس	$\hat{\mu} = ۱۹/۸۷\ (۱/۱۷)$	۴۲۱/۳۴	$۰/۰۶\ (p = ۰/۹۸)$
	$\hat{\lambda} = ۰/۰۱\ (۰/۰۰۲)$		
پیرسون نوع سه	$\hat{\alpha} = ۲/۷۳\ (۰/۸۲)$	۴۲۳/۵۰	$۰/۰۷۳\ (p = ۰/۹۰)$
	$\hat{s} = ۵/۴۳\ (۱/۳۵)$		
	$\hat{\lambda} = ۵/۰۶\ (۱/۳۰)$		

مراجع

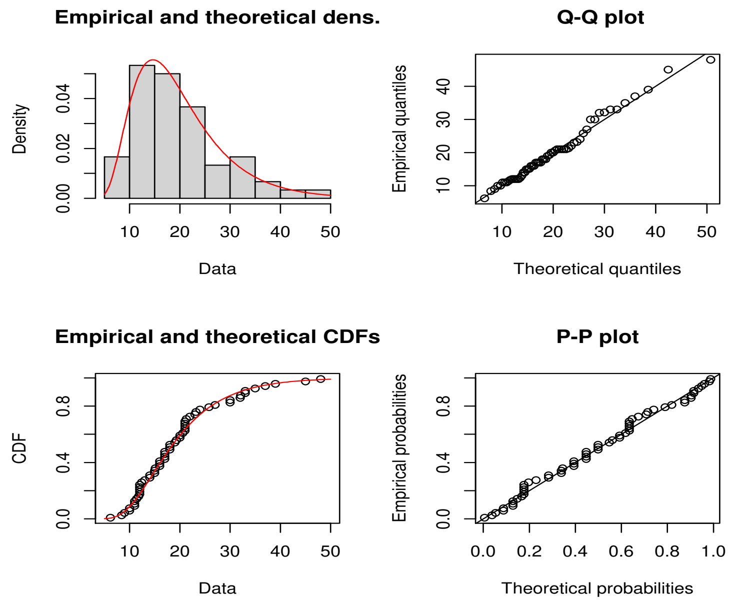
[1] M. T. Amin, M. Rizwan and A. A. Alazba, A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan, *Open Life Sci.*, **11** (2016) 432–440.

[2] F. Caeiro and A. Mateus, randtests: Testing randomness in *R*, *R* package manual vailable at: https://CRAN.R-project.org/package=randtests, (2014).

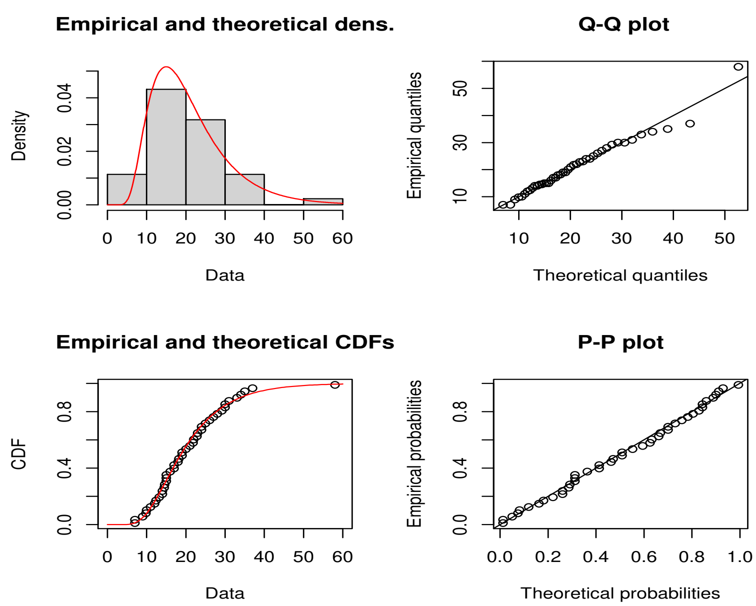
[3] J. D. Cryer and K-S. Chan, *Time Series Analysis: With Applications in R*, Springer, New York, 2008.

[4] M. L. Delignette-Muller and C. Dutang, fitdistrplus: An *R* Package for Fitting Distributions Distributions, *Journal of Statistical Software*, **64** (2015) 1–34.

[5] R. A. Fisher, The influence of rainfall on the yield of wheat at Rothamsted, *Phil. Trans. R. Soc. Lond. B*, **213** (1925) 89–142.



شکل ۲. نمودارهای نیکوئی برازش توزیع نرمال معکوس برای بیشینهٔ بارش سالانهٔ ایستگاه اصفهان.



شکل ۳. نمودارهای نیکوئی برازش توزیع لگ–نرمال برای بیشینهٔ بارش سالانهٔ ایستگاه کاشان.

[6] H. Ghorbani and M. Irshad, Correction to: The Zografos-Balakrishnan Lindley distribution properties and applications, *Statistica*, **82** (2022) 45–64.

http://dx.doi.org/10.22108/MSCI.2022.134147.1512

جدول ۴. پارامترهای برآورد شده و خطای معیار نظیر آنها، معیار آکائیک و آماره آزمون کلموگروف–اسمرنوف برای دادهٔ کاشان

نام توزیع	برآورد (خطای معیار)	AIC	KS (پی–مقدار)
گاما	$\hat{\alpha} =$ ۵٫۱۴ (۱٫۰۶)	۳۱۸٫۳۱	۰٫۰۷۶ ($p =$ ۰٫۹۶)
	$\hat{\lambda} =$ ۴٫۰۶ (۰٫۸۸)		
لگ–نرمال	$\hat{\mu} =$ ۲٫۹۴ (۰٫۰۷)	۳۱۷٫۷۴	۰٫۰۵۸ ($p =$ ۰٫۹۹۸)
	$\hat{\sigma} =$ ۰٫۴۵ (۰٫۰۵)		
وایبل	$\hat{\alpha} =$ ۲٫۲۸ (۰٫۲۴)	۳۲۲٫۹۰	۰٫۰۶۷ ($p =$ ۰٫۹۹)
	$\hat{\beta} =$ ۲۳٫۶۳ (۱٫۶۵)		
گامبل معکوس	$\hat{\mu} =$ ۱۶٫۶۳ (۱٫۱۵)	۳۱۸٫۱۶	۰٫۰۷۸ ($p =$ ۰٫۹۵)
	$\hat{\sigma} =$ ۷٫۲۴ (۰٫۸۷)		
نرمال معکوس	$\hat{\mu} =$ ۲۰٫۹۰ (۱٫۵۰)	۳۱۷٫۷۶	۰٫۰۵۷ ($p =$ ۰٫۹۹۹)
	$\hat{\lambda} =$ ۰٫۰۱ (۰٫۰۰۲)		
پیرسون نوع سه	$\hat{\alpha} =$ ۲٫۷۷ (۱٫۴۰)	۳۱۹٫۳۰	۰٫۰۵۷ ($p =$ ۰٫۹۹۹)
	$\hat{s} =$ ۵٫۸۴ (۲٫۱۷)		
	$\hat{\lambda} =$ ۴٫۷۵ (۲٫۶۷)		

[7] A. F. Jenkinson, The frequency distribution of the annual maximum (or minimum) of meteorological elements, *Q. J. R. Meteorol. Soc.*, **81** (1955) 158–171.

[8] E. Gilleland and R. W. Katz, extRemes 2.0: An Extreme Value Analysis Package in *R*, *Journal of Statistical Software*, **72** (2016) 1–39.

[9] R. W. Katz and B. G. Brown, Extreme events in a changing climate: Variability is more important than averages, *Climatic Change*, **21** (1992) 289–302.

[10] M. M. Khudri and F. Sadia, Determination of the best fit probability distribution for annual extreme precipitation in Bangladesh, *Eur. J. Sci. Res.*, **103** (2013) 391–404.

[11] A. Kumar, Prediction of annual maximum daily rainfall of Ranichauri (Tehri Garhwal) based on probability analysis, *Ind. J. Soil Conserv.*, **28** (2000) 178–180.

[12] G. Lazoglou and C. Anagnostopoulou, An overview of statistical methods for studying the extreme rainfalls in Mediterranean.

[13] R. Maity, *Statistical Methods in Hydrology and Hydroclimatology*, Springer, Singapore, 2022.

[14] S. Nadarajah and D. Choi, Maximum daily rainfall in South Korea, *J. Earth Syst. Sci.*, **116** (2007) 311–320.

[15] M. Naghettini, *Fundamentals of Statistical Hydrology*, Springer, Switzerland, 2017.

[16] R. A. Rigby and D. M. Stasinopoulos, Generalized additive models for location, scale and shape, (with discussion), *J. R. Stat. Soc., Ser. C, Appl. Stat.*, **54** (2005) 507–554.

[17] R. T. Sahu, M. K. Verma, M. K. and I. Ahmad, Regional frequency analysis using L-moments methodology, a review, In: K. K. Pathak, J. M. S. J. Bandara, R. Agrawal and (eds) Recent, *Trends in Civil Engineering, Lecture Notes in Civil Engineering*, **77** (2021) 811–832.

[18] R. K. Singh, Probability analysis for prediction of annual maximum rainfall of Eastern Himalaya (Sikkim mid hills), *Ind. J. Soil Conserv.*, **29** (2001) 263–265.

[19] E. Smith, *Bayesian modelling of extreme rainfall data*, PhD thesis, University of Newcastle, 2005.

[۲۰] پ. کردوانی، خشکسالی و راه‌های مقابله با آن در ایران، آب: در کشاورزی، صنعت و شهر، انتشارات دانشگاه تهران، ۱۳۸۶.

حمید قربانی

دانشگاه کاشان، دانشکده علوم ریاضی

hamidghorbani@kashanu.ac.ir

حمید قربانی متولد اردیبهشت ۱۳۵۶ در شهر اصفهان است. وی در سال ۱۳۷۳ وارد مقطع کارشناسی رشته آمار دانشگاه اصفهان و در سال ۱۳۷۷ وارد مقطع کارشناسی ارشد رشته آمار ریاضی دانشگاه شهید بهشتی شد. وی دکترای خود را در سال ۱۳۸۳ از دانشگاه صنعتی فرایبرگ آلمان در زمینه‌ی آمار فضایی و هندسه تصادفی دریافت کرد و هم اکنون استادیار گروه آمار دانشگاه کاشان می‌باشد.

http://dx.doi.org/10.22108/MSCI.2022.134147.1512

Determination of best-fit probability distribution of annual maximum daily precipitation (Case study- Isfahan and Kashan stations)

Hamid Ghorbani

Abstract: The goal of this paper is to determine the best-fit probability distribution to describe the annual maximum daily rainfall for the Isfahan (period 1951-2010) and Kashan (period 1966-2010) stations located distantly in Isfahan province as an arid geographical region of Iran. The Gamma, Lognormal, Weibull, Pearson type III, inverse Gumbel, inverse Gamma, inverse Normal, and Generalized Extreme Value distributions are fitted for these purposes. Parameters of these distributions were estimated by the method of maximum likelihood. The performances of the distributions are evaluated using the Kolmogorov-Smirnov goodness-of-test. Finally, the inverse Normal distribution and Lognormal distribution were empirically proved to be the most appropriate distribution of the annual maximum daily rainfall for Isfahan and Kasha stations, respectively using the Akaike information criterion.

Keywords: nonparametric methods, test of randomness, annual maximum daily precipitation, probabilistic distributions, maximum likelihood method, nonlinear minimizationlikelihood method.

Hamid Ghorbani

Faculty of Mathematical Sciences, Kashan University, Kashan, Iran.

Email: hamidghorbani@kashanu.ac.ir

Communicated by Afshin Parvardeh.

Article Type: Research Paper.

Received: 21/06/2022, Accepted: 20/07/2022.

http://dx.doi.org/10.22108/MSCI.2022.134147.1512