آنالیز تکسونومی

(روش طبقه‌بندی گروههای همگن)

و کاربرد آن در طبقه‌بندی شهرستانها و ایجاد شاخصهای توسعه جهت برنامه‌ریزی منطقه‌ای

تهیه کننده : بیژن بیدآباد

سازمان برنامه و بودجه استان مرکزی (اراک)

تیرماه ۱۳۶۲

مقدمه

اگر چه ممکن است بسادگی اظهار نظر کرد که شهرستانی از شهرستان دیگر از هر لحاظ توسعه یافته‌تر است یا عقب‌افتاده‌تر، ولی اندازه‌گیری کمّی توسعه یافتگی کار ساده‌ای نمی‌تواند باشد. واژه توسعه دارای معانی بسیاری است. آیا منظور از توسعه یافتگی توسعه اقتصادی، اجتماعی، آموزشی، فرهنگی و یا بهداشتی است؟ و یا ترکیبی از آنها یا جنبه‌های دیگری را در بر می‌گیرد؟ حتی اگر توسعه را مفهوم دقیق تعریف شده‌ای در نظر بگیریم، اندازه‌گیری آن تازه ایجاد مسئله می‌کند و بزرگترین ایرادی که در این زمینه بوجود می‌آید نبودن آمار و ارقام صحیح و کافی است. مشکل دوم روش ساختن خود شاخص است که اگر یک متدولوژی قوی جهت ساختن شاخص بتوانیم پیدا کنیم کمک زیادی به رفع نواقص و نبودن آمارها می‌کند. روشی که در این مقاله اتخاذ می‌کنیم روش تکسونومی است که احتیاج به آمارهای سری زمانی نداشته و در بعضی موارد که در متن توضیح داده خواهد شد می‌تواند بعضی از کمبودها را نیز برطرف کند. از جهتی دیگر تمام شاخص‌ها مجبور هستند که ارزشهائی را با ضرائب یا پارامترهائی در مدل‌های محاسبه خود وارد کنند که شاید از لحاظ علمی قابل قبول نباشد و تغییرات زیادی به دلخواه در نتایج نهائی مدل بدهد. بدین دلیل باید روشی اتخاذ کرد که مقادیر ارزشی را یا حذف کند یا آنها را بطور خودکار مانند متغیرهای درون‌زا از داخل خود مدل تعیین کند. مشکل دیگر حذف مقیاسهای متفاوت در بدست آوردن این شاخص‌ها می‌باشد که می‌تواند سنگ بزرگی در راه ایجاد اینگونه شاخص‌ها گردد.

در این مطالعه از روش‌های مختلف طبقه بندی – درجه بندی و مقایسه، روش آنالیز تکسونومی بکار گرفته شده است که تا حدی اشکالات مذکور در فوق را برطرف می‌کند ولی با این همه روشی است که احتیاج به بررسیهای بیشتری در قسمت چارچوب آمارهای مورد نیاز جهت ساختن شاخص دارد که البته احتیاج به آنالیزهای سینکرونیک Synchronic و یا دایاکرونیک Diachronic با استفاده از ضرائب همبستگی و محاسبات رگرسیون دارد.

می‌توان هر جنبه از توسعه را به عنوان یک شاخص در نظر گرفت و نتایج آنها را به عنوان یک شاخص ترکیبی محاسبه کرد بدین ترتیب که اول باید قبلاً شاخصی برای توسعه بهداشت در تمام شهرستانهای مورد نظر بوجود آورد که می‌توان از روش تکسونومی استفاده کرد که نمایانگر یک شاخص است که جنبه‌های مختلف توسعه بهداشت را در نظر می‌گیرد و یک شاخص خام نیست. سپس می‌توان شاخص‌های دیگری را مانند شاخص توسعه بهداشت برای مثلاً توسعه اقتصادی، فرهنگی، آموزش و تعلیم و تربیت و سایر جنبه‌های مورد نظر دیگر بوجود آورد و بعد همه شاخص‌ها را دوباره تبدیل به یک مدل تکسونومی کرد و یک شاخص ترکیبی توسعه بدست آورد که تمام جنبه‌های مذکور را در بر دارد، یا می‌توان از اول با استفاده از تمام آمارهای موجود یک شاخص بدست آورد.

به هر حال در این مطالعه سعی می‌شود که وضعیت بهداشت شهرستانهای استان مرکزی تا حدودی مورد بررسی واقع شود ولی به علت عدم دسترسی به کامپیوتر نمی‌توانیم نتایج محاسبات خود را بسط دهیم و بـه همـین دلیل فقط به بررسی وضعیت بهداشت اکتفا می‌کنیم.

منظور ما از این مقاله ارائه روشی می‌باشد که مورد نیاز برنامه‌ریزان در تمام سـطوح مـی‌باشـد مخصوصـاً برای کارشناسان برنامه‌ریزی استانها، البته باید عنوان کرد که این متن مقدماتی بوده و ایرادات زیـادی بـه آن وارد است که انشاءالله در آینده با همکاری سایر همکاران از بین خواهد رفت. در آخر باید اضافه کـرد کـه در نوشـتن این مقاله از کتاب :

Quantitative Analysis of Modernization and Development F. H. Harbison, J. R. Resnick, Princeton University, Newjercy.

خیلی کمک گرفته شده است.

آنـالیز تکسـونومی را مـی‌تـوان بـا اسـتفاده از روشـهای تکمیلـی دیگـری ماننـد نمـودار زکانوسـکی Czekanowski’s diagram نمودار خط شکسته dendrid روش درخت ربط linkage tree و غیره گسـترش داد که شرح آن خارج از حوصله این مقاله است و می‌توان به مقاله زیر رجوع نمود:

Methods and techniques helpful on the planning and development of rural areas, Michael Chilczuk, Plan and Budget Organization, United Nations Development Program, Center for Research and Training in Regional Planning, Dec. 1974, Iran.

آنالیز تکسونومی

از میان روش‌های مختلف درجه‌بندی مناطق از لحاظ توسعه یافتگی یکی روش آنالیز تکسونومی می‌باشد. آنالیز تکسونومی برای طبقه‌بندیهای مختلف در علوم بکار برده می‌شود نوع خاص آن تکسونومی عددی Numerical Taxonomy است که بنا به تعریف P.H.A. Sneath, R.R. Sokal [1] «ارزشیابی عددی شباهت‌ها و نزدیکی‌ها» بین واحدهای تکسونومیک و درجه‌بندی آن عناصر به گروههای تکسونومیک (تکسون) می‌باشد. این روش اولین بار توسط M. Adanson در سال ۱۷۶۳ پیشنهاد شد. استفاده از این روش از توسعه‌های اخیر آن است که توسط دسته‌ای از ریاضی‌دانان لهستانی در اوائل دهه ۱۹۵۰ بسط داده شد و در سال ۱۹۶۸ به عنوان وسیله‌ای برای طبقه‌بندی و درجه توسعه یافتگی بین ملل مختلف توسط پروفسور Zygmunt Hellwig از مدرسه عالی اقتصاد Wroclaw در یونسکو مطرح شد.[2] روش تکسونومی روکلا Wroclaw یک روش عالی درجه‌بندی طبقه‌بندی و مقایسه کشورها یا مناطق مختلف با توجه به درجه توسعه و مدرن بودن آنها می‌باشد.

تکسونومی روکلا یک روش آماری برای تعیین واحدها یا انواع چیزهای همگن در یک فضای برداری n بعدی بدون استفاده از رگرسیون، واریانس یا آنالیز همبستگی می‌باشد. روش تکسونومی قادر است که یک مجموعه را به زیر مجموعه‌های کم و بیش همگن تقسیم کرده، ابزار مفیدی برای انترپولاسیون و اکستراپولاسیون آمارها بدست دهد و یک مقیاس برای شناخت درجه توسعه اقتصادی و اجتماعی که مورد استفاده در برنامه‌ریزی باشد قرار بگیرد.

حال شروع می‌کنیم به بررسی این روش، مجموعه Z را در نظر می‌گیریم که شامل N عضو بوده که بیانگر شهرستانهای مختلف ۱ و ۲ و ....و N باشد. برای یک گروه از متغیرهای ۱ و ۲ و ....و m (که عبارت است از m شاخص (خصوصیت) مورد نظر ما در مطالعه فعلی) پس می‌توان گفت که داده‌های ما به شکل زیر است.

$$(1) \quad P_1\,(X_1, X_2, ..., X_m), P_2\,(X_1\, X_2, ...X_m), ..., P_N\,(X_1, X_2, ..., X_n)$$

شکل فوق را می‌توانیم با ماتریس زیر نشان دهیم:

[1] P.H.A. Sneath and R.R. Sokal, Numerical Taxonomy 1963.

[2] K.Florek, et al, Taksonomia Wroclawska (Wroclaw Taxonomy), Poznan, 1952, Z. Hellwig, "Procedure of evaluating high–level manpower data and Typology of countries by means of the taxonomic method". (Unpublished UNESCO Working Paper 1967).

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} & & X_{1m} \\ X_{21} & X_{22} & X_{23} & & X_{2m} \\ X_{31} & X_{32} & X_{33} & & X_{3m} \\ . & . & & & \\ . & . & & & \\ . & . & & & \\ . & . & & & \\ . & . & & & \\ X_{n1} & N_{2} & X_{n3} & & X_{nm} \end{bmatrix} \quad \begin{array}{l} i = 1,2,...,n \\ j = 1,2,...,m \end{array}$$

ماتریس الف

بدین ترتیب هر شهرستانی توسط یک بردار در یک فضای m بعدی نشان داده می‌شود که X_{ij} نشان دهنده خصوصیت j ام از شهرستان iام می‌باشد. با توجه به اینکه تمام این خصوصیات دارای مقیاس‌های متفاوتی می‌باشند باید کاری کرد که دخالت مقیاسهای متفاوت را از داخل مدل از بین برد. بدین لحاظ در قدم اول میانگین ستونها را بدست می‌آوریم.

$$(2) \quad \bar{X}_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij}$$

قدم بعدی انحراف استاندارد را برای هر ستون از ماتریس الف پیدا کرده :

$$(3)\ S_j = \sqrt{\frac{\sum_{t=1}^{n}(X_{ij} - \bar{X}_j)^2}{n}}$$

قدم سوم این است که عضوهای جدید ماتریسی بنام D را تشکیل می‌دهیم که عبارتند از:

$$(4)\ D_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j}$$

و ماتریس D. یک ماتریس nxm خواهد بود:

$$\begin{bmatrix} D_{11} & D_{12} & .... & D_{1m} \\ D_{21} & D_{22} & .... & D_{2m} \\ . & . & & \\ . & . & & \\ . & . & & \\ D_{n1} & D_{n2} & .... & D_{nm} \end{bmatrix}$$

ماتریس ب

حال ماتریس D خالی از هرگونه مقیاس می‌باشد، میانگین هر ستون برابر صفر بوده $\bar{D}_j = 0$ چون اگـر از دو طرف تساوی (۴) $\sum_{i=1}^{n}$ بگیریم و بر n تقسیم کنیم حاصل برابر خواهد بود با:

$$\frac{\sum_{i=1}^{n} D_{ij}}{n} = \frac{\sum_{i=1}^{n}(X_{ij} - \bar{X}_j)}{nS_j} = \frac{0}{nS_j} = 0 \qquad (۵)$$

و انحراف استاندارد هر ستون آن برابر با یک خواهد بود چون :

$$S_j = \sqrt{\frac{\sum_{i=1}^{n}(D_{ij} - \bar{D}_j)^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(D_{ij})^2}{n}} == \sqrt{\frac{\sum_{i=1}^{n}\left(\frac{X_{ij} - \bar{X}}{S_j}\right)^2}{n}} =$$

(۶)

$$\sqrt{\frac{\sum (X_{ijx} - \bar{X}_j)^2}{n(S_j)^2}} = \sqrt{\frac{\frac{\sum_{i=1}^{n}(X_{ij} - \bar{X}_j)^2}{n}}{S_j}} = \frac{S_j}{S_j} = 1$$

داشتن میانگین صفر و انحراف استاندارد یک برای هر ستون کمک به کنترل صحت مـاتریس D مـی‌کنـد. بـدین معنی می‌توان آزمایش کرد که آیا نتیجه محاسبات تا رسیدن به ماتریس D درست بوده یا غلط.

با داشتن ماتریس اسـتاندارد D قـدم بعـدی بدسـت آوردن اخـتلاف یـا فاصـله هـر نقطـه از نقطـه دیگـر (۱و۲و ....وn) برای هر کدام از m متغیر یا خصوصیت می‌باشد که حاصل آن ماتریس (ج) خواهد بود. مـاتریس ج متشکل از n − ۱ ماتریس است کـه مـی‌تـوان آنهـا را توسـط n − ۲ پارتیشـن افقـی از هـم مجـزا سـاخت ابعـاد ماتریس‌های پارتیشن شده به ترتیب از بالا به پائین برابر خواهـد بـود بـا: (n-1)×m و (n-2)×m و (n-2)×m و ... و 2×m و ۱×m در نتیجه ابعاد ماتریس ج برابر خواهد بود با:

$$\left(\sum_{i=1}^{n-1} i\right) \times m \qquad (۷)$$

یا

$$(1 + 2 + 3 + ... + n - 1) \times m \qquad (۸)$$

$$
\begin{array}{cc}
\left[\begin{array}{llll}
D_{11}-D_{21} & D_{12}-D_{22} & \cdots\cdots & D_{1m}-D_{2m} \\
D_{11}-D_{31} & D_{12}-D_{32} & \cdots\cdots & D_{1m}-D_{3m} \\
D_{11}-D_{41} & D_{12}-D_{43} & \cdots\cdots & D_{1m}-D_{4m} \\
. & . & & \\
. & . & & \\
. & . & & \\
D_{11}-D_{n1} & D_{12}-D_{n2} & \cdots\cdots & D_{1m}-D_{nm} \\
\cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\
D_{21}-D_{31} & D_{22}-D_{32} & \cdots\cdots & D_{2m}-D_{3m} \\
D_{21}-D_{41} & D_{22}-D_{42} & \cdots\cdots & D_{2m}-D_{4m} \\
D_{21}-D_{51} & D_{22}-D_{n2} & \cdots\cdots & D_{2m}-D_{5m} \\
. & & & \\
. & & & \\
D_{21}-D_{n1} & D_{22}-D_{n2} & \cdots\cdots & D_{2m}-D_{nm} \\
\cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\
. & & & \\
. & & & \\
. & & & \\
\cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\
D_{(n-2)1}-D_{(n-1)1} & D_{(n-2)2}-D_{(n-1)2} & \cdots\cdots & D_{(n-2)m}-D_{(n-1)m} \\
D_{(n-2)1}-D_{(n\ )1} & D_{(n-2)2}-D_{(n\ )2} & \cdots\cdots & D_{(n-2)m}-D_{(n-1)m} \\
\cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\
D_{(n-1)1}-D_{n1} & D_{(n-1)2}-D_{n2} & \cdots\cdots & D_{(n-1)m}-D_{nm}
\end{array}\right]
&
\begin{array}{l}
\\ \\ \\ \\
(n-1)\times m \\
\\ \\ \\ \\ \\
(n-2)\times m \\
\\ \\ \\ \\ \\ \\ \\ \\
2\times m \\
\\ \\
1\times m
\end{array}
\end{array}
$$

ماتریس ج

حالا برای پیدا کردن فاصله بین دو نقطه P_a و P_b بـرای هـر مجموعـه یـا زیرمجموعـه از متغیرهـای m فرمول زیر را بکار می‌بندیم:

$$C_{ab}=\sqrt{\sum_{k=1}^{m}(D_{ak}-D_{bk})^2} \qquad (۹)$$

C_{ab} را برای n و ... و ۲ و ۱ = b و a بدست می‌آوریم. واضح است که C_{aa} یعنی فاصله a از a مساوی صـفر است و $C_{ab}=C_{ba}$ یعنی فاصله a تا b مساوی b تا a است و $C_{ab}\le C_{ak}+C_{kb}$. حاصل را درون یک ماتریس قرار می‌دهیم به نام ماتریس فواصل یا ماتریس (د).

$$C = \begin{bmatrix} 0 & C_{12} & C_{13} \cdots\cdots\cdots & C_{1n} \\ C_{21} & 0 & C_{23} \cdots\cdots\cdots & C_{2n} \\ C_{31} & C_{32} & 0 \cdots\cdots\cdots & C_{3n} \\ . & . & & \\ . & . & & \\ . & . & & \\ . & . & & \\ C_{n1} & C_{n2} & C_{n3} \cdots\cdots\cdots & 0 \end{bmatrix}$$

ماتریس (د)

ماتریس (د) یا ماتریس فواصل دارای این خصوصیت است که اولاً قرینه بوده و قطر اصلی آن صفر می‌باشد. عضوهای ماتریس (د) فاصله ترکیبی هر شهرستان را از شهرستان دیگر را نشان می‌دهد به عبارت دیگر بیان ریاضی چند فاصله بر هرکدام از چند ابعادی است که شهرستانها می‌توانند با هم مقایسه شوند.

در هر ردیف کوچکترین فاصلهٔ C_a از آن شهرستان تا شهرستانهای دیگر را می‌توان پیدا کرد که شاخصی است برای شباهت آن شهرستان به شهرستان‌های دیگر، از آنجائی که حداقل فاصله بین نقطه P_a و سایر نقاط در ردیف a عدد C_{ab} است، P_a را «الگوی» P_a و P_a را «سایهٔ» P_b می‌نامیم. به عبارت دیگر در هر ردیف کمترین مقدار نشان دهنده کوتاهترین فاصله و شماره ستون مربوط به این کوتاهترین فاصله نمایانگر شهرستانی است که به شهرستان مزبور (شماره ردیف) از همه نزدیکتر است مثلاً اگر C_{ab} را در نظر بگیریم که ردیف a ام کوچکترین مقدار است شهرستان bام نزدیکترین شهرستان به شهرستان aام می‌باشد که b الگوی شهرستان a و a سایه شهرستان b می‌باشد. ابهامی ممکن است به نظر برسد که اگر زمانی بیش از یک فاصله مساوی در یک ردیف نسبت به نقطه Pa باشد چه باید کرد؟ در جواب باید گفت که شانس چنین واقعه‌ای تقریباً نزدیک به صفر است بنابراین این فرض را قرار می‌دهیم که یک نقطه و فقط یک نقطه وجود دارد که نزدیکترین نقطه است.

قدم بعدی تعیین روابط نموداری (گراف) یا ارتباطی است به این معنی که باید کوتاه‌ترین نمودار (گراف) خطی که بیانگر گروههای شهرستان‌ها است را پیدا کنیم.

در مرحلهٔ اول باید هر سایه‌ای را به الگویش متصل کنیم که در گراف مرتبه اول عمل متمرکز کردن گراف گروههای همگن نامتصل تهیه شود. سپس عمل متمرکز کردن مرتبه دوم را باید انجام داد. کوتاهترین فاصله بین هر دو node (گره) متعلق به دو گروه همگن مختلف می‌باشد (با بدست آوردن دومین کوتاه‌ترین

فواصل از ماتریس فواصل) و این عمل را تکرار می‌کنیم تا تمام node ها یا گره‌ها در داخل یک گراف پیوسته قرار بگیرند.[۳]

زمانیکه گراف پیوسته واحد تعیین می‌شود، در مرتبهٔ اول ۱- n پیکان node ها را به هم متصل می‌کنند که در مرتبه‌های بعدی تعداد پیکان‌ها هم کمتر می‌شود ۱ − k پیکان را از میان آنها حذف می‌کنیم عدد k توسط «فاصله (یا الگوی) حداقل بحرانی» تعیین می‌شود. فاصله حداقل بحرانی از فرمول زیر محاسبه می‌شود.

$$C_{(+)} = \bar{C} + 2S_c \qquad (۱۰)$$

که مقدار $\bar{C}$ میانگین حسابی فاصله‌های C_j (کوچکترین فاصله در هر ردیف از ماتریس فواصل) است:

$$\bar{C} = \frac{1}{N}\sum_{j=1}^{N} C_j \qquad (۱۱)$$

و S_c انحراف استاندارد و کوتاهترین فواصل در تمام ردیف‌ها است.

$$S_c \sqrt{\frac{1}{n}\sum_{j=1}^{n} (C_j - \bar{C})^2} \qquad (۱۲)$$

عدد k تعداد زیر گروههای متصل شده توسط خطوط رابط (پیکانها) است که در گراف اپتیمال (پیوسته واحد) بلندتر از $C_{(+)}$ است. تعداد n عضو در مجموعه شهرستانها را نیز می‌توان با استفاده از ارزش بحرانی دیگری کاهش داد:

$$C_{(-)} = \bar{C} - 2S_c \qquad (۱۳)$$

اعداد ۲ در فرمول‌ها بدین منظور در نظر گرفته شده‌اند که در زیر منحنی توزیع نرمال به اندازه دو انحراف استاندارد از چپ و راست را در نظر گرفته‌ایم و احتمال اینکه کمیتهای استاندارد شده در یک مجموعه خارج از سطح مورد نظر ما باشد ۰/۰۵٪ است. حالا تمام node هائی که فواصل آنها از $C_{(-)}$ کوتاه‌تر هستند حذف می‌شوند. زیرمجموعه‌های مجموعه حاصل شده توسط پارتیشن کشیدن در مجموعه اصلی تقسیم آن به k قسمت گروههایی پیدا می‌شوند که به آن گروههای تیپولوژیک می‌گوئیم (این عمل در گراف پیوسته واحد انجام می‌شود). هر زیرگرافی یک گروه تیپولوژیک متمایز ایجاد می‌کند که در برگیرنده عضوهائی است که واحد یا شبیه بنظر می‌رسند. ارزش بحرانی نیز می‌تواند به عنوان مقیاس شباهت در نظر گرفته شود، هرچه $C_{(+)}$ بزرگتر باشد شباهت بین تمام جفت نقطه‌های ممکن کمتر است. این مقیاس برای مقایسه آنالیزهای تکسونومیک مختلف کاربرد زیادی دارد. به هر حال باید به خاطر داشت که در مطالعه گراف اپتیمال فقط طول و جهت پیکان‌ها را باید در نظر گرفت نه محل node ها را. به عبارت دیگر گراف اپتیمال تصویرهای دوبعدی از یک فضای n بعدی است.

---

[۳] – پیوسته شدن گراف توسط Florek et al ثابت شده است که گراف پیوسته بوجود آمده کوتاهترین گراف است که نام آن را گراف متراکم می‌گذاریم. رجوع کنید به:

K.Florek et al . Taksonomia wroclawska, Poznan 1952 , Z. Hellwig, Procedure of evaluating high–level man power data and typology of countries … UNESCO.

از اشکالاتی که ممکن است در سر راه روش تکسونومی پیدا شود این است که اگر بطور مثال یک شهرستان جدید با m متغیر (خصوصیت) به مجموعه n شهرستان مورد مطالعه خود اضافه کنیم که در بعضی از خصوصیت‌ها آمار لازم را نداشته باشد، در درجهٔ اول لازم است که ارزشهای استاندارد شده متغیرهای موجود برای شهرستان A (شهرستانی که اضافه کرده‌ایم) را با فرمول‌های استاندارد کردن (۴ و ۳ و ۲) بر مبنای شهرستانی که قبلاً در گروه بوده‌اند را محاسبه کنیم. بر این مبنا فاصله شهرستان A از تمام شهرستانها در ماتریس اصلی قابل محاسبه‌اند. کوتاهترین فاصله یا نزدیکترین شهرستان الگوی فرضی یا پتانسیل شهرستان A و اعداد (ارزشهای) شهرستان الگو را میتوان بعنوان تقریبهائی از آمارهای ندانسته شهرستان A دانست. به هر حال این روش ممکن است خطای تصادفی زیادی را وارد مدل کند و به این دلیل فقط برای برآورد اولیه می‌تواند مفید واقع شود.

روش دیگری که می‌توانیم بکار بندیم این است که میانگین ارزشهای آن متغیر را برای همه شهرستانهایی که فاصله آنها به شهرستان A کمتر از فاصله الگوی بحرانی است محاسبه کنیم (معادله ۱۰) بنابراین یک میانگین به عنوان تقریب استفاده می‌شود تا فقط یک مقدار.[۴]

باید در اینجا یادآور شد که وقتی یک شهرستانی فقط به یک گروه معینی می‌تواند اضافه شود که فاصله به حداقل یک شهرستان در ماتریس از فاصله الگوی بحرانی کمتر باشد شاید ساده‌ترین روش تعیین این موضوع انتخاب یک متغیر تست است که باید آن را نسبت به چندین گروه استاندارد کرده و سپس آن گروهی انتخاب شود که کمترین مقدار را دارد.[۵]

در برنامه‌ریزی انتخاب هدف همیشه اولین عامل است و هدف ما در برنامه‌ریزی منطقه‌ای ممکن است یکسان کردن درجه توسعه یافتگی شهرستانها باشد با توجه به این موضوع، اشکال و اندازه‌های توسعه یافتگی در بین شهرستانها را از روش تکسونومی می‌توان بدست آورد. از این لحاظ و برای برنامه‌ریزی برای یکسان کردن شهرستانها می‌توانیم برای هر متغیر در n شهرستان بزرگترین مقدار را به عنوان مقدار ایده‌آل در نظر بگیریم. بدین ترتیب که در هر ستون از ماتریس ب بزرگترین مقدار استاندارد شده را بدست می‌آوریم به شرطی که بدانیم توسعه یافتگی یک تابع افزاینده از آن متغیر است در صورتی که توسعه یافتگی یک تابع کاهنده از آن متغیر باشد باید بزرگترین مقدار منفی استاندارد شده از ماتریس ب را در هر ستون در نظر بگیریم.[۶]

حال واژه‌ای به عنوان «سرمشق توسعه» در نظر می‌گیریم و آن را با C_{io} نشان می‌دهیم که عبارت است از فاصله شهرستان i در ماتریس تا فاصله شهرستان ایده‌آل o که از فرمول زیر محاسبه می‌شود.

---

۴ – می‌توان تست‌های آماری را هم برای با معنی بودن مورد استفاده قرار داد بدین ترتیب که انحراف استاندارد اعداد آن شهرستانهایی که برای میانگین استفاده می‌شوند را محاسبه کرد.

۵ – اگر مسئله کمبود آمار را می‌خواستیم با تجزیه و تحلیل رگرسیون چند بعدی برطرف کنیم تعداد زیادی معادلات رگرسیون مجزا می‌بایست محاسبه شود بدون اینکه قادر به آزمایش نوع رابطه فانکشال باشد در نتیجه لزوماً می‌بایست با بکاربردن توابع خطی و قبول ریسک خطای بزرگ ایجاد شده توسط فرض خطی بودن در امکان وجود و همبستگی متقابل استوکستیک بین متغیرها درگیر شویم.

۶ – مثلاً توسعه یافتگی یک تابع افزاینده از مصرف گوشت می‌باشد در صورتیکه یک تابع کاهنده از مصرف کره نباتی (مارگارین) است.

$$C_{io}=\sqrt{\sum_{k=1}^{m}(D_{ik}-D_{ok})^2} \quad (۱۴)$$

که N و ..... و ۲ و ۱ = i و o بزرگترین مقدار استاندارد شده است که از ماتریس ب پیدا می‌شود. هر چقدر عدد C_{io} بزرگتر باشد نشان دهنده فاصله بیشتر شهرستان i تا شهرستان ایده‌آل است. اندازه‌گیری توسعه یک روش سیمولیتینگ درصد توسعه یافتگی در یک منطقه بخصوص است به عبارت دیگر یک تابعی است از سرمشق توسعه و فاصله بحرانی از شهرستان ایده‌آل. فرمول‌های زیر را می‌توان برای این موضوع بکار بست.

$$d_i=\frac{C_{io}}{C_o}$$

اندازه توسعه یافتگی شهرستان $d_i=i$

که

$$C_o=\bar{C}_{io}+2S_{io}$$

و

$$\bar{C}_{io}=\frac{1}{n}\sum_{i=1}^{n}C_{io}$$

$$S_{io}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(C_{io}-\bar{C}_{io})^2}$$

میانگین سرمشق توسعه = $\bar{C}_{io}$

انحراف استاندارد سرمشق توسعه = S_{io}

هر چقدر d_i به صفر نزدیکتر باشد نشانه توسعه یافتگی بیشتر و هرچقدر به یک نزدیکتر باشد علامت توسعه نیافتگی بیشتر است و حدود دامنه تغییرات d بین یک و صفر می‌باشد. ممکن است که d از یک بیشتر هم بشود و احتمال وقوع چنین حالتی خیلی کم است و در اکثر حالات نامساوی زیر برقرار است. $0<d<1$

حال فرض کنیم می‌خواهیم الگوی یک شهرستانی را در زمینه‌های بخصوص توسعه یافتگی پیدا کنیم. فرض می‌کنیم C_{ba} فاصله دو شهرستان B و A (p_b+P_a); C_{bo} سرمشق توسعه (فاصله بین شهرستانها B و شهرستان ایده‌آل) و d_b نمایانگر اندازه توسعه است. برای تعیین الگوهای توسعه اجتماعی و اقتصادی برای شهرستان B باید تمامی نقاطی را که فاصله آنها کمتر از فاصله الگوی بحرانی است و همه آنها اندازه توسعه بزرگتری دارند را پیدا می‌کنیم. بنابراین تمام نقاط P_a را پیدا می‌کنیم. بطوریکه:

$$\begin{cases} d_a \langle\ d_b \\ \quad , \\ C_{ba} \langle C_{(+)} \end{cases}$$

مجموعه نقاطی که شرایط بالا را دارند را با Z^* نشان می‌دهیم، میانگین حسابی همه مولفه‌های مجموعه Z^* را حساب کرده و نقطه $P_b^*(X_1,X_2,...,X_m)$ را ایجاد می‌کنیم که می‌توان به عنوان الگوی توسعه برای شهرستان B در نظر گرفته شود.

خصوصیت شهرستان	تخت بیمارستان برای هر ۱۰۰۰۰ نفر	پزشک برای هر ۱۰۰۰۰ نفر	مراکز بهداشتی برای هر ۱۰۰۰۰ نفر	خانه بهداشت برای هر ۱۰۰۰۰ نفر	$X_{i1}-\bar{X}_1$	$X_{i2}-\bar{X}2$	$X_{i3}-\bar{X}3$	$X_{i4}-\bar{X}4$	$(X_{i1}-\bar{X}_1)^2$	$(X_{i2}-\bar{X}2)^2$	$(X_{i3}-\bar{X}3)^2$	$(X_{i4}-\bar{X}4)^2$
اراک	۴۵	۳	۰/۵۹	۰/۸۲	۳۵	۰/۷۳۵	−۰/۵۷۶	−۰/۷۲۳	۱۲۲۵	۰/۵۳۹	۰/۳۳۱	۰/۵۲۲
سربند	۰	۱/۴۳	۰/۶۷	۱/۳	−۱۰	۰/۸۳۵	−۰/۴۹۶	۰/۲۴۳	۱۰۰	۰/۶۹۸	۰/۲۴۶	۰/۰۵۹
ساوه	۸	۱/۵	۱	۳/۵	−۲	−۰/۷۶۵	−۰/۱۶۶	۱/۹۵۷	۴	۰/۵۸۶	۰/۰۲۷	۳/۸۲۹
آشتیان	۸	۲	۱	۰	−۲	−۰/۲۶۵	−۰/۱۶۶	−۱/۵۴۳	۴	۰/۰۷۰	۰/۰۲۷	۲/۳۸۰
تفرش	۰	۲	۲	۰/۷۷	−۱۰	−۰/۲۶۵	۰/۸۳۴	−۰/۷۷۳	۱۰۰	۰/۰۷۰	۰/۶۹۵	۰/۵۹۷
قم	۲۰	۱/۹۶	۰/۴۶	۰/۵۰	۱۰	−۰/۳۰۵	−۰/۷۰۶	−۱/۰۴۳	۱۰۰	۰/۰۹۳	۰/۴۹۸	۱/۰۸۷
خمین	۹	۱	۰/۷۸	۰/۵	−۱	−۱/۲۶۵	−۰/۳۸۶	−۱/۰۴۳	۱	۱/۶۰۰	۰/۱۴۸	۱/۰۸۷
دلیجان	۰	۴	۲	۱/۵	−۱۰	۱/۷۳۵	۰/۸۳۴	−۰/۰۴۳	۱۰۰	۳/۰۰۸	۰/۶۹۵	۰/۰۰۱۸
محلات	۰	۳/۵	۲	۵	−۱۰	۱/۲۳۵	۰/۸۳۴	۳/۴۵۷	۱۰۰	۱/۵۲۳	۰/۶۹۵	۱۱/۹۵۰
$\sum_j$	۹۰	۲۰/۳۹	۱۰/۵	۱۳/۸۹	−	−	−	−	۱۷۳۴	۸/۱۸۷	۳/۲۶۲	۲۱/۵۱۲۸
$\bar{X}^i$	۱۰	۲/۲۶۵	۱/۱۶۶	۱/۵۴۳	−	−	−	−	−	−	−	−
S_i	−	−	−	−	−	−	−	−	۱۳/۸۸۰	۰/۹۵۳	۰/۶۱۱	۱/۵۴۶

جدول (۱) ماتریس الف و محاسبات منتج از آن

ماتریس ب جدول ۲

خصوصیت / شهرستان	تعداد تخت بیمارستان برای هر ۱۰۰۰۰ نفر	تعداد پزشک برای هر ۱۰۰۰۰ نفر	تعداد مراکز بهداشتی برای هر ۱۰۰۰۰ نفر	تعداد خانه بهداشت برای هر ۱۰۰۰۰ نفر
اراک	۲/۵۲۱ D11	۰/۷۷۱ D12	−۰/۹۴۲ D13	−۰/۴۶۷ D14
سربند	−۰/۷۲۰ D21	−۰/۸۷۶ D22	−۰/۸۱۱ D23	−۰/۱۵۷ D24
ساوه	−۰/۱۴۴ D31	−۰/۸۰۲ D32	−۰/۲۷۱ D33	۱/۲۶۵ D34
آشتیان	−۰/۱۴۴ D41	−۰/۲۷۸ D42	−۰/۲۷۱ D43	−۰/۹۹۸ D34
تفرش	−۰/۷۲۰ D51	−۰/۲۷۸ D52	۱/۳۶۴ D53	−۰/۵۰۰ D54
قم	۰/۷۲۰ D61	−۰/320 D62	−۱/۱۵۵ D63	−۰/۶۷۴ D64
خمین	−۰/۰۷۲ D71	−۱/۳۲۷ D72	−۰/۶۳۲ D73	−۰/۶۷۴ D74
دلیجان	−۰/۷۲۰ D81	۱/۸۲۰ D82	۱/364 D83	−۰/۰۲۷ D84
محلات	−۰/۷۲۰ D91	۱/۲۹۵ D92	۱/۳۶۴ D93	۲/۲۳۶ D94
$\bar{X}_j$	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
S.D.	۱	۱	۱	۱

ماتریس ج   جدول ۳

D11-D21	۳/۲۴۱	D12– D22	۱/۶۴۷	D13-D23	-۰/۱۳۱	D14-D24	-۰/۳۱۰
D11-D31	۲/۶۶۵	D12-D32	۱/۵۷۳	D13-D33	-۰/۶۷۱	D14-D34	-۱/۷۳۲
D11 –D41	۲/۶۶۵	D12-D42	۱/۰۴۹	D13-D43	-۰/۶۷۱	D14-D44	۰/۵۳۱
D11-D51	۳/۲۴۱	D12-D52	۱/۰۴۹	D13-D53	-۲/۳۰۶	D14-D54	۰/۰۳۳
D11-D61	۱/۸۰۱	D12-D62	۱/۰۹۱	D13-D63	۰/۲۱۳	D14-D64	۰/۲۰۷
D11-D71	۲/۵۹۳	D12-D72	۲/۰۹۸	D13-D73	-۰/۳۱۱	D14-D74	۰/۲۰۷
D11-D81	۳/۲۴۱	D12-D82	-۱/۰۴۹	D13-D83	-۲/۳۰۶	D14-D84	-۰/۴۴
D11-D91	۳/۲۴۱	D12-D92	-۰/۵۲۴	D13-D93	-۲/۳۰۶	D14-D94	-۲/۷۰۳
D21-D31	-۰/۵۷۶	D22- D32	-۰/۰۷۴	D23-D33	-۰/۵۴۰	D24-D34	-۱/۴۲۲
D21-D41	-۰/۵۷۶	D22-D42	-۰/۵۹۸	D23-D43	-۰/۵۴۰	D24-D44	۰/۸۴۱
D21-D51	۰	D22-D52	-۰/۵۹۸	D23-D53	-۲/۱۷۵	D24-D54	۰/۳۴۳
D21-D61	-۱/۴۴	D22-D62	-۰/۵۵۶	D23-D63	۰/۳۴۴	D24-D64	۰/۵۱۷
D21-D71	-۰/۶۴۸	D22-D72	۰/۴۵۱	D23-D73	-۰/۱۸۰	D24-D74	۰/۵۱۷
D21-D81	۰	D22-D82	2/696-	D23-D83	-۲/۱۷۵	D24-D84	-۰/۱۳۰
D21-D91	۰	D22-D92	-۲/۱۷۱	D23-D93	-۲/۱۷۵	D24-D94	-۲/۳۹۳
D31-D41	۰	D32-D42	-۰/۵۲۴	D33-D43	۰	D34-D44	۲/۲۶۳
D31-D51	۰/۵۷۶	D32-D52	-۰/۵۲۴	D33-D53	-۱/۶۳۵	D34-D54	۱/۷۶۵
D31-D61	-۰/۸۶۴	D32-D62	-۰/۴۸۲	D33-D63	۰/۸۸۴	D34-D64	۱/۹۳۹
D31-D71	-۰/۰۷۲	D32-D72	۰/۵۲۵	D33-D73	0/360	D34-D74	۱/۹۳۹
D31-D81	۰/۵۷۶	D32-D82	-۲/۶۲۲	D33-D83	-۱/۶۳۵	D34-D84	۱/۲۹۲
D31-D91	۰/۵۷۶	D32-D92	-۲/۰۹۷	D33-D93	-۱/۶۳۵	D34-D94	-۰/۹۷۱
D41-D51	۰/۵۷۶	D42-D52	۰	D43-D53	-۱/۶۳۵	D44-D54	-۰/۴۹۸
D41-D61	-۰/۸۶۴	D42-D62	۰/۰۴۲	D43-D63	۰/۸۸۴	D44-D64	-۰/۳۲۴
D41-D71	-۰/۰۷۲	D42-D72	۱/۰۴۹	D43-D73	۰/۳۶۰	D44-D74	-۰/۳۲۴
D41-D-81	۰/۵۷۶	D42-D82	-۲/۰۹۸	D43-D83	-۱/۶۳۵	D44-D84	-۰/۹۷۱
D41-D91	۰/۵۷۶	D42-D92	-۱/۵۷۳	D43-D93	-۱/۶۳۵	D44-D94	-۳/۲۳۴
D51-D61	-۱/۴۴	D52-D62	۰/۰۴۲	D53-D63	۲/۵۱۹	D54-D64	۰/۱۷۴

D51-D71	-۰/۶۴۸	D52-D72	۱/۰۴۹	D53-D73	۱/۹۹۵	D54-D74	۰/۱۷۴
D51-D81	۰	D52-D82	-۲/۰۹۸	D53-D83	۰	D54-D84	-۰/۴۷۳
D51-D91	۰	D52-D92	-۱/۵۷۳	D53-D93	۰	D54-D94	۲/۷۳۶
D61-D71	۰/۷۹۲	D62-D72	۱/۰۰۷	D63-D73	۰/۵۲۴	D64-D74	۰
D61-D81	۱/۴۴	D62-D82	-۲/۱۴	D63-D83	-۲/۵۱۹	D64-D84	-۰/۶۴۷
D61-D91	۱/۴۴	D62-D92	-۱/۶۱۵	D63-D93	-۲/۵۱۹	D64-D94	-۲/۹۱۰
D71-D81	۰/۶۴۸	D72-D82	-۳/۱۴۷	D73-D83	۱/۹۹۵	D74-D84	-۰/۶۴۷
D71-D91	۰/۶۴۸	D72-D92	-۲/۶۲۲	D73-D93	-۱/۹۹۵	D74-D94	-۲/۹۱۰
D81-D91	۰	D82-D92	۰/۵۲۵	D83-D93	۰	D84-D94	-۲/۲۶۳

ماتریس فواصل (د) جدول ۴

کوتاهترین فاصله	محلات	دلیجان	خمین	قم	تفرش	آشتیان	ساوه	سربند	اراک	
۲/۱۲۶	۴/۸۳۷	۴/۱۳۷	۳/۳۵۶	۲/۱۲۶	۴/۱۱۳	۲/۹۸۹	۳/۶۰۹	۳/۶۵۱	۰	اراک
۵/۹۶۰	۳/۸۹۴	۳/۴۶۶	۰/۹۶۰	۱/۶۶۳	۲/۲۸۱	۱/۲۹۹	۱/۶۲۸	۰	۳/۶۵۱	سربند
۱/۶۲۸	۲/۸۸۸	۳/۳۹۸	۲/۰۴۲	۲/۳۴۹	۲/۵۲۸	۲/۳۲۲	۰	۱/۶۲۸	۳/۶۰۹	ساوه
۱/۱۵۷	۳/۹۹۲	۲/۸۸۹	۱/۱۵۷	۱/۲۷۸	۱/۸۰۳	۰	۲/۳۲۲	۱/۲۹۹	۲/۹۸۹	آشتیان
۱/۸۰۳	۳/۱۵۵	۲/۱۵۰	۲/۳۵۱	۲/۹۰۷	۰	۱/۸۰۳	۲/۵۲۸	۲/۲۸۱	۴/۱۱۳	تفرش
۱/۲۷۸	۴/۴۱۵	۳/۶۶۲	۱/۳۸۴	۰	۲/۹۰۷	۱/۲۷۸	۲/۳۴۹	۱/۶۶۳	۲/۱۲۶	قم
۰/۹۶۰	۴/۴۴۳	۳/۸۳۶	۰	۱/۳۸۴	۲/۳۵۱	۱/۱۵۷	۲/۰۴۲	۰/۹۶۰	۳/۳۵۶	خمین
۲/۱۵۰	۲/۳۲۳	۰	۳/۸۳۶	۳/۶۶۲	۲/۱۵۰	۲/۸۸۹	۳/۳۹۸	۳/۴۶۶	۴/۱۳۷	دلیجان
۲/۳۲۳	۰	۲/۳۲۳	۴/۴۴۳	۴/۴۱۵	۳/۱۵۵	۳/۹۹۲	۲/۸۸۸	۳/۸۹۴	۴/۸۳۷	محلات

$\bar{C}$ =۱/۵۹۸

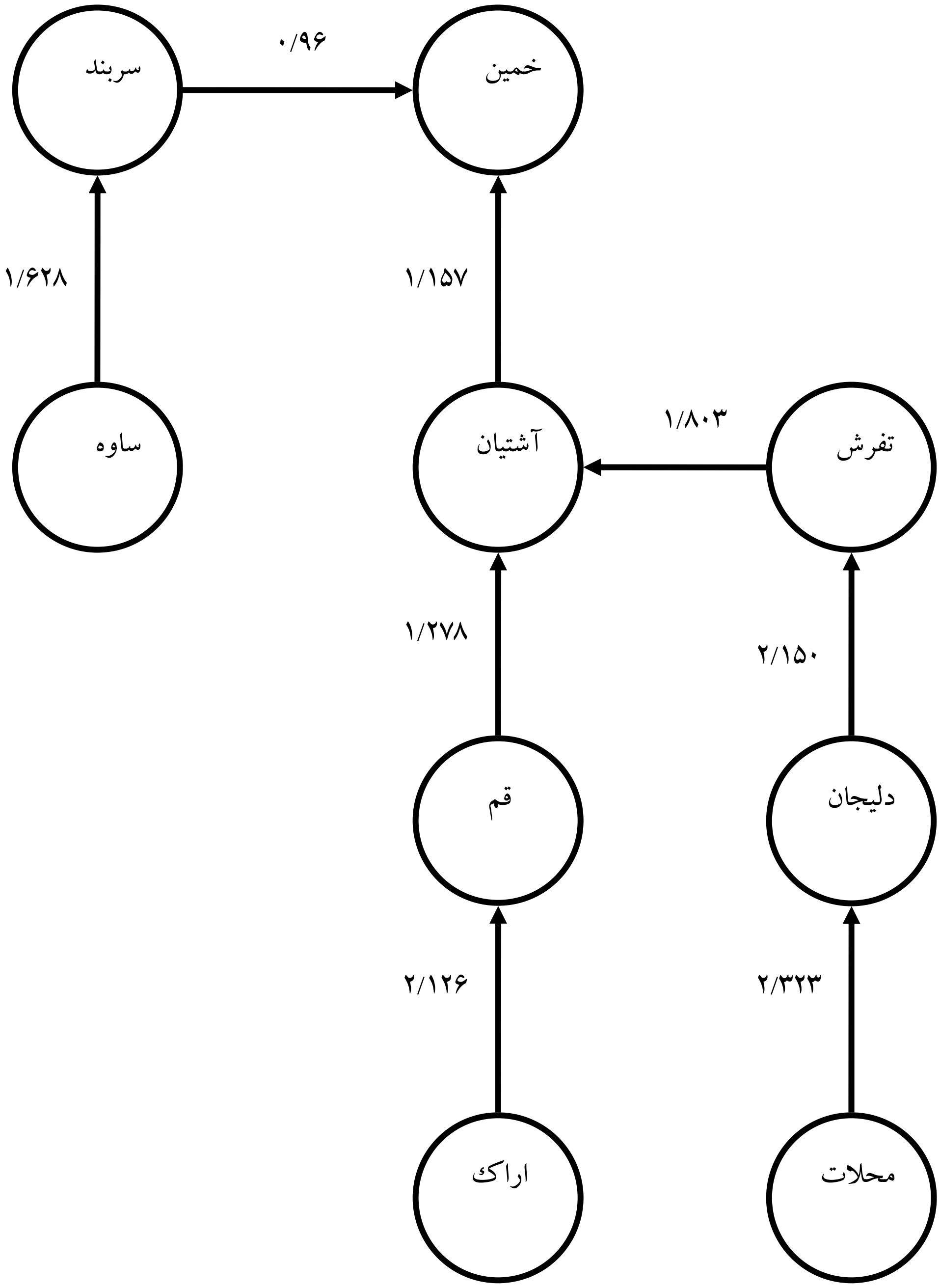
S_c =۰/۵۰۱

$C_{(+)}$ = ۱/۵۹۸ + ۲ x ۰/۵۰۱ = ۲/۶

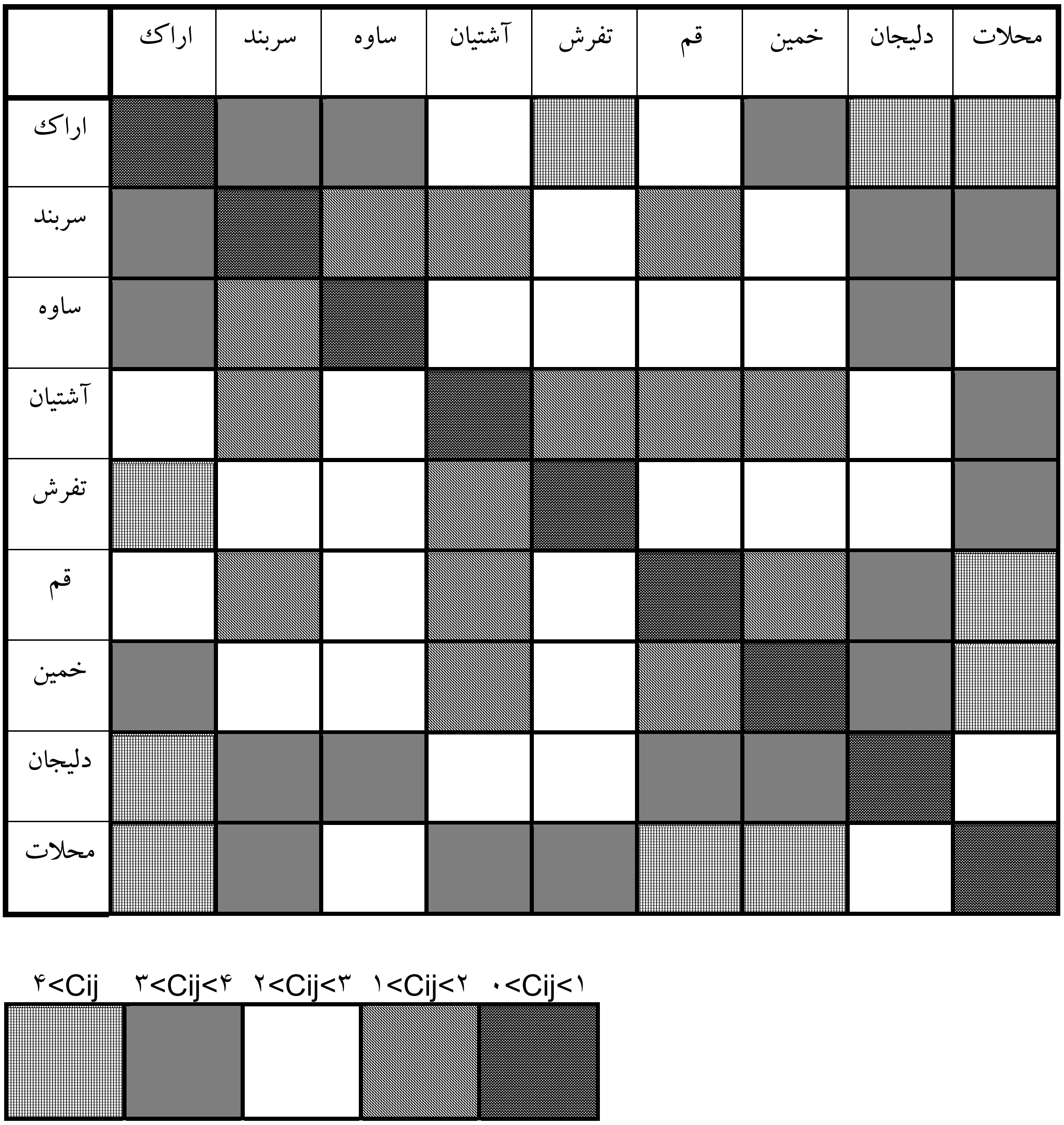
$C_{(-)}$ = ۰/۵۹۶

با توجه به محاسبات بالا تمام شهرستانهای ما درون یک گروه واقع می‌شوند و همه آنها در داخل فواصل بحرانی قرار می‌گیرند و عمل متمرکز نمودن گراف پس از پایان مرتبه اول خاتمه می‌یابد.

گراف پیوسته واحد   نمودار ۱

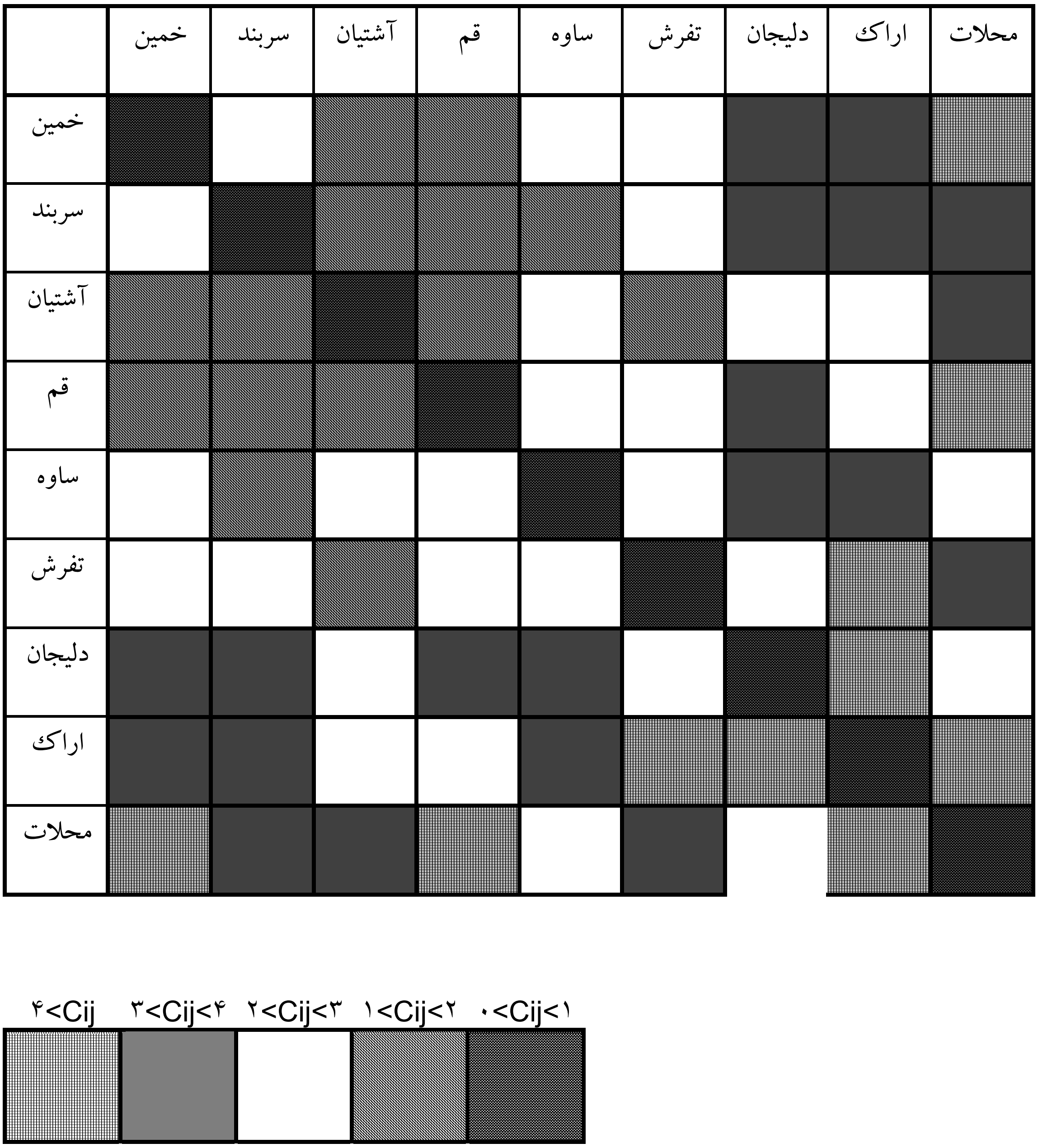


نمودار ۲   نمودار زکانوسکی (نامرتب)



نمودار ۳

نمودار رزکانوسکی (مرتب)



محاسبه d_i ها

جدول ۵

	$(D_{i1}-D_{o1})^2$	$(D_{i2}-D_{o2})^2$	$(D_{i3}-D_{o3})^2$	$(D_{i4}-D_{o4})^2$	C_{io}	$(C_{io}-\bar{C}_{io})^2$	d_i
بزرگترین D_{ij} در ستون jام	۲/۵۲۱	۱/۸۲۰	۱/۳۶۴	۲/۳۳۶	–	–	–
اراک	۰	۱/۱۰۰	۵/۳۱۷	۷/۳۰۶	۳/۷۰۴	۰/۵۲۵	۰/۶۳۸
سربند	۱۰/۵۰۴	۷/۲۶۸	۴/۷۳۰	۵/۷۲۶	۵/۳۱۳	۰/۷۸۱	۰/۹۱۶
ساوه	۷/۱۵۲	۶/۸۷۴	۲/۶۷۳	۰/۹۴۲	۴/۱۹۴	۰/۰۵۵	۰/۷۲۳
آشتیان	۷/۱۰۲	۴/۴۰۱	۲/۶۷۳	۱۰/۴۵۸	۴/۹۶۳	۰/۲۸۵	۰/۸۵۶
تفرش	۱۰/۵۰۴	۴/۴۰۱	۰	۷/۴۸۵	۴/۷۳۱	۰/۰۹۱	۰/۸۱۶
قم	۳/۲۴۳	۴/۵۷۹	۲/۴۷۴	۸/۴۶۸	۴/۳۳۱	۰/۰۰۹	۰/۷۴۷
خمین	۶/۷۲۳	۹/۹۰۳	۳/۹۸۰	۸/۴۶۸	۵/۳۹۲	۰/۹۲۷	۰/۹۳۰
دلیجان	۱۰/۵۰۴	۰	۰	۵/۱۲۱	۳/۹۵۲	۰/۲۲۷	۰/۶۸۱
محلات	۱۰/۵۰۴	۰/۲۷۵	۰	۰	۳/۲۸۳	۱/۳۱۳	۰/۵۶۶
C_{io}, S_{io}	–	–	–	–	۴/۴۲۹ C_{io}	۰/۶۸۴ S_{io}	–

C_0 = ۴/۴۲۹ + ۲ x ۰/۶۸۴

C_0 = ۵/۷۹۷