

# استنباط آماری در توزیع دایره‌ای ون میزس تحت نمونه‌گیری وزنی

فاطمه شاه سنایی<sup>۱</sup>، رحیم چینی پرداز<sup>۲</sup>

تاریخ دریافت: ۱۴۰۲/۰۸/۱۸

تاریخ پذیرش: ۱۴۰۳/۱۲/۱۵

## چکیده:

در بسیاری از زمینه‌های مختلف علمی اندازه‌گیری‌ها جهت‌دار هستند. تحقیق راجع به ویژگی‌های این نوع داده‌ها نیازمند استفاده از ابزارهای جدید آماری است. روش‌های آماری برای تحلیل این گونه داده‌ها به آمار غیرخطی معروف هستند. آمار دایره ای مثالی از این حوزه از آمار است. در بررسی آمار دایره ای ممکن است نمونه‌گیری به‌گونه‌ای باشد که هر مشاهده شانس متناسب با تابعی نامنفی از آن ثبت شود. چنین نمونه‌ای را نمونه دایره ای وزنی و توزیع متناظر با این نمونه‌ها را توزیع دایره ای وزنی می‌نامند. در این مقاله به توزیع‌های دایره‌ای وزنی پرداخته می‌شود. با توجه به اینکه توزیع ون میزس کاربرد بسیار وسیعی در مدل بندی داده‌های دایره‌ای دارد، برآورد ماکسیمم درست‌نمایی پارامترهای توزیع دایره‌ای ون میزس وزنی مورد بررسی قرار می‌گیرد. از آنجاکه ممکن است برآوردها شناسا پذیر نباشند، به کمک یک توزیع پیشین مناسب برآوردهای بیزی منحصربه‌فرد به دست می‌آید. در یک مطالعه شبیه‌سازی وزن‌های مختلف در توزیع دایره‌ای ون میزس مقایسه می‌شوند. **واژه‌های کلیدی:** داده‌های دایره‌ای، توزیع ون میزس، تابع وزن، برآورد ماکسیمم درست‌نمایی، شبیه‌سازی رد-پذیرش

## ۱ مقدمه

برای کاربردهای دیگر می‌توان به گورتمن و پنکوئیس [۱۰]، گورتمن [۹] و لیری [۱۳] مراجعه کرد.

از طرف دیگر، در بررسی‌های آماری ممکن است نمونه‌ها به‌صورت غیر تصادفی و متناسب با تابعی غیر منفی در اختیار محقق قرار گیرند. در چنین شرایطی توزیع پایه متغیر تصادفی بیانگر رفتار دقیق متغیر تصادفی نخواهد بود. بلکه لازم است از توزیع‌های وزنی برای تحلیل مشاهدات استفاده شود ([۱]). استفاده از توزیع‌های وزنی در تحلیل بسیاری از داده‌هایی که در نمونه‌گیری از آن‌ها محدودیت‌هایی وجود دارد، معمول است. به همین دلیل استفاده از توزیع‌های وزنی را در تولیدات جنگل وارن [۱۳]، شیلات چن و کولینگ [۵] و در ترافیک موریسون [۹] را می‌توان نام برد. در حوزه‌های نظری برای توزیع‌های وزنی نیز می‌توان به آنی کریشنان و سونج [۱۲] و جاکوبو [۱] اشاره کرد.

فرض کنید  $w(x)$  که تابع وزن گفته می‌شود، تابع متناسب با شانس ثبت مشاهده در نمونه  $x$  باشد. به این معنی که گرچه رفتار متغیر تصادفی را تابع چگالی احتمال  $f(x)$  تعیین می‌کند، اما در نمونه‌گیری مشاهده  $x$  از طریق تابع مثبت  $w(x)$  در دسترس محقق قرار می‌گیرد.

ابتدا فیشر [۶] توزیع‌های وزنی را مطرح کرد؛ اما این راثو [۱۰]

در بعضی از پدیده‌های تجربی، پژوهشگران با داده‌هایی مواجه هستند که ذاتاً ماهیت اقلیدسی ندارند. داده‌های دایره‌ای که برای اندازه‌گیری زاویه و یا جهت به کار می‌روند، از این نوع هستند. تفاوت اصلی چنین داده‌هایی که به‌صورت یک‌بعدی، دوبعدی و یا با ابعاد بالاتر ظاهر می‌شوند، با داده‌های خطی متناوب بودن این داده‌ها است. به این معنی که توزیع آماری آن‌ها نیز تناوبی و متفاوت با توزیع‌های خطی خواهد بود. به‌عنوان مثال در داده‌های زاویه‌ای، زاویه  $۱^\circ$  و  $۳۵۹^\circ$  بسیار نزدیک هستند. روش تحلیلی این نوع داده‌ها، در نظر گرفتن آن‌ها روی محیط دایره‌ای با شعاع واحد می‌باشد. در چنین شرایطی می‌توان برای تحلیل آن‌ها از توزیع‌های دایره‌ای استفاده کرد.

چنین داده‌هایی می‌توانند در زمینه‌های مختلف علمی کاربرد داشته باشند. ماردیا [۱۶] در علوم زیستی، بولبرت و همکاران [۴] در زیست‌شناسی رفتاری، کاس و ون مایر [۲] در روان‌شناسی شناختی و ماردیا و همکاران [۷] از این داده‌ها در بیوانفورماتیک استفاده کرده‌اند. در علوم دیگر مانند علوم سیاسی گیل و هانگارتز [۸]، علوم محیطی آرنولد و سن گوپتا [۳] نیز دامنه‌ای از علوم به‌کارگیری این داده‌ها هستند.

<sup>۱</sup> گروه برق، پردیس صنعتی شهدای هویزه، دانشگاه شهید چمران اهواز، اهواز، ایران (نویسنده مسئول: f.shahsanaei@scu.ac.ir)

<sup>۲</sup> گروه آمار، دانشکده علوم ریاضی و کامپیوتر، دانشگاه شهید چمران اهواز، اهواز، ایران

## ۲ توزیع دایره‌ای وزنی

فرض کنید  $\theta$  متغیر تصادفی دایره‌ای با تابع چگالی احتمال دایره‌ای  $f(\theta)$  با دوره تناوب  $2\pi$  باشد؛ یعنی

$$\int_0^{2\pi} f(\theta) d\theta = 1, \quad f(\theta \pm 2k\pi) = f(\theta) \geq 0 \quad (2)$$

که در آن  $k$  عدد صحیح است. با به‌کارگیری توزیع وزنی در داده‌های دایره‌ای و در حالت کلی‌تر که پارامتر موردنظر در تابع وزن نیز قرار گرفته باشد:

$$f^w(\theta; \eta) = \frac{w(\theta, \eta) f(\theta; \eta)}{E[w(\Theta, \eta)]} \quad (3)$$

خواص تابع چگالی داشتن  $f^w(\theta)$  همانند تابع خطی  $f^w(x)$  آشکار است؛ اما لزوماً شرط دایره‌ای بودن (۲) برقرار نخواهد بود. لم ۱۰۲ زیر شرط کافی خاصیت تناوبی بودن  $f^w(\theta)$  را نشان می‌دهد.

لم ۱۰۲. اگر  $f(\theta)$  در شرایط (۲) صدق کند و اگر  $E[w(\Theta)]$  موجود باشد، برای هر تابع پیوسته  $g(\theta) = g(\theta + 2\pi)$ ، تابع  $f^w(g)$  تابع چگالی دایره‌ای است.

*اثبات.* اثبات تابع چگالی احتمال بودن به‌صورت مستقیم به دست می‌آید. برای متناوب بودن تابع  $f^w(g)$  داریم:

$$\begin{aligned} f^w(g(\theta + 2\pi)) &= \frac{w(g(\theta + 2\pi)) f(g(\theta + 2\pi))}{E[w(g(\Theta + 2\pi))]} \\ &= \frac{w(g(\theta)) f(g(\theta))}{E[w(g(\Theta))]} = f^w(g(\theta)) \end{aligned}$$

اگر  $g(\theta + 2\pi) = g(\theta)$  باشد، لم ۱۰۲ برای هر تابع  $w$  برقرار است؛ اما اگر تابع  $g$  دارای این شرط نباشد، هنوز با متناوب بودن تابع،  $w(\theta + 2\pi) = w(\theta)$  لم برقرار خواهد بود. در ساده‌ترین شکل می‌توان تابع  $w(\theta)$  را با خاصیت تناوبی  $2\pi$  در نظر گرفت. در آن صورت  $g(\theta) = \theta$  کفایت می‌کند. سه حالت ممکن می‌توان توزیع وزنی دایره‌ای داشت:

۱. فرض کنید  $w(\theta) = c$  مقدار ثابتی باشد، در این صورت

$$f^w(\theta) = f(\theta) \text{ و نمونه‌گیری وزنی و تصادفی یکسان خواهند بود.}$$

۲. اگر  $w(\theta) = 1 + \lambda_1 \sin j\theta + \lambda_2 \cos j\theta$  که در آن  $j$  عدد

صحیحی است، با شرط

$$\lambda_1, \lambda_2 \in [-1, 1], \quad |\lambda_1| + |\lambda_2| \leq 1 \quad (4)$$

در این صورت  $\Theta^w$  در خانواده انعطاف‌پذیری از توزیع‌های دایره‌ای قرار می‌گیرد که توزیع‌های متقارن، غیرمتقارن، تک مدی

بود که اصول توزیع‌های وزنی و فرمول‌بندی آن‌ها را ارائه داد. هرگاه متغیر تصادفی خطی  $X$  با تابع چگالی احتمال  $f(x; \eta)$ ، همراه با تابع وزن  $w(x, \beta)$  باشد، تابع چگالی وزنی  $X^w$  به‌صورت

$$f^w(x; \eta, \beta) = \frac{w(x, \beta) f(x; \eta)}{E[w(X, \beta)]} \quad (1)$$

است. در اینجا  $\eta$  بردار پارامترهای توزیع  $f(x; \eta)$  و  $\beta$  برداری از مقادیر مجهول در تابع وزن است. این دو بردار می‌توانند تابعی از همدیگر باشند. بدیهی است  $f^w(x; \eta, \beta)$  دارای خاصیت تابع چگالی احتمال و به‌طور معمول ساختار تابعی آن با تابع  $f(x; \eta)$  یکسان نیست.

استفاده از توزیع‌های وزنی تاکنون در داده‌های خطی متداول بوده، اما مطالعه‌ای این توزیع‌ها در داده‌های دایره‌ای صورت نگرفته است. درحالی‌که ممکن است این داده‌ها نیز تحت تأثیر یک تابع وزن در اختیار محقق قرار گیرند. دستگاه سنجش زاویه مانند دوربین چرخانی را در نظر بگیرید که چرخش آن به‌جای فاصله  $[0, 2\pi]$  زاویه‌های در دامنه  $(\pi/10, 11\pi/6)$  را فیلم‌برداری کند. در اینجا همیشه زاویه‌های ثبت‌شده در این دامنه بریده‌شده هستند. یا فرض کنید این دستگاه در چرخش خود، در بعضی از زوایا بعضی از داده‌ها را ثبت نکند. در چنین مواردی با زوایای ثبت‌شده در نمونه مشروط به تابع وزن هستند. در مثال اول تابع نشانگر  $w(\theta) = I_{(\pi/10, 11\pi/6)}(\theta)$  تابع وزن مناسب خواهد بود؛ اما در حالت دوم لازم است وزن‌های مختلفی موردبررسی قرار گیرند. ازآنجاکه برآورد پارامترهای دایره‌ای به‌طور معمول فرم پیچیده‌ای دارند، مطالعه برآوردها وقتی مشاهدات با تابع وزنی به دست می‌آیند، می‌تواند موضوع جالبی برای مطالعه باشد. به‌ویژه اگر مقادیر مجهول تابع وزن با پارامترهای توزیع مرتبط باشند. مقاله حاضر به این موضوع اختصاص یافته است.

در میان توزیع‌های دایره‌ای، توزیع ون میزس مهم‌ترین توزیع و در حقیقت همان نقشی را در این نوع داده‌ها دارد که توزیع نرمال در داده‌های خطی دارد. تأکید اصلی این مقاله روی توزیع ون میزس خواهد بود. باوجوداین، نتایج حاصل‌شده می‌تواند به‌راحتی برای توزیع‌های دیگر هم مورد استفاده قرار گیرد. مقاله در هفت بخش تنظیم‌شده که در بخش دوم مقاله معرفی توزیع‌های وزنی دایره‌ای معرفی و خواص آن‌ها موردبررسی قرار می‌گیرد. در بخش سوم به توزیع ون میزس وزنی با تأکید بیشتری مطالعه می‌شود. برآورد ماکسیمم درست‌نمایی پارامترها تحت سه تابع وزنی برای توزیع ون میزس بررسی و در بخش بعدی در یک شبیه‌سازی توانایی روش پیشنهادشده در برآورد پارامترها نشان داده خواهد شد. بخش نهایی مقاله به نتیجه‌گیری از بخش‌های قبلی و پیشنهاد برای تحقیقات بعدی خواهد بود.

خواهد بود. در اینجا  $vM(\mu, \kappa)$  عضو خانواده نمایی است ولی لزوماً  $f^w(\theta)$  چنین نخواهد بود. لم (۱۰۳) نشان می‌دهد، شرط کافی برای عضو خانواده نمایی  $f^w(\theta; \mu, \kappa)$  این است که

$$w(\theta, \lambda_1, \lambda_2) = \exp\{\lambda_1 \cos j\theta + \lambda_2 \sin j\theta\}. \quad (10)$$

لم ۱۰۳. فرض کنید  $\Theta \sim vM(\mu, \kappa_1)$  و تابع وزن (۱۰) با  $\lambda_1 = \kappa_1 \cos \mu$  و  $\lambda_2 = \kappa_1 \sin \mu$  و همچنین  $j = 1$  باشد. پس توزیع خاصیت فرم پایایی دارد و  $\Theta^w \sim vM(\mu, \kappa_1 + \kappa_2)$ .

اثبات. به صورت مستقیم و به کارگیری (۱۰) در (۹) به دست می‌آید.

لم ۲۰۳. فرض کنید  $\Theta \sim vM(\mu_1, \kappa_1)$  و تابع وزن (۱۰) با  $\lambda_1 = \kappa_1 \cos \mu_1$  و  $\lambda_2 = \kappa_1 \sin \mu_1$  و همچنین  $j = 2$  باشد. پس توزیع ون میزس تعمیم یافته  $\Theta^w \sim GvM(\mu_1, \mu_2, \kappa_1, \kappa_2)$ ، به صورت

$$f^w(\theta; \eta) = \frac{\exp\{\kappa_1 \cos(\theta - \mu_1) + \kappa_2 \cos 2(\theta - \mu_2)\}}{2\pi G_0(\delta, \kappa_1, \kappa_2)} \quad (11)$$

خواهد داشت که در آن  $\eta = (\mu_1, \mu_2, \kappa_1, \kappa_2)$  و  $\delta = (\mu_1 - \mu_2) \bmod \pi$ ،  $\kappa_1, \kappa_2 > 0$ ،  $\mu_1, \mu_2 \in [0, 2\pi)$ .

اثبات. به صورت مستقیم به دست می‌آید.

با تعریف تابع وزن  $w(\theta, \mu, \lambda_1, \lambda_2) = 1 + \lambda_1 \sin j(\theta - \mu) + \lambda_2 \cos j(\theta - \mu)$  و محاسبه  $E[w(\Theta, \mu, \lambda_1, \lambda_2)]$ ، توزیع ون میزس وزنی به صورت

$$f^w(\theta; \eta) = \frac{[1 + \lambda_1 \sin j(\theta - \mu) + \lambda_2 \cos j(\theta - \mu)]}{2\pi(I_0(\kappa) + \lambda_2 I_j(\kappa))} \times \exp\{\kappa \cos(\theta - \mu)\}. \quad (12)$$

به دست می‌آید. در اینجا  $\eta = (\mu, \lambda_1, \lambda_2)$  و  $I_j(\kappa) = \int_0^{2\pi} \cos(j\theta) \exp\{\kappa \cos \theta\} d\theta$  تابع بسل اصلاح شده نوع اول مرتبه  $j$  است.

## ۴ برآورد پارامترهای توزیع ون میزس وزنی

فرض کنید  $\theta^w = (\theta_1^w, \dots, \theta_n^w)$  نمونه تحقق یافته از توزیع ون میزس (۷) و با وزن کلی تر  $w(\theta, \lambda)$  در دسترس باشد. پس:

$$f^w(\theta; \kappa, \lambda) = \frac{\prod_{i=1}^n w(\theta_i, \lambda) \exp\{\kappa \sum_{i=1}^n \cos(\theta_i - \mu)\}}{(2\pi I_0(\kappa))^n E^n[w(\theta, \lambda)]}. \quad (13)$$

<sup>3</sup>Location parameter

<sup>4</sup>Concentration parameter

<sup>5</sup>Modified Bessel function of the first kind

و چند مدی را پوشش می‌دهد. (شکل (۱) حالت‌های خاصی از این توزیع ون میزس وزنی را نشان می‌دهد).

۳. اگر تابع وزن به صورت

$$w(\theta, \lambda_1, \lambda_2) = \exp\{\lambda_1 \cos j\theta + \lambda_2 \sin j\theta\} \quad (5)$$

و توزیع پایه  $\Theta$  توزیع ماکسیمف ([۱۴])

$$f^w(\theta; \lambda_1, \lambda_2) \propto \exp\left\{\sum_{i=1}^k \lambda_{1i} \cos j\theta + \lambda_{2i} \sin j\theta\right\}, \quad (6)$$

باشد، توزیع وزنی به صورت خانواده توزیع‌های ماکسیمف و خانواده نمایی است. در حقیقت توزیع دایره‌ای وزنی دارای فرم پایایی است.

هرگاه ویژگی پایایی در یک توزیع و یا یک خانواده از توزیع‌ها برقرار باشد، استنباط آماری تحت نمونه‌گیری وزنی مشابه توزیع پایه و فقط با اندکی تفاوت در مقادیر پارامترها خواهد بود.

## ۳ توزیع ون میزس وزنی

توزیع ون میزس نقش مهمی در داده‌های دایره‌ای ایفا می‌کند. چنین نقشی را توزیع نرمال در داده‌های خطی بازی می‌کند؛ بنابراین نتایج حاصل از آن در نمونه‌گیری وزنی می‌تواند بسیاری از کاربردهای دایره‌ای را پوشش دهد. تابع چگالی ون میزس که با نماد  $vM(\mu, \kappa)$  نشان داده می‌شود، به صورت

$$f(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}, \quad 0 \leq \theta < 2\pi \quad (7)$$

تعریف و در آن  $\mu \in [0, 2\pi)$  پارامتر مکان<sup>۳</sup> و  $\kappa > 0$  پارامتر تمرکز<sup>۴</sup> است.  $I_0(\kappa)$  تابع بسل اصلاح شده نوع اول<sup>۵</sup> مرتبه صفر و از رابطه

$$I_0(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} e^{\kappa \cos \theta} d\theta \quad (8)$$

به دست می‌آید ([۲]). با توجه به اهمیت توزیع ون میزس ادامه نتایج را بر روی این توزیع دایره‌ای معطوف می‌کنیم.

فرض کنید  $\Theta^w \sim vM(\mu, \kappa)$  و  $\Theta$  مشاهده تحقق یافته با وزن  $w(\theta)$  باشد. در این صورت توزیع ون میزس وزنی به صورت

$$f^w(\theta; \mu, \kappa) = \frac{w(\theta, \beta) \exp\{\kappa \cos(\theta - \mu)\}}{2\pi I_0(\kappa) E[w(\Theta, \beta)]}. \quad (9)$$

لگاریتم تابع درست‌نمایی به صورت

$$\ell^w(\kappa, \lambda) = \sum_{i=1}^n \ln(w(\theta_i; \lambda)) + \sum_{i=1}^n \kappa \cos(\theta_i - \mu) - n \ln(\Upsilon \pi I_0(\kappa)) - n \ln(E[w(\Theta; \lambda)]). \quad (14)$$

و بنابراین برآوردهای ماکسیمم درست‌نمایی برای پارامترهای  $\mu$  و  $\kappa$  از دو رابطه

$$\sum_{i=1}^n \frac{\partial}{\partial \mu} \ln w(\theta_i, \lambda) + \kappa \sum_{i=1}^n \sin(\theta_i - \mu) - n \frac{\partial}{\partial \mu} \ln(E[w(\Theta, \lambda)]) = 0, \quad (15)$$

$$\sum_{i=1}^n \frac{\partial}{\partial \kappa} \ln(w(\theta_i; \lambda)) + \sum_{i=1}^n \cos(\theta_i - \mu) - n \frac{I_1(\kappa)}{I_0(\kappa)} - n \frac{\partial}{\partial \kappa} \ln(E[w(\Theta, \lambda)]) = 0, \quad (16)$$

به دست می‌آیند. در حالت کلی معادلات بالا جواب تحلیلی ندارند و باید با روش‌های عددی به دست آیند.

از آنجاکه به‌طور معمول تابع وزن از پارامتر مورد نظر مستقل است. جمله‌های اول دو معادله (۱۵) و (۱۶) برابر صفر و معادلات به صورت

$$\kappa \sum_{i=1}^n \sin(\theta_i - \mu) - n \frac{\partial}{\partial \mu} \ln(E[w(\Theta, \lambda)]) = 0 \quad (17)$$

$$\sum_{i=1}^n \cos(\theta_i - \mu) - n \frac{I_1(\kappa)}{I_0(\kappa)} - n \frac{\partial}{\partial \kappa} \ln(E[w(\Theta, \lambda)]) = 0 \quad (18)$$

تغییر می‌یابند. از حل این دو معادله

$$\begin{cases} \hat{\kappa} = \frac{\frac{\partial}{\partial \mu} \ln(E[w(\Theta, \lambda)])}{S \sin \mu - C \cos \mu} \\ \frac{I_1(\hat{\kappa})}{I_0(\hat{\kappa})} + \frac{\partial}{\partial \kappa} \ln(E[w(\Theta, \lambda)]) = C \cos \mu + S \sin \mu. \end{cases}$$

که در آن  $\bar{C} = \frac{1}{n} C$ ,  $\bar{S} = \frac{1}{n} S$  و  $C = \sum_{i=1}^n \cos \theta_i$ ,  $S = \sum_{i=1}^n \sin \theta_i$  اگر  $E[w(\Theta)]$  مستقل از پارامترها باشد، برآوردهای ماکسیمم درست‌نمایی وزنی و تصادفی یکسان خواهند شد.

## ۵ برآورد ماکسیمم درست‌نمایی با تابع وزن‌های خاص

در این بخش به برآورد پارامترهای توزیع ون میزس تحت چند تابع وزن خاص پرداخته می‌شود.

۱. فرض کنید  $\Theta \sim vM(\mu_1, \kappa_1)$  و تابع وزن نمایی

$$w(\theta, \lambda_1, \lambda_2) = \exp\{\lambda_1 \cos \Upsilon \theta + \lambda_2 \sin \Upsilon \theta\}$$

با  $\lambda_2 = \kappa_2 \sin \Upsilon \mu_2$  و  $\lambda_1 = \kappa_2 \cos \Upsilon \mu_2$  در آن صورت توزیع ون میزس تعمیم‌یافته  $\Theta^w \sim GvM(\mu_1, \mu_2, \kappa_1, \kappa_2)$  حاصل می‌شود. برآورد پارامترهای توزیع ون میزس تعمیم‌یافته توسط [۷] به دست آمده است.

۲. فرض کنید برای  $\Theta \sim vM(\mu, \kappa_1)$  تابع وزن نمایی (۱۰)

با  $\lambda_2 = \kappa_2 \sin \mu$  و  $\lambda_1 = \kappa_2 \cos \mu$  در نظر گرفته شود. پس  $\Theta^w \sim vM(\mu, \kappa_1 + \kappa_2)$ . برآورد پارامترها از حل معادلات (۱۵) و (۱۶) منجر به  $\hat{\mu} = \bar{\theta}$  و اگر  $\kappa_2$  معلوم باشد  $A(\kappa) = \frac{I_1(\kappa)}{I_0(\kappa)}$  در اینجا  $\hat{\kappa}_1 = A^{-1}(\bar{R}) - \kappa_2$  و  $\bar{R} = \frac{1}{n} \sqrt{C^2 + S^2}$  که به‌سادگی از مقادیر نمونه به دست می‌آیند.

اگر  $\kappa_2$  مجهول باشد، برآوردهای  $\kappa_1$  و  $\kappa_2$  غیرقابل تشخیص هستند. در نتیجه برآوردها غیر یکتا و غیر مفید خواهند بود. کمک گرفتن از توزیع‌های پیشین روی پارامترها و منجر شدن به برآوردهای ماکسیمم درست‌نمایی تعمیم‌یافته توسط [۱۸] در داده‌های خطی پیشنهاد شده است. با به‌کارگیری این روش و انتخاب توزیع پیشین

$$\begin{aligned} \pi(\kappa_1, \kappa_2) &= \pi_0(\kappa_2 | \kappa_1) \pi_1(\kappa_1) \\ &= \kappa_1 e^{-\kappa_1 \kappa_2} e^{-\kappa_1} = \kappa_1 e^{-\kappa_1(\kappa_2 + 1)}, \quad \kappa_1, \kappa_2 > 0 \end{aligned}$$

توزیع پسین به صورت زیر حاصل می‌شود:

$$GL(\kappa_1, \kappa_2, \mu) \propto \frac{\kappa_1 \exp\{(\kappa_1 + \kappa_2) \sum_{i=1}^n \cos(\theta_i - \mu) - \kappa_1(\kappa_2 + 1)\}}{\Upsilon \pi I_0(\kappa_1 + \kappa_2)^n}$$

و برآورد پارامترها با ماکسیمم سازی تابع  $GL(\kappa_1, \kappa_2, \mu)$  نسبت به سه پارامتر  $\mu$ ,  $\kappa_1$  و  $\kappa_2$  به دست می‌آید. با مشتق‌گیری از  $GL(\kappa_1, \kappa_2, \mu)$  نسبت به سه پارامتر  $\mu$ ,  $\kappa_1$  و  $\kappa_2$  و برابر صفر قرار دادن آن‌ها داریم:

$$\begin{aligned} (\kappa_1 + \kappa_2) \sum_{i=1}^n \sin(\theta_i - \mu) = 0 &\implies \hat{\mu} = \bar{\theta} \\ \left(\frac{1}{\kappa_1} - \kappa_2 - 1\right) + \sum_{i=1}^n \cos(\theta_i - \mu) &= 0 \\ -n \frac{I_1(\kappa_1 + \kappa_2)}{I_0(\kappa_1 + \kappa_2)} = 0 & \quad (19) \\ -\kappa_1 + \sum_{i=1}^n \cos(\theta_i - \mu) - n \frac{I_1(\kappa_1 + \kappa_2)}{I_0(\kappa_1 + \kappa_2)} = 0 & \quad (20) \end{aligned}$$

$\lambda_1 = \lambda_2 = 1$  که در آن  $w(\theta, \lambda_1, \lambda_2) = \exp\{\lambda_1 \cos 2\theta + \lambda_2 \sin 2\theta\}$  به تعداد ۱۰۰۰۰ بار انتخاب شد. از آنجاکه  $\Theta^w \sim GVM(2, 2, 1, 1)$  انتظار می‌رود برآورد پارامترها برای این مدل نسبت به مدل‌های با وزن‌های دیگر برازش بهتری را نشان دهد. در این مرحله، نمونه صدتایی از توزیع ون میزس با وزن بالا را به تعداد ۱۰۰۰۰ بار شبیه‌سازی کرده و برآورد ماکسیمم درست‌نمایی پارامترها را مطابق بخش قبلی به دست آورده شد. جهت مقایسه دو وزن دیگر نیز مورد بررسی قرار گرفت. وزن اصلی  $w_2(\theta, \lambda_1, \lambda_2)$  و دو وزن دیگر به صورت

$$w_1(\theta, \lambda_1, \lambda_2) = \exp\{\lambda_1 \cos \theta + \lambda_2 \sin \theta\},$$

$$\lambda_1 = \kappa_2 \cos \mu_1, \quad \lambda_2 = \kappa_2 \sin \mu_1$$

$$w_2(\theta, \lambda_1, \lambda_2) = \exp\{\lambda_1 \cos 2\theta + \lambda_2 \sin 2\theta\},$$

$$\lambda_1 = \kappa_2 \cos 2\mu_2, \quad \lambda_2 = \kappa_2 \sin 2\mu_2$$

$$w_3(\theta, \lambda_1, \lambda_2) = 1 + \lambda_1 \sin(\theta - \mu_1) + \lambda_2 \cos(\theta - \mu_1)$$

بوده‌اند. همچنین برای مقایسه وزن‌ها از معیارهای ماکسیمم لگاریتم تابع درست‌نمایی (MLL)، ضریب آکائیکه (AIC) و شوارتز (BIC) استفاده شد. نتایج در جدول (۱) نشان می‌دهد مدل واقعی بهتر از دیگر مدل‌ها به داده‌ها برازش می‌شود. برآوردها نیز به مقدار واقعی آن‌ها نزدیک است.

## ۷ بحث و نتیجه‌گیری

در این مقاله توزیع‌های وزنی در داده‌های دایره‌ای مطالعه شده‌اند. هرگاه مشاهدات همراه با یک تابع از مشاهدات در دسترس قرار گیرند، متغیر تصادفی جدید متفاوت با توزیع پایه خواهد بود. در داده‌های دایره‌ای برخلاف داده‌های خطی که هرگاه تابع وزن غیر منفی باشد متغیر تصادفی توزیع جدیدی خواهد داشت، در این نوع از داده‌ها ممکن است توزیع جدید خاصیت تناوبی نداشته باشد.

در این مقاله شرط کافی برای متناوب بودن تابع توزیع وزنی ارائه شده است. توزیع وزنی برای داده‌های توزیع مهم ون میزس به دست آمده و سه تابع وزن پیشنهاد شده است. برآورد ماکسیمم درست‌نمایی پارامترهای ون میزس برای هرکدام از توابع وزن به دست آمده و در یک شبیه‌سازی روش محاسبه برآوردها مورد ارزیابی قرار گرفته و نشان داده شده است، با روش پیشنهادی می‌توان هرگاه داده‌ها غیر تصادفی و همراه با یک وزن در اختیار قرار می‌گیرند پارامترها را برآورد کرد. ضمن اینکه می‌توان

با جایگذاری  $\bar{\theta}$  به جای  $\mu$  و ساده کردن معادلات (۱۹) و (۲۰) برآوردهای  $\kappa_1$  و  $\kappa_2$  با روش عددی از دو رابطه

$$\left(\frac{1}{\kappa_1} - \kappa_2 - 1\right) - n \frac{I_1(\kappa_1 + \kappa_2)}{I_0(\kappa_1 + \kappa_2)} = C \cos(\bar{\theta}) + S \sin(\bar{\theta})$$

$$\kappa_1 + n \frac{I_1(\kappa_1 + \kappa_2)}{I_0(\kappa_1 + \kappa_2)} = C \cos(\bar{\theta}) + S \sin(\bar{\theta})$$

به دست می‌آیند؛ بنابراین از بین جواب‌های رابطه

$$\frac{1}{\kappa_1} - \kappa_1 - 1 = \kappa_2$$

که در (۱۹) صدق کنند، برآوردها محاسبه خواهند شد.

۳. در حالتی که تابع وزن به صورت  $w(\theta, \mu) = 1 + \lambda_1 \sin j(\theta - \mu) + \lambda_2 \cos j(\theta - \mu)$  باشد، لگاریتم تابع درست‌نمایی:

$$\begin{aligned} \ell^w(\mu, \kappa, \lambda_1, \lambda_2) &= \sum_{i=1}^n \ln[1 + \lambda_1 \sin j(\theta_i - \mu) + \lambda_2 \cos j(\theta_i - \mu)] \\ &+ \kappa \sum_{i=1}^n \cos(\theta_i - \mu) - n \ln[\pi(I_0(\kappa) + \lambda_2 I_j(\kappa))]. \end{aligned}$$

خواهد بود. در این حالت روش‌های عددی برای برآورد چهار پارامتر  $(\lambda_1, \lambda_2, \mu, \kappa)$  به کار می‌رود. باید یادآوری کرد جهت محاسبه برآورد پارامترها، علاوه بر شرط (۴) استفاده از بسته نرم‌افزاری *Rsolnp* در نرم‌افزار *R* ضروری است. هرگاه پارامترهای  $\lambda_1$  و  $\lambda_2$  معلوم هستند، برآوردهای  $\mu$  و  $\kappa$  به صورت عددی از دو معادله

$$\begin{aligned} \sum_{i=1}^n \frac{\lambda_1 \sin j(\theta_i - \mu) - \lambda_2 \cos j(\theta_i - \mu)}{1 + \lambda_1 \sin j(\theta_i - \mu) + \lambda_2 \cos j(\theta_i - \mu)} \\ + \kappa \sum_{i=1}^n \sin(\theta_i - \mu) = 0 \\ \sum_{i=1}^n \cos(\theta_i - \mu) - n \frac{\lambda_1 [A_{j+1}(\kappa) - A_1(\kappa)A_j(\kappa)]}{1 + \lambda_2 A_j(\kappa)} = 0. \end{aligned}$$

به دست می‌آیند که  $A_j(\kappa) = \frac{I_j(\kappa)}{I_0(\kappa)}$  است.

## ۶ شبیه‌سازی

در این بخش با استفاده از شبیه‌سازی رد - پذیرش توزیع‌های ون میزس وزنی تحت تابع وزن‌های مختلف را مقایسه می‌کنیم. فرض کنید  $\theta^w = (\theta_1^w, \dots, \theta_n^w)$  نمونه‌ای وزنی از  $\Theta \sim VM(\mu_1, \kappa_1)$  با  $\mu_1 = 2$  و  $\kappa_1 = 1$  باشد. نمونه‌ای تصادفی  $100$  تایی از توزیع وزنی با تابع وزن

در نتیجه وزن‌های متفاوتی حاصل خواهد شد. چنین مطالعه‌ای می‌تواند پیش و روی پژوهشگران قرار گیرد.

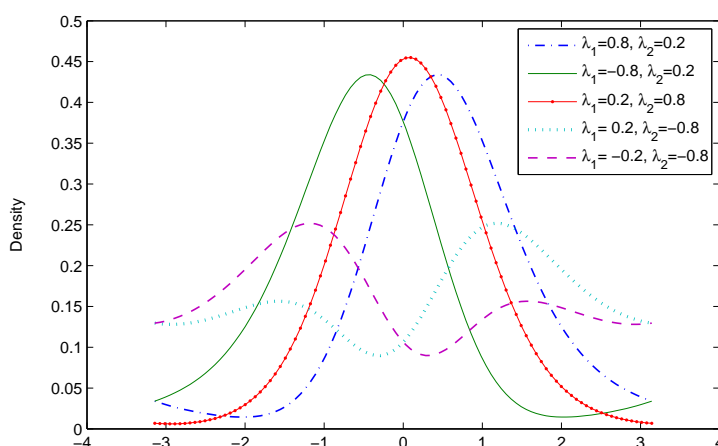
تابع وزن نامعلومی که داده‌ها از آن طریق به دست آمده‌اند را از تابع وزن‌های دیگر تشخیص داد.

به دست آوردن توزیع‌های وزنی می‌تواند به عنوان روشی برای تولید توزیع‌های جدید از توزیع پایه نیز به کار گرفته شود.

اگرچه در بخش اول مقاله توزیع وزنی دایره‌ای را در حالت کلی مورد بحث قرار داده است، اما نتایج اصلی مقاله توزیع ون میزس را مدنظر قرار داده است. بدیهی است برای توزیع‌های دیگر دایره‌ای نتایج متفاوت و

## تقدیر و تشکر

نویسندگان مقاله نهایت تشکر خود را از داوران و ویراستار محترم مجله ابراز می‌دارند.



شکل ۱: تابع چگالی ون میزس وزنی به ازای مقادیر مختلف  $\lambda_1$  و  $\lambda_2$  زمانی که  $j = 1$  و  $k = 1$ .

جدول ۱: برآورد پارامترهای توزیع ون میزس تحت سه تابع وزن با معیارهای ماکسیمم لگاریتم تابع درست‌نمایی، ضریب آکائیکه و شوارتز برای داده‌های شبیه‌سازی شده از توزیع ون میزس وزنی تحت تابع وزن  $w_2(\theta)$

تابع وزن	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	MLL	AIC	BIC
$w_1(\theta)$	۰/۱/۲	-	۰۲/۱	۴۷۸/۰	-	-	-۰۵۳/۱۴۴	۱۰/۲۹۲	۵۲/۳۰۶
$w_2(\theta)$	۹۹/۱	۰۰/۲	۰۳/۱	۰۲/۱	-	-	-۰۳/۱۲۸	۰۶/۲۶۴	۴۸/۲۷۴
$w_3(\theta)$	۹۹/۱	-	۰۷/۱	-	-۰۰۳/۰	۷/۰	-۸۸/۱۲۸	۷۶/۲۶۵	۱۸/۲۷۶

## مراجع

[۱] علوی، س.م.ر؛ و چینی پرداز، ر (۱۳۸۶)، استنباط در توزیع نرمال بر اساس نمونه‌گیری وزنی، مجله علوم آماری، ۱(۱)، ۷۳-۸۸.

[2] Abramowitz, M. and Stegun, I.A. (Eds.) (1972). *Handbook of Mathematical Functions*, Dover, New York.

- [3] Arnold, B.C., and SenGupta, A. (2006). Recent advances in the analyses of directional data in ecological and environmental sciences. *Environmental and Ecological Statistics*, **13(3)**, 253-256.
- [4] Bulbert, M.W., Page, R.A., and Bernal, X.E. (2015). Danger comes from all fronts: predator-dependent escape tactics of Túngara frogs. *PLoS One*, **10(4)**, e0120546.
- [5] Chen, S. X., and Cowling, A. (2001), Measurement Error in Line Transect Surveys where Detectability Varies with Distance and Size, *Biometrics*, **57**, 732-742.
- [6] Fisher, R. A. (1934), The Effect of Models of Ascertainment upon the Estimation of Frequencies., *Annal Eugen*, **6**, 13-25.
- [7] Gatto, R. and Jammalamadaka, S.R. (2007). The generalized von Mises distribution. *Stat Methodol*, **4**, 341-353.
- [8] Gill, J., and Hangartner, D. (2010). Circular data in political science and how to handle it. *Political Analysis*, **18(3)**, 316-336.
- [9] Gurtman, M.B. (2009). Exploring personality with the interpersonal circumplex. *Social and Personality Psychology Compass*, **3(4)**, 601-619.
- [10] Gurtman, M.B., and Pincus, A.L. (2003). *The circumplex model: Methods and research applications*. In: *Handbook of Psychology*. In: Handbook of Psychology. Wiley Online Library.
- [11] Jacobo, D. (2003), Large Sample Results under Biased Sampling when Covariables are Present, *Statistics and Probability Letters*, **63**, 287-293.
- [12] Kaas, A.L., and Van Mier, H.I. (2006). Haptic spatial matching in near peripersonal space. *Experimental Brain Research*, **170(3)**, 403-413.
- [13] Leary, T. (1957). *Interpersonal Diagnosis of Personality*, Ronald Press, New York.
- [14] Maksimov, V. T. (1967). Necessary and sufficient to sufficient statistics for the family of shifts probability distributions on continuous bicomact groups *Theoria Veroyanta*, **12**, 307-321.
- [15] Mardia, K.V. (1972). *Statistics of Directional Data*, Academic Press, New York.
- [16] Mardia, K.V., (2011). How new shape analysis and directional statistics are advancing modern life-sciences. In: *International Statistical Institute: Proc, 58th World Statistical Congress*
- [17] Mardia, K.V., Hughes, G., Taylor, C.C., and Singh, H., (2008). A multivariate von Mises distribution with applications to bioinformatics. *Canadian Journal of Statistics*, **36(1)**, 99-109.
- [18] Martins, E. S. and Stedinger, J. R. (2000), Generalized maximum likelihood GEV quantile estimators for hydrologic data. *Water Resource Res.*, **36**, 737-744.
- [19] Morrison, D. G. (1973), Some results for waiting times with an application to survey data., *American Statistician*, **27**, 226-227.
- [20] Rao, C.R. (1965). On discrete distributions arising out of methods of ascertainment. *Sankhya*, **27**, 311-324.
- [21] Ye, Y. (1967). *Interior Algorithms for Linear, Quadratic, and Linearly Constrained non-linear programming*. Ph. D. thesis, Department of ESS, Stanford University.

- [22] Unnikrishnan Nair, N., and Sunoj, S. M. (2003), Form-Invariant Bivariate Weighted Models., *Statistics*, **37(3)**, 259-269.
- [23] Warren, W. G.(1975), Statistical Distributions in Forestry and Forest Products Research, in *Statistical Distributions in Scientific Work*, Patil, G. P., Kotz, S., Ord, K. eds., **2**, 369-384.

## Statistical Inference for von Mises Circular Distribution under Weighted Sampling

Fatemeh Shahsanaei<sup>1</sup> and Rahim Chinipardaz<sup>2</sup>

Abstract:

In many scientific fields, measurements are directional. Research into the characteristics of these types of data requires the use of new statistical tools. Statistical methods for analyzing these kinds of data are known as nonlinear statistics. Circular statistics is an example of this field of statistics. In the study of circular statistics, sampling may be such that each observation is recorded with a probability proportional to a non-negative function of it. Such a sample is called a weighted circular sample and the corresponding distribution of these samples is called a weighted circular distribution. This paper addresses weighted circular distributions. Considering that the von Mises distribution is widely used in modeling circular data, the maximum likelihood estimation of the parameters of the weighted von Mises circular distribution is examined. Since the estimators may not be identifiable, unique Bayesian estimates are obtained using a suitable prior distribution. In a simulation study, different weights in the von Mises circular distribution are compared.

**Keywords:** Circular data, von Mises distribution, Weight function, Maximum likelihood estimation, Accept-Rejection simulation.

---

<sup>1</sup>Department of Electrical Engineering, Shohadaye Hoveizeh Campus of Technology Shahid Chamran University of Ahvaz, Iran

<sup>2</sup> Department of Statistics, Shahid Chamran University of Ahvaz, Ahvaz, Iran