سیزدهمین کنفرانس آمار ایران
۲-۴ شهریور ۱۳۹۵

آماره کاوشی فضایی انعطاف‌پذیر

فاطمه رهایی[۱]، امید کریمی[۱]
[۱] گروه آمار، دانشگاه سمنان

چکیده: آماره‌های کاوشی فضایی به طور گسترده ابزارهای استفاده برای آشکارسازی خوشه‌های بیمار هستند. به ویژه آماره کاوشی فضایی دایره‌ای مطرح شده توسط کالدورف (۱۹۹۷) و آماره کاوشی فضایی انعطاف‌پذیر به وسیله تانگو و تاکاهاشی (۲۰۰۵) که دارای محدودیت‌هایی هستند. در این مقاله، یک آماره کاوشی فضایی انعطاف‌پذیر به وسیله نسبت درستنمایی مقید که توسط تانگو (۲۰۰۸) مطرح شده است، را با دو شرط بیان می‌کنیم: ۱. حذف محدودیت ماکزیمم، ۳۰ همسایگی نزدیک ۲. کاهش زمان محاسباتی قابل توجه نسبت به آماره کاوشی فضایی انعطاف‌پذیر اصلی. این آماره، خوشه را با هر شکل منطقی آشکار می سازد، اما به همان نسبت ریسک نسبی خوشه را از طریق شبیه‌سازی مونت کارلو بزرگ می‌کند. در انتها آماره کاوشی فضایی روی داده‌های مرگ و میر بیماران سکته مغزی در منطقه متروپلیتان توکیو به کار گرفته شده است.

واژه‌های کلیدی: آماره کاوشی فضایی، آشکارسازی خوشه، آماره نسبت درستنمایی مقید.

کد موضوع‌بندی ریاضی (۲۰۱۰): 62H11، 62H30، 62M30.

۱ مقدمه

بسیاری از نویسندگان نظریه‌هایی در چندین کتاب ارائه کردند و تعدادی از روش‌های آماری مختلف را برای آشکارسازی خوشه‌های بیمار نشان دادند. (به عنوان نمونه؛ لاوسن و همکاران (۱۹۹۹) ، والر و گوتوی (۲۰۰۴) و تانگو (۲۰۱۰)). آماره کاوشی فضایی به وسیله کالدورف و ناگاوالا (۱۹۹۵) و کالدورف (۱۹۹۷) به همراه نرم‌افزار SatScan پیشنهاد شده است که در تنوع وسیعی از مطالعات اپیدمیولوژیک و هم‌چنین در مراقبت بیماری برای آشکارسازی خوشه‌های بیمار قابل اجرا می‌باشد. آماره کاوشی فضایی سعی می‌کند تا محتمل ترین خوشه (MLC) [۱] را آشکار سازد که عبارت است از مجموعه‌ای از ناحیه‌های به هم مرتبط که دارای ماکزیمم نسبت درستنمایی هستند. به دلیل استفاده ناحیه دایره‌ای برای کاوش مناطق خوشه‌های ممکن، در آشکارسازی درست خوشه‌های غیر دایره‌ای واقعی مشکلاتی وجود دارد. برای شناسایی خوشه‌های غیر دایره‌ای که به وسیله آماره کاوشی فضایی دایره‌ای قابل آشکارسازی

[۱] فاطمه رهایی : rahaei.pnu@gmail.com

[۱] Cluster Likely Most

نیستند، تعدادی از نویسندگان داکزمال و اسانکو (۲۰۰۴) ، پتلی و تیلی (۲۰۰۴) ، تانگو و تاکاهاشی (۲۰۰۵) ، اسانکو و همکاران (۲۰۰۶) نظریه‌های مختلفی برای آماره‌های کاوشی فضایی پیشنهاد کردند. همچنین کالدورف و همکاران (۲۰۰۶) نسخه بیضوی از آماره کاوشی فضایی را مطرح کرده‌اند. علاوه بر این، آماره کاوشی فضایی انعطاف‌پذیر به وسیله تانگو و تاکاهاشی (۲۰۰۵) به همراه نرم‌افزار FlexScan تاکاهاشی و همکاران (۲۰۱۰) بیان شده است که برای آشکارسازی خوشه‌های نامنظم نیز استفاده می‌شود. هر چند محدودیت روش آن‌ها به دلیل بار محاسباتی سنگین فقط ماکزیمم، تا ۳۰ همسایگی نزدیک را برای تعیین خوشه‌ها در نظر می‌گیرد. چندین نویسنده‌ای که مقایسه‌ای از آزمون‌ها را برای خوشه‌های بیمار بیان کردند، این محدودیت را نشان دادند. به عنوان نمونه؛ هوانگ و همکاران (۲۰۰۸) . از طرف دیگر، تانگو و تاکاهاشی (۲۰۰۵) نشان دادند که این آماره کاوشی فضایی براساس ماکزیمم کردن نسبت درستنمایی معمولی منجر به آشکارسازی یک MLC ای می شود که به دلیل جذب ناحیه‌های همسایه با ریسک غیرمرتفع رخداد بیماری، خیلی بزرگ‌تر از خوشه درست است. برای اجتناب از این خاصیت نامطلوب، تانگو (۲۰۰۸) یک آماره کاوشی فضایی جدید به وسیله مقید کردن نسبت درستنمایی که فقط ناحیه‌ای با خطر مرتفع را کاوش می کند، مطرح کرد. تانگو (۲۰۰۸) انجام آشکارسازی خوشه‌های دایره‌ای را به وسیله دو روش مقایسه کرد: آماره کاوشی فضایی دایره‌ای کالدورف و آماره کاوشی فضایی دایره‌ای انجام شده با نسبت درستنمایی مقید مطرح شده از طریق مطالعه شبیه‌سازی مونت کارلو. نتایج نشان داد که روش دوم توانایی آشکارسازی کل و یا یک قسمتی از خوشه درست را بیشتر از اولی دارد. در این مقاله، یک آماره کاوشی فضایی انعطاف‌پذیر انجام شده با نسبت درستنمایی مقید را مطرح می کنیم. انجام آماره کاوشی مطرح شده با نسبت درستنمایی مقید با آماره کاوشی فضایی انعطاف‌پذیر اصلی تانگو و تاکاهاشی و آماره کاوشی فضایی دایره‌ای در بخش ۲ مورد بررسی قرار می‌گیرند. در بخش ۳ روش مطرح شده روی داده های میزان مرگ و میر بیماران سکته مغزی در مناطقی از متروپلیس توکیو و حوزه اداری کاناگاوا در ژاپن مورد مطالعه قرار می‌گیرد.

۲ آماره کاوشی فضایی انعطاف‌پذیر

آماره کاوشی فضایی برای اهداف بهداشت عمومی جامعه در پیدا کردن خوشه‌های فضایی معنی‌دار، بروز یا شیوع بیماری و همچنین برای شناسایی عوامل خطر زیست محیطی به کار می‌رود و تعمیم حالت دو بعدي یا چند بعدي آماره کاوشي است. آماره کاوشی فضایی سعی می‌کند تا محتمل‌ترین خوشه (MLC) را آشکار سازد که عبارت است از مجموعه‌ای از ناحیه‌های به هم مرتبط که دارای ماکزیمم نسبت درستنمایی هستند. فرض کنید تمام مناطق مطالعه شده به m ناحیه قابل شمارش تقسیم شده است. تعداد حالت ها در ناحیه i به وسیله متغیر تصادفی N_i با مقدار مشاهده شده $n_i(i=1,\cdots,m)$ و $n=n_1+\cdots+n_m$ نشان داده می شود. تحت فرض H_0 ، N_i متغیر مستقل پواسون به صورت

$$H_0:\mathrm{E}(N_i)=\xi_i,\quad N_i\sim Poisson(\xi_i),\quad i=1,\cdots,m \qquad (۱.۲)$$

است که در آن $Poisson(\xi_i)$ ، توزیع پواسون با میانگین ξ_i است و ξ_i تعداد موارد مورد انتظار ξ_i در ناحیه i تحت فرض صفر است. برای برآورد تعدیل تعدادی از موارد مورد انتظار برای متغیرهای کمکی ممکن مثل سن، می‌توانیم از استانداردسازی غیر مستقیم و یا یک مدل رگرسیونی پواسون استفاده کنیم. ξ_i به صورت

$$\xi_i=n\frac{w_i}{\sum_{k=1}^{m}w_k},\qquad i=1,\cdots,m \qquad (۲.۲)$$

محاسبه می شود، که در آن w_i اندازه جامعه در ناحیه i ام است. برای مشخص کردن موقعیت جغرافیایی هر ناحیه از مختصات مرکز جامعه اداری استفاده می‌کنیم.

آماره کاوشی فضایی کالدورف براي آشکارسازی مناطقی مورد استفاده قرار می‌گیرد که ریسک q_i (ریسک مورد نظر در مناطق) در داخل منطقه مورد بررسی به طور معنی‌داري بیشتر از ریسک خارج از منطقه مورد بررسی است. آماره کاوشی فضایی کالدورف، یک پنجره دایره‌ای Z روی مرکز هر ناحیه رسم می‌کند. برای هر یک از این مراکز، شعاع دایره به طور پیوسته از صفر تا ۵۰ درصد از ریسک جامعه تغییر می‌کند. اگر پنجره شامل مرکز ناحیه باشد، پس تمام ناحیه در پنجره قرار می‌گیرد. در مجموع تعداد زیاد و متفاوت پنجره‌ها باعث تداخل پنجره‌های دایره‌ای می‌شود که هر یک از آنها ممکن است موقعیت و اندازه متفاوتی در هر خوشه داشته باشد. فرض کنید $Z_{ik}(k = ۱, \cdots, K_i)$ به پنجره ساخته شده از $(k-۱)$ همسایه نزدیک به ناحیه i ام دلالت کند. پس همه پنجره‌ها به وسیله آماره کاوشی فضایی بررسی شده که به صورت زیر تعریف می‌شود:

$$z = z_۱ = \{Z_{ik} | ۱ \leq i \leq m \quad ۱ \leq k \leq K_i\} \qquad (۳.۲)$$

آماره کاوشی فضایی انعطاف‌پذیر تانگو و تاکاهاشی، برای آشکارسازی خوشه‌های نامنظم نیز استفاده می‌شود. هر چند محدودیت این روش، به دلیل بار محاسباتی سنگین، فقط ماکزیمم، تا ۳۰ همسایگی نزدیک را برای تعیین خوشه‌ها در نظر می‌گیرد. از طرف دیگر، تانگو و تاکاهاشی نشان دادند که این آماره کاوشی فضایی براساس ماکزیمم کردن نسبت درستنمایی معمولی منجر به آشکارسازی یک MLC ای می‌شود که به دلیل جذب ناحیه‌های همسایه با ریسک غیرمرتفع رخداد بیماری، خیلی بزرگ‌تر از خوشه درست است. هم‌چنین این آماره، یک پنجره شکلی انعطاف پذیر Z را روی هر مرکز از ناحیه به وسیله اتصال آنها به ناحیه‌های مجاور ایجاد می‌کند. برای چشم‌پوشی از آشکارسازی یک خوشه به شکل عجیب و نامحتمل، ناحیه‌های متصل، به زیرمجموعه‌هایی از مجموعه ناحیه‌های i و k همسایه نزدیک به ناحیه i محدود می‌شود. در مجموع تعداد زیاد و متفاوت پنجره‌ها باعث تداخل پنجره‌های دایره‌ای می‌شود که هر یک از آنها ممکن است موقعیت و اندازه متفاوتی در هر خوشه داشته باشد. فرض کنید $Z_{ik(j)}(j = ۱, \cdots, j_{ik})$ بر پنجره‌ی j ام دلالت کند که مجموعه ای از k ناحیه متصل به ناحیه i است و در آن j_{ik} برابر با تعدادی از j هایی است که در رابطه زیر صدق می‌کند:

$$Z_{ik(j)} \subseteq Z_{ik}, \quad k = ۱, \cdots, K_i = K$$

بنابراین همه‌ی پنجره‌های کاوشگر شامل مجموعه زیر هستند:

$$z = z_۲ = \{Z_{ik(j)} | ۱ \leq i \leq m \quad ۱ \leq k \leq K \quad ۱ \leq j \leq j_{ik}\}$$

برای هر ناحیه داده شده i ، آماره کاوشی فضایی دایره‌ای، به صورت K دایره هم‌مرکز در نظر گرفته می‌شود. در صورتی که فرض می‌شود آماره کاوشی انعطاف‌پذیر، K دایره هم مرکز بعلاوه همه مجموعه ناحیه‌های متصل (شامل ناحیه مفرد i) باشد که درون K تا از بزرگترین دایره‌های هم‌مرکز قرار داده شده‌اند. بنابراین اندازه $z_۲$ بزرگتر از $z_۱$ است که تقریباً $m \times K$ است. به دلیل بار محاسباتی سنگین، ماکزیمم طول K باید کمتر از ۳۰ باشد؛ در غیر اینصورت ممکن است محاسبات، بیشتر از یک روز یا یک هفته با توجه به داده‌ها زمان صرف کند.

۱.۲ آماره نسبت درستنمایی کالدورف

با استفاده از $Z \in z$ ، فرض صفر (۱.۲) به شکل زیر بیان می‌شود:

$$H_0 : \mathrm{E}(N(Z)) = \xi(Z), \quad for\ all\ Z \in z$$

که در آن $N(\cdot)$ تعداد حالات و $\xi(\cdot)$ تعداد حالات مورد انتظار تحت فرض صفر می‌باشد. تحت فرض H_1 حداقل یک پنجره $Z \in z$ وجود دارد که ریسک درون پنجره در مقایسه با بیرون آن، بیشتر است. فرضیه H_1 به شکل زیر بیان می‌شود:

$$H_1 : \mathrm{E}(N(Z)) > \xi(Z), \quad for\ some\ Z \in z$$

برای هر پنجره Z ممکن است به ترتیب درستنمایی مشاهدات برای تعدادی از موارد درون و بیرون پنجره محاسبه شود. تحت فرض توزیع پواسون (۱.۲) آماره نسبت درستنمایی کالدورف به شکل زیر مطرح می‌شود:

$$\lambda_K = \max_{Z\in z} \lambda_K(Z) = \max_{Z\in z} \left(\frac{n(Z)}{\xi(Z)}\right)^{n(Z)} \left(\frac{n-n(Z)}{n-\xi(Z)}\right)^{n-n(Z)} I\left(\frac{n(Z)}{\xi(Z)} > \frac{n-n(Z)}{n-\xi(Z)}\right)$$

که در آن $n(\cdot)$ تعداد موارد مشاهده شده درون پنجره مشخص شده و $I(\cdot)$ تابع نشانگر است. پنجره Z^* که ماکزیمم نسبت درستنمایی را به دست می‌آورد به وسیله MLC تعریف می شود. هر چند به نظر نمی رسد که Z^* خیلی خوب ساخته شده باشد؛ چون آماره کاوشی فضایی با استفاده از نسبت درستنمایی تعریف شده، یک MLC خیلی بزرگتر از خوشه واقعی را به وسیله ناحیه‌های همسایگی با ریسک غیرمرتفع، آسان‌تر آشکار می‌سازد. (تانگو و تاکاهاشی (۲۰۰۵)، تانگو (۲۰۰۰))

تانگو (۲۰۰۸) بیان کرد که در فرایند کاوش پنجره براساس $\lambda_K(Z)$ ، ممکن است دو پنجره مجزا Z_1 و Z_2 و چندین ناحیه $\{i_1\}, \cdots, \{i_r\}$ وجود داشته باشد به طوریکه:

$$\lambda_K(\{Z_1, Z_2, \{i_1\}, \cdots, \{i_r\}\}) > max\{\lambda_K(Z_1), \lambda_K(Z_2)\}$$

است. که در آن

$$\frac{n(Z_1)}{\xi(Z_1)} > 1, \quad \frac{n(Z_2)}{\xi(Z_2)} > 1 \quad and \quad \frac{n_i}{\xi_i} \leq 1 \ (i = 1 \cdots r)$$

به این معنی که اگر ما اجازه دهیم هر ناحیه‌ای برای MLC انتخاب شود، امکان آشکارسازی یک MLC بزرگ غیر واقعی وجود دارد.

۲.۲ آماره نسبت درستنمایی مقید

تانگو (۲۰۰۸) نسبت درستنمایی مقید به وسیله تفسیر خطر ناحیه منحصر به فرد را مطرح کرده است که در زیر نشان داده می‌شود:

$$\lambda_T(Z) = \left(\frac{n(Z)}{\xi(Z)}\right)^{n(Z)} \left(\frac{n-n(Z)}{n-\xi(Z)}\right)^{n-n(Z)} I\left(\frac{n(Z)}{\xi(Z)} > \frac{n-n(Z)}{n-\xi(Z)}\right) \times \prod_{i\in Z} I(p_i < \alpha_1)$$

که در آن p_i ، $p-Value$ یک طرفه آزمون

$$H_0 : \mathrm{E}(N_i) = \xi_i$$

است که به وسیله $P-Value$ میانی به صورت

$$p_i = Pr\{N_i \geq n_i + 1 | N_i \sim Poisson(\xi_i)\} + \frac{1}{2} Pr\{N_i = n_i | N_i \sim Poisson(\xi_i)\}$$

محاسبه می‌شود و α_1 سطح معنی‌داری از قبل تعیین شده برای ناحیه منحصر به فرد است. دلیل استفاده از $P-Value$ میانی، محافظت از تعریف معمولی $P-Value$ برای ξ_i کوچک و تعداد خروجی‌ها می‌باشد. در این فرمول، $I(p_i < \alpha_1)$ ملاک آزمایش در نظر گرفته می‌شود.

تانگو (۲۰۰۸) خواص آماره کاوشی فضایی دایره‌ای با نسبت درستنمایی مقید را از طریق مطالعه شبیه‌سازی مونت‌کارلو تحقیق کرده است و آن نشان می‌دهد که توانایی بهتری برای آشکارسازی خوشه دایره‌ای درست در مقایسه با اصل کالدورف در همه مدل‌های خوشه‌های دایره‌ای مطرح شده دارد. نتایج مطالعه شبیه‌سازی مونت‌کارلو یک راهکار مناسب درخصوص انتخاب α_1 برای یک آماره نسبت درستنمایی مقید از α در سطح ۰.۰۵ به صورت زیر پیشنهاد می‌کند: (۱) $\alpha_1 = 0.10 - 0.20$ که خوشه‌های کوچک با افزایش خطر نوک‌دار را آشکار می‌سازد.

(۲) $\alpha_1 = 0.20 - 0.30$ که خوشه‌های کوچک به اندازه میانی با افزایش مقید در خطر را آشکار می‌سازد.

(۳) $\alpha_1 = 0.30 - 0.40$ که خوشه‌های بزرگتر با یک افزایش ناچیز در خطر را آشکار می‌سازد.

تانگو (۲۰۰۸) $\alpha_1 = 0.20$ را به طور پیش فرض معرفی کرده است.

۳ مثال کاربردی

سه آماره کاوشی فضایی انعطاف‌پذیر اصلی(FS) با $\alpha_1 = 0.2$ ، آماره کاوشی فضایی انعطاف‌پذیر (FSR) و آماره کاوشی فضایی دایره‌ای (CS) را برای داده‌های مرگ و میر بیماران سکته مغزی (۱۹۹۳-۱۹۹۷) در زنان مناطق متروپلیس توکیو و حوزه اداری کاناگاوا در ژاپن به کار برده ایم.

شکل ۱: ۱۱۳ناحیه شامل بخش‌ها، شهرها و دهکده‌ها در توکیومتروپلیس و حوزه اداری کاناگاوا در ژاپن

مجموع اعداد مشاهده شده مرگ و میر بیماران سکته مغزی در زنان بالای دوره ۵ ساله، ۴۵۷۰۰ در منطقه می‌باشد. در مورد ماکزیمم طول برای نزدیکترین همسایه‌های K ما باید برای آماره‌های کاوشی فضایی انعطاف‌پذیر اصلی $K = 20$ و برای آماره‌های کاوشی فضایی دایره‌ای کالدورف و انعطاف‌پذیر مطرح شده $K = 50$ را انتخاب کنیم. مقدار $K = 50$ در این مثال مطابق با نیمی از ناحیه‌ها می‌باشد. ما آماره‌های کاوشی فضایی را با استفاده از ۹۹۹ تکرار برای آزمون فرضیه مونت‌کارلو محاسبه کردیم و نتایج در جدول ۱ خلاصه شده‌اند. آماره‌های کاوشی فضایی انعطاف‌پذیر یک ترکیب MLC از ۱۰ ناحیه $\{23, 22, 21, 18, 17, 16, 8, 7, 6, 1\}$ (اعداد ناحیه در شکل ۱ نشان داده شده است.) با $log\lambda_T (= log\lambda_K) = 151.7$، $\hat{\theta} = 1.18$ و $p = \frac{1}{(999+1)} = 0.001$ را آشکار می‌کند. اکثر این ۱۰ ناحیه درون MLC نشان داده شده، خطر مرتفع شدن را به طور معنی‌داری دارند. آماره کاوشی فضایی انعطاف‌پذیر اصلی در این مثال با MLC یکسان آشکارسازی شده است. هرچند، زمان اجرای آماره کاوشی فضایی انعطاف‌پذیر مطرح شده کمتر از یک ثانیه بوده است ولی این مقدار کمتر از زمان اجرای آماره کاوشی فضایی انعطاف‌پذیر اصلی یعنی ۳۳۸ ثانیه می‌باشد.

جدول ۱: محتمل‌ترین خوشه تشخیص داده شده در $\alpha_{\circ} = ۰٫۰۵$ به وسیله آماره‌های کاوشی فضایی انعطاف‌پذیر.

p_i	ردیف	شماره ناحیه	ریسک نسبی	مقادیر مورد انتظار	مقادیر مشاهده شده
$۱٫۵ * ۱۰^{-۱۱}$	۱	۲۳	۱٫۲۱	۱۰۷۲٫۳	۱۲۹۷
$۱ * ۱۰^{-۱۴}$	۲	۲۲	۱٫۲۵	۱۰۱۳٫۲	۱۲۶۶
$< ۱ * ۱۰^{-۱۷}$	۳	۱۸	۱٫۴۱	۵۲۲٫۷	۷۳۸
$۲٫۷ * ۱۰^{-۶}$	۴	۷	۱٫۱۹	۶۲۰٫۴	۷۳۷
$۲٫۷ * ۱۰^{-۵}$	۵	۸	۱٫۱۵	۷۸۰٫۶	۸۹۶
$۷٫۷ * ۱۰^{-۸}$	۶	۶	۱٫۲۳	۵۵۰٫۵	۶۷۸
۰٫۰۵۷	۷	۱	۱٫۱۳	۱۴۴٫۷	۱۶۴
۰٫۰۰۰۲۹	۸	۱۷	۱٫۱۱	۹۹۹٫۳	۱۱۱۰
$۹٫۲ * ۱۰^{-۸}$	۹	۲۱	۱٫۱۵	۱۳۳۵٫۱	۱۵۳۰
۰٫۰۲۴	۱۰	۱۶	۱٫۰۷	۷۴۳٫۴	۷۹۸
۰٫۷۷۸	۱۱	۵	۰٫۹۷	۵۶۶٫۳	۵۴۸
۰٫۱۶۱	۱۲	۲	۱٫۰۶	۲۵۱٫۳	۲۶۷

از طرفی، آماره کاوشی فضایی دایره ای کالدورف، دو ناحیه بیشتر $\{۲, ۵\}$ را به MLC آشکارسازی داده شده توسط آماره کاوشی فضایی انعطاف‌پذیر مطرح شده با $log\lambda_K = ۱۴۰٫۶$، $\hat{\theta} = ۱٫۱۷$ و $p = ۰٫۰۰۱$ اضافه کرده است.

بحث و نتیجه‌گیری

در این مقاله خلاصه‌ای از آماره‌های کاوشی فضایی بیان شد. هم‌چنین آماره کاوشی فضایی دایره‌ای کالدورف و آماره کاوشی فضایی انعطاف‌پذیر اصلی معرفی شد و به محدودیت‌ها و مشکلات این دو آماره اشاره گردید. سپس آماره کاوشی فضایی انعطاف‌پذیر با نسبت درستنمایی مقید روی داده‌های واقعی مورد ارزیابی قرار گرفت که علاوه بر سرعت در اجرای محاسبات از همسایگی‌های بیشتری می‌توان در تعیین خوشه‌های بیمار استفاده کرد. نتایج نشان داد که آماره کاوشی فضایی انعطاف‌پذیر همیشه خوشه‌های کوچکتر از خوشه‌های درست را تشخیص می‌دهد که بدیهی است به واسطه مقدار کوچک K ، بار محاسباتی سنگینی برای آماره کاوشی فضایی انعطاف‌پذیرفراهم می‌شود. میانگین زمان اجرا نیز برای آماره کاوشی فضایی انعطاف‌پذیر کمتر از یک ثانیه، برای آماره کاوشی فضایی انعطاف‌پذیر اصلی یک ثانیه و برای آماره کاوشی فضایی دایره‌ای حدود ۱۲۰ ثانیه می‌باشد. نشان داده می‌شود که زمان اجرا برای آماره کاوشی فضایی انعطاف‌پذیر به طور واقعی بدون محدودیت روی K به طور شگفت‌آوری کمتر از آماره کاوشی فضایی انعطاف‌پذیر اصلی با $K = ۲۰$ است.

مراجع

Assuncao R. Costa M. Tavares A. and Ferreira S . (2006), Fast Detection of Arbitrarily Shaped Disease Clusters *Statistics in Medicine,* 25, 723–742.

Duczmal L. and Assuncao R . (2004), A Simulated Annealing Strategy for the Detection of Arbitrarily Shaped Spatial Clusters *Computational Statistics & Data Analysis*,45, 269-286.

Huang L. Pickle LW. and Das B . (2008), Evaluating Spatial Methods for Investigating Global Clustering and Cluster Detection of Cancer Cases *Statistics in Medicine,* 27, 5111–5142 .

Kulldorff M . (1997), A Spatial Scan Statistic *Communications in Statistics: Theory and Methods.*, 26, 1481-1496.

Kulldorff M. and Nagarwalla N . (2004), Spatial Disease Clusters: Detection and Inference *Statistics in Medicine* , 14, 799–810.

Kulldorff M. Huang L. Pickle L. and Duczmal L . (2006), An Elliptic Spatial Scan Statistic *Statistics in Medicine,* 25, 3929–3943.

Lawson AB. Biggeri A. Böhning D. Lesaffre E. Viel JF. and Bertollini R (eds). (1999), *Disease Mapping and Risk Assessment for Public Health,* New York, John Wiley.

Patil GP. and Taillie C . (2004), Upper Level Set Scan Statistic for Detecting Arbitrarily Shaped Hotspots *Environmental and Ecological Statistics* ,11, 183-197.

Takahashi K. Yokoyama T. and Tango T. (2010), FleXScan v3.1: Software for the Flexible Scan Statistic *National Institute of Public Health: Japan,* .

Tango T . (2000), A Test for Spatial Disease Clustering Adjusted for Multiple Testing *Statistics in Medicine,* 19, 191–204 .

Tango T . (2008), A Spatial Scan Statistic with a Restricted Likelihood Ratio *Japanese Journal of Biometrics,* 29, 75–95.

Tango T . (2010), *Statistical Methods for Disease Clustering.*, New Springer.

Tango T. and Takahashi K . (2005), A Flexibly Shaped Spatial Scan Statistic for Detecting Clusters *International Journal of Health Geographics* ,411.

Waller LA. and Gotway CA . (2004), *Applied Spatial Statistics for Public Health Data*, New York, John Wiley.