

مجله علوم آماری، بهار و تابستان ۱۳۸۷

جلد ۲، شماره ۱، ص ۷۱-۵۱

برازش مدل رگرسیونی رشد به مجموعه‌های تصادفی بولی

مجتبی خزائی

گروه آمار، دانشگاه شهید بهشتی

تاریخ دریافت: ۱۳۸۷/۴/۲۳ تاریخ آخرین بازنگری: ۱۳۸۷/۶/۲۹

چکیده: یکی از مدل‌هایی که می‌تواند در مطالعه‌ی رابطه‌ی بین مجموعه‌های تصادفی بولی و متغیرهای کمکی به کار رود مدل رگرسیونی رشد است که با تعمیم مدل بولی و وابسته کردن توزیع دانه‌های آن به متغیرهای کمکی تعریف می‌شود. این مدل می‌تواند در مطالعه‌ی رفتار مجموعه‌های تصادفی بولی وقتی تغییر در ناحیه‌ی پوشیده شده توسط آنها با تغییر در اندازه‌ی دانه‌ها همراه است استفاده شود. در این مقاله شناسایی و برآزش مدل مناسب، با تکیه به اطلاعات موجود در مصادیق مدل‌های بولی و مقادیر متغیرهای کمکی را امکان پذیر ساخته‌ایم. همچنین روش مناسبی برای برآزش مدل رگرسیونی رشد ارائه و خواص برآوردهای حاصل از آن را به کمک شبیه‌سازی مورد مطالعه قرار داده‌ایم.

واژه‌های کلیدی: مجموعه‌ی تصادفی بسته، تابع اصابت، کسر حجمی، مدل بولی، مدل رگرسیونی رشد.

آدرس الکترونیک مسئول مقاله: مجتبی خزائی، m_khazaei@sbu.ac.ir

کد موضوع بنای ریاضی (۲۰۰۰): 62G99

۱ مقدمه

پیامد بعضی از پدیده‌های تصادفی، اجسام یا تصاویری هستند که می‌توان آنها را به عنوان زیر مجموعه‌هایی از R^d مطالعه کرد. به عنوان مثال شکل یک تومور، ناحیه‌ی حاصل از اصابت یک شهاب سنگ یا بمب ناحیه‌ی پوشیده شده توسط نوعی گیاه یا آتش در جنگل، ناحیه‌ای از مغز که در پاسخ به یک محرک عکس العمل نشان می‌دهد و غیره. در مطالعه رفتار تصادفی این پدیده‌ها ماترون (۱۹۷۵) و کندال (۱۹۷۴) به طور مستقل از یکدیگر نظریه‌ی مجموعه‌های بسته تصادفی را ارائه کردند. شرح مختصراً از نظریه‌ی آنها و تعریف یک مجموعه‌ی بسته تصادفی را در بخش ۲ این مقاله آورده‌ایم. طی سال‌های گذشته مدل‌های پارامتری مختلفی برای تبیین رفتار تصادفی این پدیده‌ها و تولید چنین مجموعه‌های تصادفی پیشنهاد شده است که ساده‌ترین و پرکاربردترین آنها چه در تئوری و چه در عمل مدل بولی است. با مدل بولی و بعضی از خواص آن در بخش ۲ آشنا می‌شویم. طی سال‌های گذشته عمده‌ی فعالیت آماری محققین در مطالعه‌ی مجموعه‌های تصادفی روی برآورد پارامترها و مطالعه‌ی ویژگی‌های یک مجموعه تصادفی به کمک تحقیقی از آن متتمرکز بوده است. به عنوان مثال می‌توان به کرسی (۱۹۹۳)، اشتولیان و همکاران (۱۹۹۵ و ۲۰۰۱) و مولچانوف (۱۹۹۷) اشاره کرد. اما خزائی (۲۰۰۴) و خزائی و شفیعی (۲۰۰۶) به منظور گسترش روش‌های آماری به مجموعه‌های بسته تصادفی بحث رگرسیون در مجموعه‌های تصادفی بولی را مطرح کردند. ایده‌ی آنها این بود که در بسیاری از تحقیقات علاوه بر تصویر ثبت شده اطلاعات کمکی دیگری نیز وجود دارند که می‌توانند بر توزیع مجموعه‌ی تصادفی تأثیر بگذارند. به عنوان مثال جنسیت، سن و دیگر عوامل زیستی می‌توانند بر شکل توموری که مشاهده می‌شود تأثیر بگذارند. بنابر این مطالعه رابطه‌ی بین مجموعه‌های تصادفی و متغیرهای کمکی به منظور استفاده از مقدار متغیرهای کمکی در پیش‌بینی رفتار مجموعه‌ی تصادفی وابسته به آنها ضروری می‌نماید. آنها با تعمیم مدل بولی مدل‌های رگرسیونی انتشار، رشد و انتشار-رشد را تعریف کرده (بخش ۳) و روش‌هایی برای برآش مدل‌های انتشار ارائه کردند.

در این مقاله مدل‌های رگرسیونی رشد را به منظور یافتن روشی برای برآشش این گونه مدل‌ها بررسی کردہ‌ایم. در این مدل‌ها افزایش یا کاهش ناحیه‌ی پوشیده شده توسط مجموعه‌ی تصادفی با تغییر در اندازه‌ی دانه‌ها^۱ همراه است و متوسط تعداد دانه‌ها در واحد سطح ثابت فرض می‌شود. همچنین فرض می‌کنیم دانه‌ها دیسکهایی تصادفی هستند، حالتی که بسیاری از پدیده‌های واقعی را به خوبی تقریب می‌کند. علی‌رغم ساختار ظاهرآ ساده‌ی این مدل با توجه به غیرقابل مشاهده بودن شعاع‌ها در حالت کلی برآش مدل با مشکل رویرو است در بخش ۴ برآورد کننده‌ی مناسبی برای این شعاع‌ها شناسایی و پیشنهاد کردہ‌ایم. با جایگزین کردن این برآوردها و روش کمترین توانهای دوم عادی پارامترهای مدل رشد را برآورد کرده و خواص آنها را به کمک شبیه‌سازی مطالعه کرده‌ایم.

۲ مجموعه‌های بسته تصادفی و مدل بولی

در نظریه ماترون (۱۹۷۵) و کنдал (۱۹۷۴)، یک مجموعه‌ی بسته تصادفی روی یک مجموعه‌ی پایه‌ی E تعریف می‌شود. در حالت کلی E یک فضای هاووسدورف تفکیک پذیر موضعآ فشرده در نظر گرفته می‌شود. یک مجموعه‌ی بسته تصادفی روی چنین فضایی به ترتیب زیر تعریف می‌شود. اگر \mathcal{F} گردایه‌ی زیر مجموعه‌های بسته‌ی E و K گردایه‌ی زیر مجموعه‌های فشرده‌ی E باشند برای هر $A \subset E$ تعریف می‌کنیم

$$\mathcal{F}_A = \{F \in \mathcal{F} : F \cap A \neq \emptyset\}, \quad \mathcal{F}^A = \{F \in \mathcal{F} : F \cap A = \emptyset\}.$$

می‌توان نشان داد که گردایه‌ی مجموعه‌هایی به شکل $\mathcal{F}^K \cap \mathcal{F}_{G_1} \cap \dots \cap \mathcal{F}_{G_n}$ که در آن G_1, \dots, G_n به ازای $n = 1, 2, \dots$ ، زیر مجموعه‌های بازی از E و $K \in \mathcal{K}$ هستند، پایه‌ای توپولوژیک روی \mathcal{F} است (ماترون ۱۹۷۵). توپولوژی تولید شده توسط این پایه به توپولوژی اصابت - یا - عدم اصابت^۲ معروف است. اگر Σ - میدان تولید شده توسط این توپولوژی و (Ω, \mathcal{A}, P) یک فضای احتمال دلخواه باشد،

^۱ Grains

^۲ Hit-or-Miss Topology

۵۴ برآش مدل رگرسیونی رشد به مجموعه‌های تصادفی بولی

یک مجموعه‌ی بسته تصادفی چون Y نگاشتی اندازه پذیر از (\mathcal{A}, Ω) به (\mathcal{F}, Σ) تعریف می‌شود. در این صورتتابع احتمال القاء شده توسط Y روی Σ عبارت است از

$$P_Y(B) = P(Y^{-1}(B)), \quad \forall B \in \Sigma.$$

بدیهی است این تابع احتمال آنقدر پیچیده است که فاقد جنبه‌های کاربردی به خصوص در مسایل آماری است. خوشبختانه بین این تابع احتمال و تابعی با ساختار ساده‌تر، یعنی تابع اصابت Y که به صورت زیر تعریف می‌شود تناظری یک به یک وجود دارد (ماترون ۱۹۷۵).

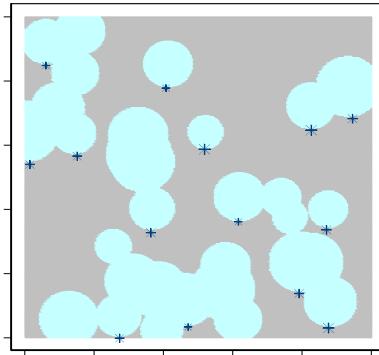
$$T_Y(K) = P_Y(Y \cap K \neq \emptyset) = P(\{Y \cap K \neq \emptyset\}), \quad K \in \mathcal{K}$$

در واقع تابع اصابت نقشی مشابه تابع توزیع تجمعی برای متغیرهای تصادفی دارد و در مطالعه‌ی توزیع مجموعه‌های بسته تصادفی مناسبتر از P_Y است. شرایط لازم و کافی برای اینکه یک تابع تعریف شده روی \mathcal{K} یک تابع اصابت باشد را می‌توان در ماترون (۱۹۷۵) یافت. در ادامه با بعضی اصطلاحات و نتایج در حالتی که آشنا می‌شویم $E = \mathbb{R}^d$.

در عمل معمولاً فقط بخشی از یک تحقیق مجموعه‌ی تصادفی Y دیده می‌شود، که به آن تحقق Y در پنجره‌ی W می‌گوییم (شکل ۱ را ببینید) و استنباطه‌های آماری به کمک آن انجام می‌شود. طبیعی است عدم تغییر رفتار مجموعه‌ی تصادفی تحت انتقال یا دوران می‌تواند انجام چنین استنباطه‌ای را ساده‌تر کند. یک مجموعه‌ی بسته تصادفی روی \mathbb{R}^d را مانا می‌گوییم هرگاه تابع اصابت آن تحت انتقال پایا باشد و آن را همسانگرد^۲ می‌نامیم هرگاه تابع اصابت آن تحت دوران پایا باشد. برای زیر مجموعه‌ی دلخواه B از \mathbb{R}^d با حجم (اندازه‌ی لبگ)^۳ ۱، کسر حجمی^۴، برابر اید ریاضی حجم پوشیده شده از B توسط Y تعریف شده و با $p(B)$ نشان داده می‌شود. در صورتی که Y مانا باشد می‌توان نشان داد کسر حجمی به شکل ناحیه B بستگی

^۲ Isotropic

^۳ Volume fraction



شکل ۱: ناحیه‌ی سفید رنگ تحقیقی از یک مجموعه‌ی تصادفی بسته‌ی مانا و همسانگرد در پنجره‌ای مربع شکل به طول ۱ را نشان می‌دهد. ستاره‌ها در این شکل نقاط مماس پایین را نشان می‌دهند.

ندارد و برابر است با

$$p = P_Y(o \in Y) = T_Y(\{o\}), \quad (1)$$

یعنی احتمال اصابت Y به مبدأ. یک برآورد کننده‌ی ناریب برای p وقتی تحقیقی از مجموعه‌ی تصادفی مانا Y در پنجره‌ی W مشاهده می‌شود عبارت است از

$$\hat{p} = \frac{\|Y \cap W\|}{\|W\|} \quad (2)$$

که در آن $\|\cdot\|$ بیانگر اندازه‌ی لیگ است. تساوی فوق در \mathbb{R}^2 برابر نسبت مساحت پوشیده شده از W توسط Y به مساحت W است. به عنوان مثال در شکل ۱ مساحت پوشیده شده توسط Y , 0.458 و با توجه به اینکه $1 = \|W\|$ داریم

$$\hat{p} = 0.458$$

تعريف ۱: اگر $D = \{d_1, d_2, \dots\}$ یک فرایند نقطه‌ای پواسون همگن باشد λ در \mathbb{R}^d و $Z_i; i = 1, 2, \dots$ مجموعه‌های بسته‌ی تصادفی مستقل و هم توزیع با Z و

مستقل از D باشد، مدل بولی عبارت است از

$$Y = \bigcup_{d_i \in D} (Z_i \oplus d_i) \quad (3)$$

که در آن $\{z + d_i : z \in Z_i\}$. $Z_i \oplus d_i = \{z + d_i : z \in Z_i\}$ مجموعه‌ی تصادفی بولی نامیده می‌شود. معمولاً نقاط فرایند پواسون یعنی d_i ها جرم‌ها و Z_i ها دانه‌های مدل بولی نامیده می‌شوند. دانه‌هایی که در عمل بیشترین کاربرد را دارند عبارتند از پاره خط‌هایی با طول و جهت تصادفی، دیسکهایی با شعاع تصادفی و خوش‌های متناهی از نقاط. شکل ۱ تحقیقی از یک مدل بولی در پنجره‌ی W را نشان می‌دهد که در آن $\lambda = 35$ و $Z_0 = C(0, R)$ دیسکی با شعاع تصادفی یکنواخت روی بازه‌ی $(0/05, 0/1)$ است.

می‌توان نشان داد (اشتویان و همکاران، ۱۹۹۵) کهتابع اصابت یک مدل بولی به صورت

$$T_Y(K) = 1 - \exp\{-\lambda E[\|Z_0 \oplus \check{K}\|]\} \quad (4)$$

است، که در آن $\{\cdot - k : k \in K\}$ اندازه‌ی لبگ مجموعه‌ی $\|Z_0 \oplus \check{K}\|$ می‌باشد.

از تساوی (۴) و با توجه به تعریف نتیجه می‌شود که یک مجموعه‌ی تصادفی بولی مانا است و در صورتی که Z_0 همسانگرد باشد همسانگرد نیز است. با استفاده از تساوی‌های (۱) و (۴) کسر حجمی یک مجموعه‌ی تصادفی بولی عبارت است از

$$p = 1 - \exp\{-\lambda E[\|Z_0\|]\}. \quad (5)$$

در صورتی که Z_i محدب و «بردار یکه‌ی رو به بالا باشد اولین نقطه‌ی تماس ابر صفحه‌ای با بردار نرمال» که به سمت بالا حرکت می‌کند با Z_i را نقطه‌ی مماس پایین Z_i می‌نامیم. در یک مجموعه‌ی تصادفی بولی بعضی از این نقاط توسط سایر دانه‌ها پوشیده می‌شوند اما بعضی از آنها آشکار می‌مانند. می‌توان نشان داد که نقاط مماس پایین آشکار تشکیل یک فرایند نقطه‌ای باشد

پایین آشکار در پنجره‌ی W باشد برآورده‌ی گشتاوری
 $\lambda(1-p) = \lambda \exp\{-\lambda E[|Z_0|]\}$ می‌دهند. در صورتی که n^+ تعداد نقاط مماس

$$\hat{\lambda} = \frac{n^+}{\|W\|(1-\hat{p})}, \quad (6)$$

برای برآورد λ به دست می‌آید (کرسی، ۱۹۹۳ و مولچانوف، ۱۹۹۵). در شکل ۱
 نقاط مماس پایین با ستاره مشخص شده‌اند و تعداد آنها $n^+ = 14$ است. با توجه با
 اینکه $\hat{\lambda} = 26$ از تساوی (6) داریم $\hat{p} = 0.458$

۳ مدل‌های رگرسیونی

فرض کنید' $X = (X_1, X_2, \dots, X_k)$ بردار متغیرهای کمکی باشد که می‌خواهیم
 رابطه‌ی آنها را با مجموعه‌ی تصادفی بولی Y مطالعه کنیم. با توجه به اینکه رفتار
 یک مجموعه‌ی تصادفی بولی تحت تأثیر دو منبع تصادفی، یکی فرآیند پواسون D و
 دیگری توزیع Z است. خزائی و شفیعی (۲۰۰۶) متغیرهای کمکی را به سه دسته‌ی
 زیر تقسیم کردند.

- ۱- متغیرهای کمکی که فقط روی D تأثیر دارند را متغیرهای انتشار می‌نامیم.
- ۲- متغیرهای کمکی که فقط روی Z تأثیر دارند را متغیرهای رشد می‌نامیم.
- ۳- متغیرهای کمکی که هم روی D و هم روی Z تأثیر دارند را متغیرهای
 انتشار-رشد می‌نامیم.

تعمیمی طبیعی از مدل بولی (۳) برای نشان دادن وابستگی Y به X عبارت است
 از

$$Y_X = \bigcup_{d_j \in D_X} (Z_{jX} \oplus d_j), \quad (7)$$

که به شرط Z_{jX} ، X ها مجموعه‌های تصادفی مستقل و هم توزیع با Z_{0X} و
 D_X فرایند نقطه‌ای پواسون باشد λ_X ، مستقل از Z_{0X} است. متناظر با دسته‌بندی بالا
 مدل‌های رگرسیونی را نیز می‌توان به سه دسته‌ی مدل رگرسیونی انتشار، رشد و

انتشار-رشد دسته‌بندی کرد. مدل‌های رگرسیونی انتشار در مواردی که تغییرات ناحیه‌ی پوشیده شده توسط Y به تعداد جرم‌ها بستگی دارد نه توزیع دانه‌ها مناسب است. خزائی و شفیعی (۲۰۰۶) روش‌هایی برای برآش مدل‌های انتشار وقتی دانه‌ها دیسکهایی با شعاع معلوم، نامعلوم و تصادفی هستند ارائه کردند.

اگر X بردار متغیرهای کمکی باشد، مدل رگرسیونی رشد در مجموعه‌های تصادفی بولی به صورت

$$Y_X = \bigcup_{d_j \in D} (Z_{jX} \oplus d_j), \quad (8)$$

تعریف می‌شود، که در آن D فرایند نقطه‌ای پواسون باشد و Z_{jX} مجموعه‌های تصادفی مستقل و هم توزیع با مجموعه‌ی تصادفی $Z_{\circ X}$ و مستقل از D هستند. برای تکمیل مدل کافی است رابطه‌ای برای تبیین وابستگی $Z_{\circ X}$ به X ارائه کرد. با توجه به اینکه در غالب مواردی که مدل‌های بولی برای تقریب مشاهدات واقعی به کار می‌روند $Z_{\circ X}$ مجموعه‌های هندسی ساده‌ای مثل نقاط داخل یک دایره، مریع یا لوزی و... فرض می‌شوند، کافی است رابطه‌ی چنین مجموعه‌هایی با X شناسایی و مدل شود. به عنوان مثال همانطور که در این مقاله فرض کرده‌ایم $Z_{\circ X}$ می‌توانند دیسکهایی تصادفی با شعاع $\epsilon = f(X, \beta) + \epsilon$ بخصوص به ازای $f(X, \beta) = X'\beta$ باشند. انتظار داریم روش‌ها و نتایج این مقاله با اندکی تعدیل برای سایر موارد نیز کارآمد باشند. بنابراین در ادامه، مدل رگرسیونی رشد با دانه‌های دیسکی شکل، به صورت

$$\begin{cases} Y_X = \bigcup_{d_j \in D} (Z_{jX} \oplus d_j) \\ Z_{\circ X} = C(\circ, R_X), \quad R_X = X'\beta + \epsilon \end{cases} \quad (9)$$

را در نظر می‌گیریم، که در آن $C(\circ, R_X)$ دیسکی به مرکز مبدا با شعاع R_X است. بدیهی است که این مدل در مواردی که تغییرات ناحیه‌ی پوشیده شده توسط Y به اندازه‌ی دانه‌ها وابسته است نه به متوسط تعداد جرم‌ها می‌تواند بکار رود.

در صورتی که X و Y_X در n پسجهره مشاهده شوند (شکل‌های ۲ و ۴) برای برآورد β به کمک مشاهدات (X_i, Y_i) که $Y_i = Y_{X_i}$ ، کافی است مدل رگرسیونی $(X_i, R_i) = (X_i, R_{X_i}) = X'_i \beta + \epsilon_i, i = 1, 2, \dots, n$

برازش دهیم. اما مشکلی که وجود دارد قابل مشاهده نبودن R_i ها در حالت کلی است. پیشنهادی که داریم این است که مقادیر مناسبی را جایگزین R_i ها در این مدل رگرسیونی خطی کرده و برآوردهای کمترین توانهای دوم عادی β را به دست آوریم. در بخش بعد با برآوردکننده‌ی پیشنهادی برای R_i و نتایج حاصل از آن آشنا می‌شویم.

۴ روشهای برآورد مدل رگرسیونی رشد

مولچانوف (۱۹۹۷) پارامترهای یک مدل بولی را به دو دسته‌ی میکروسکوپی و ماکروسکوپی تقسیم کرد. به این ترتیب که پارامترهای ماکروسکوپی را می‌توان مستقیماً از روی تحقق مشاهده شده برآورد کرد. مثل کسر حجمی، مقادیر تابع اصابت و کوواریانس. در مقابل پارامترهای میکروسکوپی مستقیماً قابل برآورد نیستند. و غالباً به کمک روابطی که با سایر پارامترهای میکروسکوپی یا ماکروسکوپی دارند و روش گشتاوری (در مجموعه‌های تصادفی روش شدت^۵) برآورد می‌شوند. از این پارامترها می‌توان به λ اشاره کرد. با توجه به این تقسیم بندهی R_i را می‌توان در دسته‌ی پارامترهای میکروسکوپی قرار داد. برای برآورد آن از روش شدت به واسطه‌ی کسر حجمی مدل بولی استفاده می‌کنیم. از تساوی (۵) وقتی $Z_0 = C(\circ, R)$ نتیجه می‌شود

$$p = 1 - \exp\{-\pi\lambda E(R^2)\}$$

و

$$E(R^2) = -\frac{\ln(1-p)}{\pi\lambda}. \quad (10)$$

وقتی تحقیقی از Y در پنجره‌ی W مشاهده شود پس از برآورد p و λ به کمک روابط (۲) و (۶) و جایگزینی در تساوی (۱۰) برآوردکننده‌ی زیر حاصل می‌شود.

$$r = \hat{R} = \sqrt{-\frac{\ln(1-\hat{p})}{\pi\hat{\lambda}}}. \quad (11)$$

^۵ Intensity

بحث مختصری در رفتار این برآوردهای می‌تواند در انتخاب نهایی جایگزین R_i ها کمک نماید. ابتدا برای حالتی که λ معلوم است داریم

$$r = \frac{1}{\sqrt{\pi\lambda}} \sqrt{-\ln(1 - \hat{p})}. \quad (12)$$

که با بسط این تابع از \hat{p} حول p تقریب

$$r \approx \sqrt{E(R^2)} + C_1(p)(\hat{p} - p) + C_2(p)(\hat{p} - p)^2, \quad (13)$$

حاصل می‌شود، که در آن

$$\begin{aligned} C_1(p) &= \frac{1}{2\sqrt{\pi\lambda}}(1-p)^{-1}[-\ln(1-p)]^{-1/2} \\ C_2(p) &= \frac{-1}{4\sqrt{\pi\lambda}}(1-p)^{-2}[-\ln(1-p)]^{-3/2}[1 - 2\ln(1-p)]. \end{aligned}$$

با استفاده از جملات اول و دوم در تقریب فوق داریم $E(r) \approx \sqrt{E(R^2)} \geq E(R)$. اما با اضافه شدن جمله‌ی سوم و با توجه به اینکه برای $1 < p < 0$ داریم $C_2(p) < 0$ ، این اریبی مثبت تقلیل می‌یابد و همانطور که شبیه‌سازی‌ها نیز تأیید می‌کنند (خزایی، ۱۳۸۶)، $E(r) \approx E(R)$

با استفاده از جملات اول و دوم در تقریب (13) داریم

$$\begin{aligned} Var(r) &\approx C_1(p)^2 Var(\hat{p}) \\ &= \left(\frac{1}{2\sqrt{\pi\lambda}(1-p)\sqrt{-\ln(1-p)}} \right)^2 Var(\hat{p}) \\ &= \frac{Var(\hat{p})}{4\pi^2[\lambda(1-p)]^2 E(R^2)}, \end{aligned} \quad (14)$$

که در آن $Var(\hat{p}) = \frac{1}{\|W\|^4} \int_W \int_W (C(x-y) - p^2) dx dy$ و $C(x, y) = \int_{\mathcal{P}} P(x \in Y, y \in Y)$ کواریانس مجموعه‌ی تصادفی Y و برابر است با $P(x \in Y, y \in Y)$. برای بحث در واریانس \hat{p} و برآورده آن می‌توان به میس (۱۹۸۲) مراجعه کرد.

بدیهی است در حالتی که λ معلوم نیست حصول نتایجی مشابه به طور جبری بسیار پیچیده‌تر خواهد بود. در این حالت از شبیه‌سازی کمک گرفته‌ایم. با توجه به اینکه در برآزش مدل رشد به برآورد R در n پنجمره متناظر با مقادیر متغیرهای

کمکی نیاز داریم بعلاوه فرض کرده‌ایم λ به متغیرهای کمکی وابسته نیست و مقدارش ثابت است، به همراه برآورده کننده (۱۱) یعنی $r_i = \hat{R}_i = \sqrt{-\frac{\ln(1-\hat{p}_i)}{\pi\hat{\lambda}_i}}$ برآورده کننده زیر را نیز برای برآورد R در پنجره‌ی λ پیشنهاد می‌کنیم.

$$r_i^* = \hat{R}_i = \sqrt{-\frac{\ln(1-\hat{p}_i)}{\pi\bar{\lambda}}}, \quad (15)$$

که در آن $\hat{\lambda}_i = \bar{\lambda} \sum_{i=1}^n$ برآورده‌ای برای λ مشترک است.

قبل از مقایسه‌ی این دو برآورده کننده به کمک شبیه‌سازی به یک مزیت r_i^* اشاره می‌کنیم. در صورتی که برای یک i و در نتیجه $\hat{\lambda}$ صفر باشد r_i تعریف نشده است. در حالی که بسیار بعید است که برای هر i $r_i^* = n_i^+$ و در نتیجه $\bar{\lambda}$ صفر باشد و r_i^* تعریف نشده باشد.

شبیه‌سازی‌های انجام شده توسط خزائی (۱۳۸۶) نشان می‌دهد که امید ریاضی هر دوی r_i و r_i^* با $E(R_i)$ یکی است. توزیع r_i^* فارغ از توزیع R_i ، نرمال است در حالی که توزیع r_i حتی در حالتی که R_i توزیع نرمال دارد بسیار نامنظم است. در اینجا به اختصار و برای یک مدل بعضی از این نتایج توصیف می‌شوند.

شبیه‌سازی و نتایج آن: مدل رگرسیونی رشد (۹) را وقتی D فرایند نقطه‌ای پواسون باشد $\lambda = 7 + 0.07x_i + \epsilon_i$ و $R_i = 0.0004$ است به ازای x_i اینجا $i = 1, 2, \dots, 12$ و مقادیر متغیر کمکی x که در جدول ۳ آمده، در نظر بگیرید. در هر یک از سه حالتی که $\epsilon_i \sim N(0, 0.0004)$ ، $\epsilon_i \sim U(-0.03464, 0.03464)$ یعنی توزیع یکنواخت با میانگین صفر و واریانس 0.0004 و $\epsilon_i \sim DEXP(0, 0.0004)$ یعنی توزیع نمایی دوگانه با میانگین صفر و واریانس 0.0004 ، از این مدل 1000 بار شبیه‌سازی کرده و آماره‌های مربوطه را حساب کرده‌ایم. واریانس خطرا در سه حالت برابر با 0.0004 اختیار کرده‌ایم. دلیل این انتخاب این است که، نهایتاً قصد داریم تحقیق‌هایی از مدل‌های بولی با دانه‌های دیسکی شکل و شعاع $\lambda = 7 + 0.07x_i + \epsilon_i$ در پنجره‌های 1×1 ، شبیه‌سازی کنیم. میانگین شعاع‌های بزرگ (در این مثال وقتی x بزرگ است) به پنجره‌هایی تقریباً یا کاملاً پوشیده توسط Y منجر می‌شوند. در چنین مواردی برآورده کننده‌های معمول \hat{p} و $\bar{\lambda}$ علاوه بر اینکه دقیق نیستند مشکلات محاسباتی بعدی را نیز به همراه

دارند. از طرفی در پنجره‌هایی با میانگین شعاع کوچک (در این مثال وقتی x کوچک است) توزیهای فوق می‌توانند به مقدار منفی منجر شوند. با آزمون و خطا، $\sigma^2 = 0,0004$ را طوری تعیین کرده‌ایم که علاوه بر حفظ ماهیت تصادفی شعاع‌ها، شانس مشاهده‌ی پنجره‌های تقریباً یا کاملاً پوشیده و مقادیر منفی برای شعاع حداقل گردد. شبیه‌سازی‌ها طی مراحل زیر انجام شده است.

۱: n مشاهده از توزیع خطای شبیه‌سازی کرده و $R_i = 0,07 + 0,15x_i + \epsilon_i$ را حساب می‌کنیم.

۲: $n = 12$ مشاهده m_i از توزیع پواسون با $\lambda = 7$ شبیه‌سازی می‌کنیم (تعداد جرمها در پنجره‌ی i را تعیین می‌کند).

۳: تحقیق‌هایی از مدل بولی با پارامترهای $12, \dots, 1, 2, \dots, m_i, R_i$ (شبیه‌سازی می‌کنیم. این کار را با تکرار مراحل زیر انجام می‌دهیم (اشتویان و همکاران، ۱۹۹۵).

۱-۳: $2 \times m_i$ مشاهده از توزیع یکنواخت روی (۰, ۱) شبیه‌سازی می‌کنیم. هر زوج از این مشاهدات مختصات یک نقطه از فرایند نقطه‌ای پواسون نظری مدل بولی را مشخص می‌کند.

۲-۳: m_i دایره به شعاع R_i به مرکز نقاط فوق رسم می‌کنیم.

۴: برای تحقیقات حاصل در مرحله‌ی ۳ که اشکالی مشابه شکل ۲ هستند آماره‌های $\hat{q}_i, r_i^*, \hat{\lambda}_i, n_i^+$ و غیره را محاسبه می‌کنیم.

مراحل فوق را ۱۰۰۰ بار تکرار می‌کنیم.

برای اختصار در توزیع r_i^* ‌ها به ازای $i = 4, 8, 12$ بحث خواهیم کرد. نتایج برای سایر i ‌ها کاملاً مشابه است. برای این i ‌ها داریم $E(R_4) = 0,130$, $E(R_8) = 0,175$ و $E(R_{12}) = 0,220$. میانگین r_i^* ‌های حاصل از شبیه‌سازی که در جدول ۱ آمده است بر ناریبی (تقریبی) r_i^* ‌ها تاکید دارند. برای آزمون فرض نرمال بودن از روش‌های آزمون کولموگروف – اسمیرنوف و شاپیرو – ویلک استفاده

کرده‌ایم. با توجه به تعداد نمونه‌ی بزرگ $n = 1000$ برای انجام آزمون شاپیرو-ویلک از تصحیح رویستون (۱۹۸۲) استفاده کرده‌ایم. نتایج در جدول ۲ همگی فرض نرمال بودن توزیع r_i^* ‌ها را تأیید می‌کنند.

جدول ۱: میانگین r_i^* ‌های حاصل از شبیه‌سازی به ازای $i = 4, 8, 12$

توزیع خطاهای $i = 12, 8, 4$		
$i = 12$	$i = 8$	$i = 4$
۰/۲۱۹	۰/۱۷۱	۰/۱۲۹
۰/۲۱۶	۰/۱۷۴	۰/۱۳۰
۰/۲۱۵	۰/۱۷۳	۰/۱۳۰

نرمال
یکنواخت
نمائی دوگانه

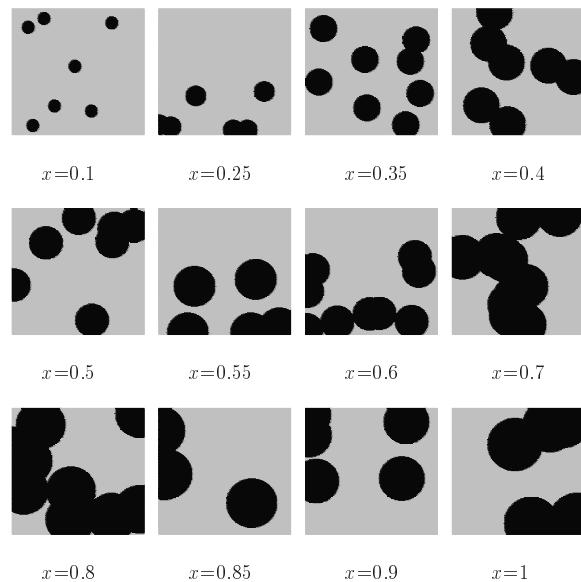
جدول ۲: مقادیر آماره‌های آزمون کولموگروف - اسمیرنف و شاپیرو- ویلک به همراه اندازه‌ی - احتمال این آزمون‌ها در آزمون فرض نرمال بودن توزیع‌های r_i^* به ازای $i = 4, 8, 12$

$i = 12$		$i = 8$		$i = 4$		آزمون	نرمال	یکنواخت	نمائی دوگانه	توزیع خطاهای $i = 4, 8, 12$
P.V	آماره	P.V	آماره	P.V	آماره					
۰/۶۰۱	۰/۰۱۷۷	۰/۱۲۲	۰/۰۲۵۵	۰/۱۱۶	۰/۰۲۵۳	K-S				
۰/۸۸۲	۰/۹۸۹۳	۰/۱۳۳	۰/۹۸۳۲	۰/۲۷۱	۰/۹۸۵۳	S-W				
۰/۳۶۸	۰/۰۲۰۶	۰/۳۷۴	۰/۰۲۰۵	۰/۰۹۸	۰/۰۲۶۳	K-S				
۰/۱۱۰	۰/۹۸۳۹	۰/۷۰۳	۰/۹۸۷۵	۰/۴۲۱	۰/۹۸۶۲	S-W				
۰/۶۹۸	۰/۰۱۸۷	۰/۲۶۱	۰/۰۲۱۶	۰/۱۲۳	۰/۰۲۸۰	K-S				
۰/۶۵۲	۰/۹۸۷۱	۰/۷۰۳	۰/۹۸۷۵	۰/۱۲۳	۰/۹۸۴۱	S-W				

با توجه به خواص r_i^* ‌ها، به نظر انتخاب مناسبی برای جایگزینی به جای R_i می‌باشدند. اگر r^* بردار شامل این مقادیر و \mathbf{X} ماتریس مشاهدات متغیرهای کمکی باشند برآوردکننده‌ی کمترین توانهای دوم عادی برای β و σ^2 عبارتند از

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{r}^*, \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (r_i^* - \hat{r}_i^*)^2}{n-p}. \quad (16)$$

با توجه به نرمال بودن توزیع r_i^* ‌ها انتظار داریم توزیع $\hat{\beta}$ نرمال باشد شبیه‌سازی‌ها این را تأیید می‌کند. اما قبل از پرداختن به آن، در مثال زیر با چگونگی استفاده از روش فوق در برآشش یک مدل رشد به مشاهدات موجود بیشتر آشنا می‌شویم.



شکل ۲: تحقیق‌هایی از مدل رشد وقتی که $\lambda = ۷$, $R_x = ۰/۰۷ + ۰/۱۵x + \epsilon$ به همراه مقادیر متغیر کمکی x در پنجره‌هایی ۱×۱ و $\epsilon \sim N(۰, ۰/۰۰۰۴)$

مثال ۱: شکل ۲ یکی از تحقیق‌های شبیه‌سازی شده از مدل مورد بحث را به همراه مقادیر متغیر کمکی x نشان می‌دهد. به کمک روابط ارائه شده مقادیر \hat{p}_i , n_i^+ , $\hat{\lambda}_i$ و r_i^* را در هر پنجره محاسبه و در جدول ۳ آورده‌ایم. به کمک آنها در این جدول $\bar{\lambda} = \frac{۱}{۷} \sum \hat{\lambda}_i = ۵/۶$ و به کمک مشاهدات ۱۲ , (x_i, r_i^*) ; $i = ۱, ۲, \dots, ۱۲$, و روابط $\hat{\beta}_0 = (۰/۰۷۶, ۰/۱۲)'$ و $\hat{\beta}_1 = (\hat{\beta}_0, \hat{\beta}_1)' = (۰/۰۷۶, ۰/۱۲)'$ به دست می‌آیند. در نتیجه مدل رشدی که به مشاهدات شکل ۲ برآش می‌شود عبارت است از:

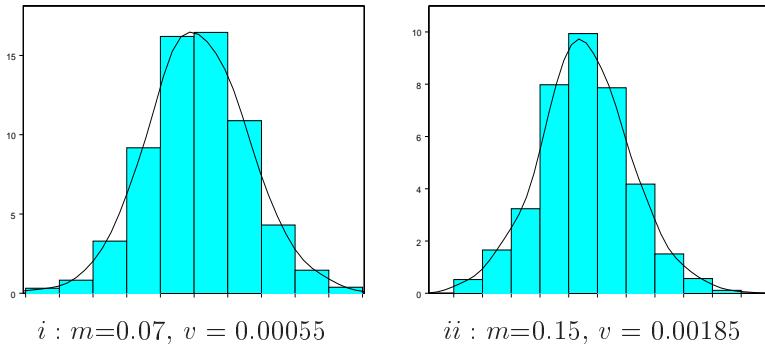
$$\begin{cases} Y_x = \bigcup_{d_j \in \hat{D}} (\hat{Z}_{jx} \oplus d_j) \\ \hat{Z}_{\circ x} = C(\circ, \hat{R}_x), \quad \hat{R}_x = ۰/۰۷۶ + ۰/۱۲x, \quad \hat{\sigma}^2 = ۰/۰۰۱۲ \end{cases} \quad (17)$$

که در آن \hat{D} فرایند نقطه‌ای پواسون همگنی باشد.

برای مطالعه توزیع $\hat{\beta}_0$ و $\hat{\beta}_1$, رفتار $\hat{\beta}_0$ و $\hat{\beta}_1$ ‌های حاصل از تحقیق‌های شبیه‌سازی شده در فوق را بررسی کرده‌ایم. نمودارها در شکل ۳ هیستوگرام و برآورد تابع چگالی به همراه میانگین و واریانس نمونه‌ای $\hat{\beta}_0$ و $\hat{\beta}_1$ ‌های حاصل را نشان می‌دهند (نمودارها در حالتی که شعاع‌های اولیه دارای توزیع نرمال‌اند

جدول ۳: مقادیر متغیر کمکی و آمارهای محاسبه شده در مثال ۱

i	۱۲	۱۱	۱۰	۹	۸	۷	۶	۵	۴	۳	۲	۱	x
۱	۰/۹	۰/۸۵	۰/۸	۰/۷	۰/۶	۰/۵۵	۰/۵	۰/۴	۰/۳۵	۰/۲۵	۰/۱	\hat{p}	x
۰/۴۳	۰/۳۳	۰/۲۶	۰/۶۱	۰/۰۱	۰/۳۴	۰/۲۹	۰/۱۹	۰/۳۷	۰/۲۸	۰/۱۰	۰/۰۶	n^+	
۲	۴	۲	۳	۳	۴	۲	۵	۴	۶	۳	۵	n^-	
۳/۰۱	۰/۹۴	۲/۷۲	۷/۶۶	۶/۱۲	۶/۰۹	۲/۸۲	۷/۰۷	۶/۳۵	۸/۳۲	۳/۳۲	۷/۴۳	\bar{x}	
۰/۱۸	۰/۱۵	۰/۱۳	۰/۲۳	۰/۲	۰/۱۵	۰/۱۴	۰/۱۴	۰/۱۶	۰/۱۴	۰/۰۸	۰/۰۶	r^*	


 شکل ۳: توزیع $\hat{\beta}_0$ و توزیع $\hat{\beta}_1$

رسم شده‌اند در دو حالت دیگر نمودارها مشابه‌اند). همانطور که مشاهده می‌شود برای $\hat{\beta}_0$ داریم $\hat{\beta}_0 = \beta_0 \approx 0.072$ و برای $\hat{\beta}_1$ داریم $\hat{\beta}_1 = \beta_1 \approx 0.145$ و این برآوردهای ناواریب هستند. مقادیر آماره‌های آزمون‌های کولموگروف-اسمیرونف و شاپیرو-ویلک و اندازه‌های احتمال بزرگتر از ۵٪ در جدول ۴ نشان می‌دهد که فرض نرمال بودن توزیع $\hat{\beta}_0$ و $\hat{\beta}_1$ را نمی‌توان رد کرد.

همانطور که در مثال ۱ ملاحظه می‌شود واریانس برآورد شده در مقایسه با مقدار واقعی بزرگ است. در واقع متوسط $\hat{\sigma}^2$ که در ۱۰۰۰ بار شبیه‌سازی به دست آمد برابر است با 0.0014 که در مقایسه با مقدار واقعی 0.0004 بزرگ است. کمی دقت نشان می‌دهد آنچه که برآورده شده واریانس ϵ نیست. زیرا مدل $r^* = R + \epsilon^* = \beta_0 + \beta_1 x + \epsilon + \epsilon^* = \beta_0 + \beta_1 x + \epsilon$ در حالی که مدل واقعی $R = \beta_0 + \beta_1 x + \epsilon$ است. در واقع در این روش واریانس ϵ برآورده شده نه واریانس ϵ و واریانس ϵ قابل برآورده نیست. البته این توانی

جدول ۴: مقادیر آماره‌های آزمون کولمتوگروف-اسمیرنف و شاپیرو-ویلک به همراه اندازه‌ی احتمال این آزمون‌ها در آزمون فرض نرمال بودن توزیع‌های $\hat{\beta}_1$ و

$\hat{\beta}_1$	P.V	آماره	P.V	آماره	آزمون	توزیع خطاهای
۰/۰۹۶	۰/۰۲۴۷	۰/۶۷۹	۰/۰۱۶۹		K-S	نرمال
۰/۲۹۳	۰/۹۸۵۴	۰/۷۴۳	۰/۹۸۸۲		S-W	
۰/۸۲۶	۰/۰۱۵۸	۰/۲۳۶	۰/۰۲۳۴		K-S	یکنواخت
۰/۴۶۷	۰/۹۸۵۹	۰/۵۳۰	۰/۹۸۶۳		S-W	
۰/۱۳۳	۰/۰۲۸۱	۰/۶۶۵	۰/۰۱۹۱		K-S	نمائی دوگانه
۰/۹۷۰	۰/۹۹۰۱	۰/۷۲۹	۰/۹۸۷۹		S-W	

در برابر مشاهده پذیر نبودن R ‌ها است و در هر روش برآورد دیگری نیز اتفاق خواهد افتاد. از طرفی اگر قرار دهیم $\hat{V}(\hat{\beta}) = \hat{V}(\hat{\beta}^*)$ برآورد واریانس $\hat{\beta}$ حاصل از روابط معمول یعنی $(\hat{X}'\hat{X})^{-1} = \hat{\sigma}^2 \hat{V}(\hat{\beta})$ با آنچه در شکل ۳ آمده سازگار است. برای $\hat{\beta}_1$ داریم $\sum \hat{Var}(\hat{\beta}_{1,i}) = ۰/۰۰۰۵۸ \approx \frac{۱}{۱۰۰} \sum \hat{Var}(\hat{\beta}_{0,i}) = ۰/۰۰۰۵۵ \approx ۰/۰۰۰۵۵$ و برای $v = ۰/۰۰۱۸۵ \approx \frac{۱}{۱۰۰} \sum \hat{Var}(\hat{\beta}_{1,i}) = ۰/۰۰۱۷۶$ داریم. بنابراین می‌توان نتیجه گرفت

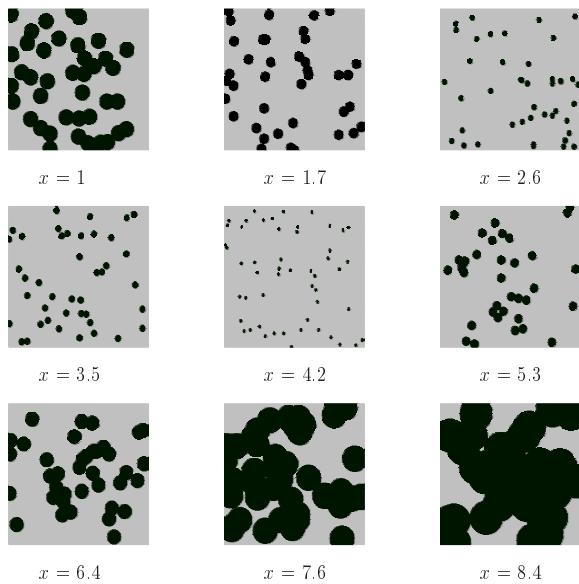
$$\hat{\beta} \approx N_p(\beta, \hat{V}(\hat{\beta})),$$

که در آن $\hat{V}(\hat{\beta}) = \frac{1}{n-p} \sum_{i=1}^n (r_i^* - \hat{r}_i^*)^2$ همچنین برای x_0 مفروض

$$\hat{R}x_0 \approx N(x_0' \beta, x_0' \hat{V}(\hat{\beta}) x_0), \quad (18)$$

بنابراین استنباط‌های آماری، با روابطی مشابه آنچه در رگرسیون خطی تحت فرض نرمال بودن خطا انجام می‌شود در اینجا هم قابل اجرا است. برآوردهای نقطه‌ای ناریباند ولی فواصل اطمینان و آزمون‌های فرض در اینجا نسبت به سطوح اطمینان و معنی داری معین محافظه کارانه‌تر هستند.

در مثال زیر علاوه بر مروری بر چگونگی شناسایی و برآش یک مدل رشد مناسب به مشاهدات از نتایج فوق در انجام استنباط‌های آماری نیز استفاده کردایم.



شکل ۴: تحقیق‌هایی از مدل‌های بولی در پنجره‌هایی 1×1 به همراه مقادیر متغیر کمکی x .

مثال ۲: می‌خواهیم به مشاهدات شکل ۴ یک مدل رشد مناسب برآذش دهیم. برای این منظور مقادیر آماره‌های مورد نیاز را محاسبه و در جدول ۵ آورده‌ایم. نمودار پراکنش در شکل ۵ رابطه‌ای درجه‌ی دوم بین r^* و x را پیشنهاد می‌کند. برآورد کمترین توانهای دوم عادی پارامترهای این مدل عبارتند از شدت فرایند نقطه‌ای پواسون مشترک نیز برابر است با $\bar{\lambda} = \frac{1}{q} \sum \hat{\lambda}_i = \frac{39}{5}$ و در نتیجه مدل رشد برآذش شده به این مشاهدات عبارت است از

$$\begin{cases} Y_x = \bigcup_{d_j \in \hat{D}} (\hat{Z}_{jx} \oplus d_j) \\ \hat{Z}_{\circ x} = C(\circ, \hat{R}_x), \quad \hat{R}_x = 0.084 - 0.033x + 0.0044x^2, \quad \hat{\sigma}^* = 0.0064 \end{cases}$$

که در آن \hat{D} فرایند نقطه‌ای پواسون همگنی با شدت $39/5$ است. جدول ۶ شامل این برآوردها، خطاهای معیار آنها همراه آماره‌های آزمون و مقدار-احتمال برای آزمون فرض‌های $H_{\circ i} : \beta_i = 0$ است. همانطور که ملاحظه

جدول ۵: مقادیر متغیر کمکی و آماره‌های محاسبه شده در مثال ۲.

i	۹	۸	۷	۶	۵	۴	۳	۲	۱	x
	۸/۴	۷/۶	۶/۴	۵/۳	۴/۲	۳/۵	۲/۶	۱/۷	۱	\hat{p}
۰/۷۵	۰/۶۵	۰/۲۷	۰/۱۱	۰/۰۳	۰/۰۸	۰/۰۶	۰/۱۵	۰/۳۶	۰/۹۸	n^+
۵	۱۶	۲۵	۲۹	۵۰	۴۲	۳۹	۳۱	۳۱	۳۱	$\hat{\lambda}$
۲۰/۰۵	۴۵/۲۴	۳۴/۴۵	۳۲/۶۹	۵۱/۴۰	۴۵/۶۱	۴۱/۳۷	۳۶/۳۴	۴۸/۲۴	۰/۰۹	r^*
۰/۱۱	۰/۰۹	۰/۰۵	۰/۰۳	۰/۰۱	۰/۰۳	۰/۰۲	۰/۰۴	۰/۰۶	۰/۰۶	

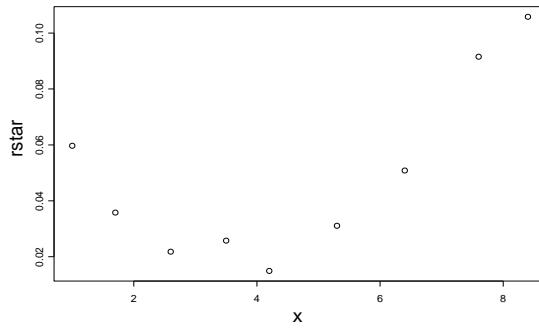
می‌شود فرض صفر بودن این ضرایب رد می‌شود. از طرف دیگر مقدار ضریب همبستگی چندگانه‌ی $R = ۰/۹۸$ نیز بر مناسب بودن این مدل تأکید می‌کند. اگر بخواهیم پیش‌بینی راجع به مدل بولی که می‌تواند مولد تحقیق‌هایی به ازای $x = ۷$ باشد انجام دهیم از مدل برآش شده به ازای $x = ۷$ انتظار داریم یک مدل بولی باشد $\hat{R}_7 = ۳۹/۵ = \hat{\lambda}$ و دانه‌هایی دایره‌ای شکل با شعاع متوسط $\hat{R}_7 = ۰/۰۶۹$ برای این منظور مناسب باشد. اگر بخواهیم برای شعاع متوسط، برآوردهای فاصله‌ای داشته باشیم می‌توانیم از رابطه‌ی (۱۸) کمک بگیریم. در این صورت یک برآورد فاصله‌ای با ضریب اطمینان $(1 - \alpha)$ درصد برای شعاع متوسط عبارت است از $\hat{R}_x \pm Z_{\alpha/2} \sqrt{x_0' \hat{V}(\hat{\beta}) x_0}$. به ازای $x_0 = (1, 7, 49)$ با کمی محاسبات جبری $\hat{R}_x = ۱۰۰$ درصد برآوردهای $\hat{V}(\hat{\beta}) x_0 = ۹/۲ \times ۱۰^{-6}$ و با ضریب اطمینان ۹۵ درصد برآوردهای $\hat{V}(\hat{\beta}) x_0 = (۰/۰۶۳, ۰/۰۷۵)$ حاصل می‌شود. اگر (Z_0, Ux_0) و (Z_u, Ux_u) به ترتیب متوسط محیط و مساحت دانه‌ها را نشان دهند وقتی $x_0 = (1, 7, 49)$ و $x_u = (0, 0, 0)$ برآوردهای نقطه‌ای برای آنها به ترتیب عبارت‌اند از $۰/۴۳۵ = ۰/۰۶۹ \pi$ و $۰/۰۱۵ = ۰/۰۶۹ \pi$ و به کمک برآوردهای فاصله‌ای فوق فوائل اطمینان ۹۵ درصدی زیر را برای (Z_0, Ux_0) و (Z_u, Ux_u) به دست می‌آوریم.

$$(2\pi r_l, 2\pi r_u) = (2\pi \times ۰/۰۶۳, 2\pi \times ۰/۰۷۵) = (۰/۳۹۶, ۰/۴۷۱)$$

$$(\pi^2 r_l, \pi^2 r_u) = (\pi^2 \times ۰/۰۶۳, \pi^2 \times ۰/۰۷۵) = (۰/۰۱۲۵, ۰/۰۱۷۷).$$

جدول ۶: برآوردهای β_i ها، خطاهای معیار، آمارهای آزمون و مقدار احتمال برای آزمون فرض‌های $H_{0,i} : \beta_i = 0$ در مثال ۲.

$P.V_i$	t_i	$\sqrt{Var(\hat{\beta}_i)}$	$\hat{\beta}_i$	i
۰/۰۰۰۰	۱۰/۳۳۸۶	۰/۰۰۸۱	۰/۰۸۴	۰
۰/۰۰۰۲	-۸/۲۹۰۲	۰/۰۰۴۰	-۰/۰۳۳	۱
۰/۰۰۰۰	۱۰/۵۰۱۷	۰/۰۰۰۴	۰/۰۰۴	۲


 شکل ۵: نمودار پراکنش r_i^* در مقابل x .

لازم به ذکر است که پارامترهای (Z_0, Ax, Ux) از پارامترهای مهم در مطالعه‌ی مجموعه‌های تصادفی هستند و معمولاً در بررسی تحقیقی از یک مجموعه‌ی تصادفی به برآورد آنها اهمیت داده می‌شود. این مثال نشان می‌دهد در صورت داشتن اطلاعات مشابه شکل‌های ۲ یا ۴ و برآزش یک مدل رشد به آنها نه تنها می‌توان در مجموعه‌ی مشاهدات، بلکه برای x ‌های دیده نشده نیز این پارامترها را به خوبی برآورد کرد.

بحث و نتیجه گیری

در این مقاله مدل‌های رشد با دانه‌های دیسکی شکل را در نظر گرفتیم. با توجه به اینکه معمولاً مدل‌های بولی با دانه‌های دیسکی شکل برای تقریب مسائل واقعی به

کار می‌روند، مدل‌های رشد مورد بحث می‌توانند برای تقریب مسائل واقعی مناسب باشند. از طرفی برای برآش این مدل‌ها فقط نیاز به \hat{p}_i و n_i^+ در پنجره‌ی Δ داریم. اولی را برای هر نوع تحقیقی و دومی را در حالتی که دانه‌های محدب باشند می‌توان محاسبه کرد. بنابراین از مدل و روش برآش ارائه شده در مواردی که دانه‌ها محدب باشند نیز به عنوان تقریب می‌توان استفاده کرد.

دیدیم که برآش مدل رشد منجر به برآش یک مدل رگرسیونی به مشاهدات (X_i, r_i^*) شد. در این مقاله به بحث در روش برآوردهای توان‌های دوم برای برآش این مدل رگرسیونی و نتایج حاصل از آن پرداختیم. بدیهی است که سایر روش‌های برآش مدل‌های رگرسیونی هم می‌توانند مورد مطالعه و استفاده قرار بگیرند.

تقدیر و تشکر

از معاونت پژوهشی دانشگاه شهید بهشتی برای حمایت از طرح پژوهشی وابسته به این مقاله تشکر و قدردانی می‌کنم.

مراجع

خزائی، م. (۱۳۸۶): روش‌هایی برای برآش مدل‌های رگرسیونی رشد، طرح پژوهشی، دانشگاه شهید بهشتی.

Cressie, N. (1993), *Statistics for Spatial Data*, John. Wiley & Sons, New York.

Kendall, D. G. (1974), Foundations of a Theory of Random Sets, In *Stochastic Geometry*, (ed. E. F. Harding and D. G. Kendall), Wiley, Chichester, 332-376.

۷۱ مجتبی خزائی

Khazaee, M. (2004), Regression Models for the Boolean Random Sets,
Ph.D. Thesis, Department of Statistics, Shahid Beheshti University,
Tehran, Iran.

Khazaee, M., and Shafie, K. (2006), Regression Model for Boolean Random Sets, *Journal of Applied Statistics*, **33(5)**, 557-567.

Mase, S. (1982), Asymptotic Properties of Stereological Estimators of Volume Fraction for Stationary Random Sets, *Journal of Applied Probability*, **19**, 111-126.

Matheron, G. (1975), *Random Sets and Integral Geometry*, John Wiley & Sons, New York.

Molchanov, I. S. (1995), Statistics of the Boolean Model: From the Estimation of Means to the Estimation of Distributions, *Advances in Applied Probability*, **27**, 63-86.

Molchanov, I. S. (1997), *Statistics of the Boolean Model for Practitioners and Mathematicians*, Wiley, Chichester.

Roystone, S. (1982), An Extension of Shapiro and Wilk's W Test for Normality to Large Samples, *Applied Statistics*, **31(2)**, 115-124.

Stoyan, D., Kendall, W. S., and Mecke, J. (1995), *Stochastic Geometry and Its Applications*, Wiley, Chichester.

Stoyan, D., Stoyan, H., Tscheschel, A. and Mattfeldt, T. (2001), On The Estimation of Distance Distribution Functions for Point Process and Random Sets, *Image Anal. Stereol*, **20**, 65-69.