مدلهای پانل دیتا

تهیه و تنظیم: محمد رضا مبشری

1- مدلهای پانل دیتا[1]

بسیاری از مطالعات اخیر که در زمینه اقتصاد صورت گرفته از مجموعه داده های پانل شده برای بررسی استفاده کرده اند، بدین ترتیب که چندین بنگاه، خانوار، کشور و... در طول زمان مورد تجزیه و تحلیل قرار گرفته است. تجزیه و تحلیل پانل دیتا یکی از موضوعات جدید و کاربردی در اقتصاد سنجی می باشد، چرا که پانل دیتا یک محیط بسیار غنی از اطلاعات را برای گسترش دادن تکنیکهای تخمین و نتایج تئوریک فراهم می آورد. در بسیاری موارد محققین می توانند از پانل دیتا برای مواردی که مسائل را نمی توان فقط به صورت سری زمانی و یا فقط به صورت مقطعی بررسی کرد، استفاده کرده و بهره گیرند. مثلاً در بررسیهای تابع تولید مسئله ای که وجود دارد این است که بتوان تغییرات تکنولوژی را از صرفه های به مقیاس تفکیک کرد. در این گونه موارد داده های مقطعی فقط اطلاعاتی را در مورد صرفه های به مقیاس فراهم می آورد. در حالی که داده های سری زمانی اثرات هر دو را بدون هیچ گونه تفکیکی نشان می دهد. تلفیق آمارهای سری زمانی[2] با آمارهای مقطعی[3] نه تنها می تواند اطلاعات سودمندی را برای تخمین مدلهای اقتصاد سنجی فراهم

1. Models For Panel Data
2. Time series Data
3. Cross- Sectional Data

آورد، بلکه بر مبنای نتایج بدست آمده همچنین می توان استنباطهای سیاستگذاری در خور توجهی به عمل آورد.

1-1 – مزیت استفاده از داده‌های تلفیقی نسبت به سریهای زمانی و مقطعی

1 –داده‌های تلفیقی اطلاعات آگاهی دهنده بیشتر، تنوع یا تغییرپذیری بیشتر، همخطی کمتر بین متغیرها، درجات آزادی و کارایی بیشتر را فراهم می‌کند. در حالیکه سریهای زمانی گرفتار همخطی می‌باشند. در داده‌های تلفیقی با توجه به اینکه ترکیبی از سریهای زمانی و مقطعی می‌باشد، بعد مقطعی موجب اضافه شدن تغییر پذیری یا تنوع بسیار زیادی می شود که با در دست داشتن این اطلاعات می توان برآوردهای معتبرتری انجام داد. مزیت عمده در این داده ها این است که داده های گروهی یعنی داده های مرکب از یک سری زمانی از نمونه‌های مقطعی بالقوه از نظر اطلاعات غنی تر از نمونه مقطعی (N) خواهد بود و اگر صرفا″ از سریهای زمانی استفاده شود تنها به اندازه مشاهدات (T) خواهد بود، اما با تلفیق این دو تعداد داده‌ها به اندازه تعداد مقاطع ضربدر تعداد مشاهدات (N.T) افزایش خواهد یافت که این امر می‌تواند منجر به برآوردهای کاراتری از پارامترها شود.

در محاسبه واریانس جامعه با توجه به مشاهدات مربوط به سری زمانی، واریانس بدست آمده از مشاهدات بر تعداد داده‌ها منهای تعداد پارامترها تقسیم می‌شود.

(رابطه1-1)

$$\hat{\delta}_i^{\,2} = \frac{\delta_i^{\,2}}{N-K}$$

در حالیکه در داده‌های گروهی داریم:

(رابطه1-2)

$$\hat{\delta}_i^{\,2} = \frac{\delta_i^{\,2}}{NT-N-K}$$

که معمولا″ در این حالت مخرج بزرگتر شده و بنابراین واریانس محاسبه شده کوچکتر از واریانس بدست آمده صرفا″ از داده‌های سری زمانی می‌باشد و بنابراین کارایی تخمین افزایش می‌یابد.

به همین قیاس چنانچه آزمون F (آزمون معنی دار بودن کل رگرسیون) را در دو حالت یعنی سری زمانی و تلفیقی مقایسه کنیم داریم:

در مدل تنها سری زمانی:

(رابطه1-3)

$$F = \frac{(\bar{e}'\bar{e} - e'e)/N-1}{e'e/N-K}$$

در صورتیکه مدل تلفیقی F به صورت زیر محاسبه می‌گردد:

(رابطه1-4)

$$F=\frac{(\bar{e}'\bar{e}-e'e)/N-1}{e'e/(NT-N-K)}$$

بوضوح مشخص است که مقدار F در مدل تلفیقی می تواند بزرگتر از مدل تنها سری زمانی باشد و لذا احتمال معنی دار بودن کل رگرسیون یعنی وجود متغیرهای توضیحی در مدل تلفیقی بیشتر خواهد بود.

2 -داده‌های تلفیقی امکان طراحی الگوهای رفتاری پیچیده تری نسبت به داده های مقطعی و سریهای زمانی صرف را فراهم می کند. برای مثال بوسیله داده های ترکیبی امکان بهتری برای بررسی و مدل‌سازی کارایی تکنیکی وجود دارد.

3 -داده‌های تلفیقی امکان بیشتری برای شناسایی و اندازه گیری اثراتی را فراهم می کند که بوسیله فقط آمارهای مقطعی و یا سری زمانی به سادگی قابل شناسایی نیست.

4 -داده‌های تلفیقی از واحدهای کوچکی مثل افراد، شرکتها و خانوارها گردآوری می شوند. خیلی از متغیرها را می توان در مقیاس کوچک با دقت بیشتری اندازه گیری نمود و انحراف‌های ناشی از تجمع افراد یا شرکتها حذف می‌شوند.

امتیاز دیگری که برای تلفیق کردن داده ها می توان در نظر گرفت این است که استفاده از مشاهدات مقطعی ممکن است منجر به برآوردهای اریبی از پارامترها شود. چنانچه از این برشهای مقطعی طی زمان نمونه‌گیری شود و به اصطلاح داده های گروهی فراهم شود براوردهای نااریب و سازگاری امکان پذیر است.

1-2- روشهای تخمین

بر آورد روابطی که در آنها از داده های ترکیبی (سری زمانی، مقطعی) استفاده می شود، غالباً با پیچیدگی هایی مواجه است. در حالت کلی، مدل زیر نشان دهنده یک مدل با داده های ترکیبی است:

(رابطه1-5)

$$Yit=\beta_{1it}+\sum_{k=2}^{k}\beta_{kit}x_{kit+e_{it}}$$

که در آن $i=۱,۲,\ldots,n$ نشان دهنده واحدهای مقطعی (مثلا کشورها) و $t=1,2,\ldots,T$ بر زمان اشاره دارد. Y_{it} متغیر وابسته را برای i امین واحد مقطعی در سال t و x_{kit} نیز k امین متغیر مستقل غیر تصادفی برای i امین واحد مقطعی در سال t ام است.

فرض می شود جمله اخلال e_{it} دارای میانگین صفر، $E[e_{it}]=0$ و واریانس ثابت $E[e_{it}^2]=\sigma^2 e$ است. B_{kit} پارامتر های مدل مجهول است که واکنش متغیر وابسته نسبت به تغییرات kامین متغیر مستقل در i امین مقطع و t امین زمان را اندازه گیری می کند. در حالت کلی فرض می شود که این ضرایب در میان تمامی واحدهای مقطعی و زمانی مختلف متفاوت است. ولی در بسیاری از مطالعات پژوهشی متغیر بودن این ضرایب هم برای تمامی مقاطع و هم برای تمامی زمانهای بسیار محدود کننده است و باید نسبت به ماهیت موضوع مورد مطالعه و سایر شرایط، پژوهشگر خود فرضهای مقتضی را در خصوص پارامترها تعیین کند.

این مدل را می توان به پنج حالت زیر تقسیم کرد:

1 -تمامی ضرایب ثابتند و فرض می شود که جمله اختلال قادر است تمام تفاوت های میان واحدهای مقطعی و زمان را دریافت کند و توضیح دهد.

$$Y_{it}=\beta_1+\sum_{k=2}^{K}\beta_k x_{kit}+e_{it}$$ (رابطه1-6)

2 -ضرایب مربوط به متغیرها (شیب ها) ثابت اند و تنها عرض از مبدا برای واحدهای مختلف مقطعی متفاوت است.

(رابطه1-7)

$$Y_{it}=\beta_{1i}+\sum_{k=2}^{K}\beta_k x_{kit}+e_{it}$$

3 -ضرایب مربوط به متغیرها (شیب ها) ثابت اند ولی عرض از مبدا مابین مقاطع و بین دوره ها متفاوتند..

(رابطه1-8)

$$Y_{it}=\beta_{1it}+\sum_{k=2}^{K}\beta_k x_{kit}+e_{it}$$

4 -همه ضرایب برای تمام واحدهای مقطعی متفاوت است.

(رابطه1-9)

$$Y_{it}=\beta_{1i}+\sum_{k=2}^{K}\beta_{ki} x_{kit}+e_{it}$$

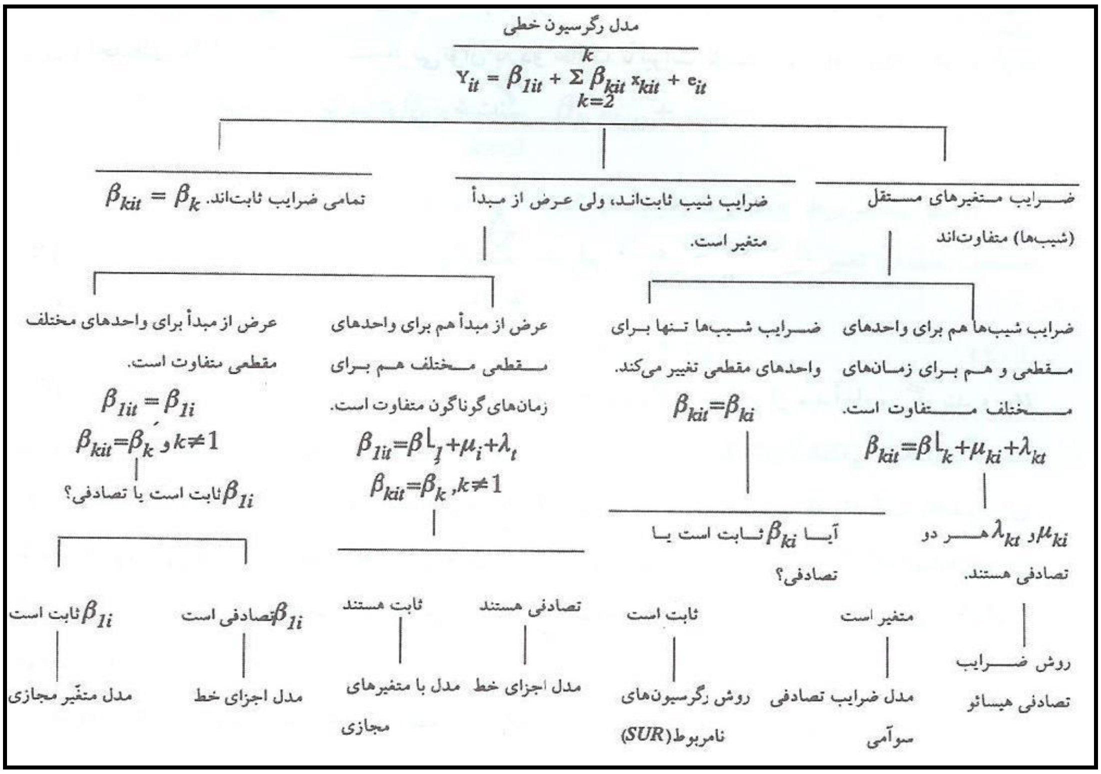
5- تمام ضرایب هم نسبت به زمان و هم نسبت به واحدهای مقطعی متفاوت است.

(رابطه1-10)

$$Y_{it}=\beta_{1it}+\sum_{k=2}^{K}\beta_{kit} x_{kit}+e_{it}$$

در خصوص روشهای تخمین مدلهای مذکور می توان گفت که در حالتهای 2،3و 4 بسته به این که کدامیک از ضرایب ثابت یا متغیر باشند، به مدلهای اثرات ثابت[4] یا تاثیرات تصادفی[5] تقسیم می شوند. که مدلهای مورد بحث را می توان در جدول4-1 خلاصه کرد.

جدول4-1- مدلهای مختلف حاصل از ترکیب داده های سریهای زمانی و مقطعی.



1-3- روش مراحل تخمین مدل بوسیله داده‌های تلفیقی

سوالی که اغلب در مطالعات کاربردی مطرح می شود این است که آیا شواهدی دال بر قابلیت ادغام شدن داده ها وجود دارد یا این که مدل برای تمام واحدهای مقطعی متفاوت است. بعبارت دیگر آیا در مدل مورد نظر برای مقاطع مختلف هم شیبها و هم عرض از مبدأها متفاوت است. این سؤال را می‌توان با فرضیه زیر مطرح نمود:

$$H_0 : \alpha_i = \alpha, \beta_i = \beta$$
$$H_1 : Not \quad H_0$$

1. Fixed Effect
2. Random Effect

فرضیه مذکور را می‌توان به عنوان یک مجموعه قیود خطی روی ضرایب در نظر گرفت و برای آزمون آن که به chow test معروف است از آماره F به صورت ذیل استفاده نمود:

(رابطه1-11)

$$F_{obs}=\frac{(\bar{e}'\bar{e}-e_1'e_1-e_2'e_2-,......,-e_N'e_N)/(N-1)k'}{(e_1'e_1+e_2'e_2+,.........,+e_N'e_N')/N(T-K')}$$

که در آن :

$\bar{e}'\bar{e}$: مجذور پسماندهای حاصل از برازش رگرسیون مقید $y_{it}=\alpha+\beta X_{it}+e_{it}$ است.

$e_i'e_i$:مجذورپسماندهای حاصل از برازش رگرسیون نامقید هر یک از معادلات $y_{it}=\alpha_i+\beta_i X_{it}+e_{it}$ با استفاده از روش حداقل مربعات معمولی می‌باشد.

در صورتیکه فرض H_0 پذیرفته نشود، دلیلی بر یکسان فرض نمودن شیبها و عرض از مبدأ واحدهای مختلف مقطعی وجود ندارد.

آزمون دیگری مطرح است که با فرض متفاوت بودن عرض از مبدأ مقاطع فرضیهٔ زیر را مطرح نمود.

$$H_0:\beta_i=\beta$$
$$H_1: Not \quad H_0$$

که این فرضیه به صورت یک مجموعه قیود خطی فقط روی ضرایب متغیرهای توضیحی (شیبها) در نظر گرفته می‌شود که برای آزمون فرضیهٔ مذکور از آمارهF به صورت ذیل استفاده می‌شود.

(رابطه1-12)

$$F_{obs}=\frac{(\bar{e}'\bar{e}-e_1'e_1-e_2'e_2-,......,-e_N'e_N)/(N-1)k}{(e_1'e_1+e_2'e_2+,.........,+e_N'e_N')/N(T-K)}$$

که در آن:

$\bar{e}'\bar{e}$: مجذور پسماندهای حاصل از برازش رگرسیون مقید $y_{it}=\alpha_i+\beta X_{it}+e_{it}$ است.

$e_N'e_N$: مجذور پسماندههای حاصل از برازش رگرسیون نامقید هر یک از معادلات $y_{it}=\alpha_i+\beta_i X_{it}+e_{it}$ با استفاده از روش حداقل مربعات معمولی (OLS) می‌باشد.

درصورتیکه فرضیه H_0 پذیرفته شود،سؤال اساسی دیگری مطرح خواهد شد و آن این است که آیا تفاوت در مقاطع مختلف می‌تواند بوسیله عرض از مبدأ خاص در واحد پاسخگو باشد. به عبارت دیگر آیا تفاوت در عرض از مبدأ واحدهای مقطعی به طور ثابت عمل می‌کند یا اینکه عملکردهای تصادفی می‌توانند این اختلاف بین واحدها را بطور واضح‌تری بیان نماید که به ترتیب این دو روش در ادبیات داده‌های تلفیقی به روشهای ثابت و اثرات تصادفی مشهور هستند که ذیلا" روشهای فوق‌الذکر به اختصار مورد بحث قرار می‌گیرد.

1-4- اثرات ثابت

یک روش متداول در فرمول بندی کردن مدل داده های تلفیقی، بر این فرض استوار است که اختلافات بین واحدها را می توان به صورت تفاوت عرض از مبدا نشان داد و بنابراین در رابطه (4-5) هر x_i یک پارامتر ناشناخته ای است که باید بر آورد گردد.

به فرض که y_i , x_i شامل T مشاهده برای واحد iام باشند و ε_{it} بردار جزء اخلال بوده و دارای ابعادT.1 بوده باشد در نتیجه رابطه (4-5) را به صورت:

(رابطه1-13)

$$Y_i = I\alpha_i + x_i\beta + \varepsilon_i \qquad i = 1,\dots n$$

$$\begin{bmatrix} Y_1 \\ Y_n \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_n \end{bmatrix} + \begin{bmatrix} X_1 \\ X_n \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_0 \\ \varepsilon_n \end{bmatrix}$$

که در این فرمولها I بردار یکه با ابعاد T.1 می باشد مدل فوق را می توان به شکل خلاصه زیر نوشت.

(رابطه1-14)

$$Y = \left[d_1, \dots d_n\right] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \varepsilon$$

که d_i متغیر مجازی برای نشان دادن iامین مقطع می باشد حال اگر ماتریش D را به صورت $D = \left[d_1 .. d_n\right]$ با ابعاد $T_{n.n}$ تعریف کنیم خواهیم داشت.

(رابطه1-15)

$$Y = D\alpha + X\beta + \varepsilon$$

که این رابطه به عنوان مدل حداقل مربعات متغیر مجازی (LSDV)[1] نامیده می شود. مدل اخیر یک مدل رگرسیونی کلاسیک بوده و هیچ شرط جدیدی برای تجزیه و تحلیل آن لازم نیست و می توان مدل را با استفاده از ارزش OLS با k رگرسور در x وn ستون در D به عنوان یک مدل چند متغیره با n+k پارامتر بر آورد کرد. لازم به ذکر است که می توان در روش اثرات ثابت، عرض از مبدا را طوری بر آورد کرد که نه تنها در مقاطع مختلف بلکه در زمانهای مختلف نیز متفاوت از هم باشند.

1-5- اثرات تصادفی

مدلهای اثرات ثابت تنها درصورتی منطقی خواهد بود که ما اطمینان داشته باشیم که اختلاف بین مقاطع را می توان به صورت انتقال تابع رگرسیون نشان داد، در حالیکه ما همیشه از

1. least square Dummy Variable Model.

وجود این موضوع مطمئن نیستیم. لذا روشهای دیگر مورد استفاده قرار می گیرند. روش دیگر بر آورد، روش اثرات تصادفی است که فرض می کند جزء ثابت مشخص کننده مقاطع مختلف به صورت تصادفی بین واحدها و مناطق توزیع شده است. با توجه به این مورد، مدل با اثرات تصادفی به شکل زیر خواهد بود:

(رابطه16-1)

$$y_{it} = \alpha + \beta X_{it} + u_i + \varepsilon_{it}$$

که دارای K رگرسور به اضافه یک عرض از مبدا می باشد. مولف u_i مشخص کننده جزء تصادفی مربوط به i امین واحد بوده و در طول زمان ثابت است. در مطالعات کاربردی، می توان u_iرا آن دسته از ویژگیهای خاص مربوط به هر مقطع در نظر گرفت که در مدل وارد نشده اند. باید توجه داشت که در این حالت واریانسهای مربوط به مقاطع مختلف با هم یکسان نبوده و مدل ما دچار واریانس ناهمسانی می باشد که باید از روش GLS استفاده نمود. با معرفی این دو روش سوالی که پیش می آید این است که در عمل ما بایستی کدامیک از روشهای مذکور را استفاده کنیم که برای تصمیم گیری از آزمون هاسمن[۷] کمک می گیریم.

1-6- آماره هاسمن

آماره این آزمون که برای تشخیص ثابت یا تصادفی بودن تفاوت های واحدهای مقطعی به صورت زیر محاسبه می شود که دارای توزیع کای- دو با درجه آزادی برابر با تعداد متغیرهای مستقل (K) است.

$$W = X^2(k) = \left[b - \hat{\beta}\right]' \sum{}^{-1} \left[b - \hat{\beta}\right] \qquad \text{(رابطه17-1)}$$

که:

(رابطه18-1)

$$Var\left[b - \hat{\beta}\right] = Var[b] - Var\left[\hat{\beta}\right] = \sum$$

فرضیه صفر بودن آزمون هاسمن، برابری بر آورد کننده هر دو روش حداقل مربعات تعمیم یافته و متغیر مجازی است یعنی داریم:

$$H_0 : \hat{\beta} = b$$

$$H_1 : \hat{\beta} \neq b$$

1. Hausman

چنانچه آماره آزمون محاسبه شده بزرگتر از χ_k^2 جدول باشد فرضیه H_0 رد می‌شود پس برابری برآوردهای این روش رد و توصیه می شود که از روش تصادفی برای دریافت در واحدهای مقطعی استفاده شود، که در شکل1-4 مراحل مختلف آزمونها مشخص گردیده است.

در حالت کلی، همانطور که قبلا″ ذکر شد مدل زیر بصورت ماتریسی نشان دهنده یک مدل با داده‌های ادغام شده‌است.

(رابطه1-19)

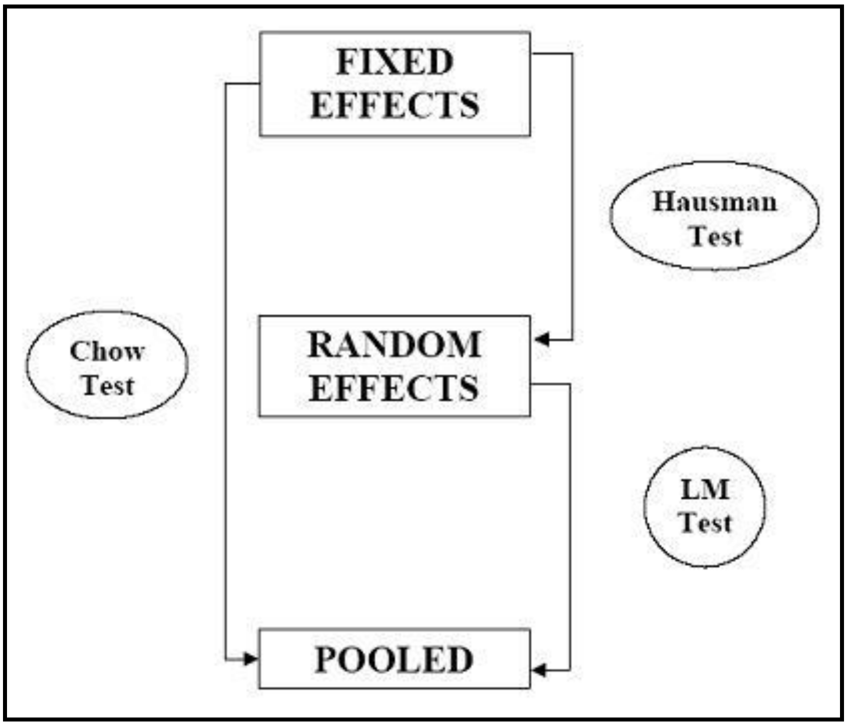
$$Y_{it} = \alpha + X_{it}\beta + W_{it}$$

که در آن Y_{it} متغیر وابسته مربوط به واحد i ام در زمان t تابعی است از X_{it} که مقادیر k متغیر مستقل مربوط به واحد i ام در زمان t را در بر دارد. اما جمله خطا، W_{it} که در ادبیات اقتصادسنجی برآیند تأثیر کلیه متغیرهایی است که بر Y_{it} تأثیر دارند، صریحا″ وارد الگو نگردیده که نقطه افتراق اقتصادسنجی اطلاعات تلفیقی با اقتصادسنجی اطلاعات سری زمانی و یا مقطعی است. در رابطه فوق جمله خطا می‌تواند توسط سه جزء هدایت شود:

الف) μ_i، متغیرهایی که برای واحدهای مقطعی متفاوت، اما در طول زمان ثابت هستند، به عنوان نمونه‌ای از این متغیرها می‌توان به مدیریت در نمونه‌ای متشکل از چند شرکت تولیدی و در یک دوره زمانی معین اشاره کرد. عامل مدیریت بین شرکتها تفاوت دارد، اما برای هر شرکت در طول زمان ثابت است. جنسیت، توانایی و متغیرهای اقتصادی، اجتماعی نیز نمونه‌های دیگر از این متغیرهاست.

ب) λ_{it}، متغیرهایی که برای تمام واحدهای مقطعی در زمان مشابه یکسان هستند اما در طول زمان تغییر می‌کنند. نمونه این نوع متغیرها، می‌تواند قیمت، نرخ بهره و یا انتظارات نسبت به آینده باشد.

ج) ε_{it}، متغیری که نه تنها در طول زمان تغییر می‌کنند بلکه در هر زمان بین واحدهای مقطعی نیز متفاوتند. مثال این متغیرها، حجم سرمایه، فروش و یا سود در واحدهای تولیدی است.



شکل4-1- مراحل مختلف آزمونهای بکاررفته جهت تخمین.

در اقتصاد سنجی اطلاعات سری زمانی و یا مقطعی، جمله خطا تنها از جزء سوم تشکیل شده‌است، یعنی $w_{it} = \varepsilon_{it}$ اما در اقتصادسنجی اطلاعات تلفیقی در حالت اصطلاحاً یکطرفه $w_{it} = \mu_i + \varepsilon_{it}$ و برای حالت دو طرفه $w_{it} = \mu_i + \lambda_t + \varepsilon_{it}$ می‌باشد که عرض از مبدأ برای مدلهای دو طرفه و یکطرفه به صورت ذیل می‌باشد.

(رابطه1-20)

$$\alpha_{it} = \alpha + \mu_i + \lambda_t$$

و برای مدلهای یک طرفه:

(رابطه1-21)

$$\alpha_i = \alpha + \mu_i$$

برای حالت یک طرفه در مدل ما فرض می‌کنیم که عرض از مبدأهای مختلف در مقاطع گوناگون اما در طول زمان یکسان است. گزینش شکل و نوع خطا از بین دو حالت ذکر شده مقوله‌ای تجربه است.

1-7- برآورد یک معادله با استفاده از اطلاعات تلفیقی برای حالت یکطرفه

مدل چند متغیره ذیل را در نظر بگیرید:

(رابطه1-22)

$$y_{it}=\beta_1'+\mu_i+\sum_{k=2}^{k}\beta_k\chi_{kit}+e_{it}$$

به گونه‌ای که $\beta_{1i}=\beta_1'+\mu_i$ است. β'را اصطلاحا" میانگین عرض از مبدأ می‌گویند و μ_i نشان دهندهٔ تفاوتهای موجود در میانگین عرض از مبدأ در بین واحدهای مقطعی مختلف است. روش تخمین معادله به ثابت یا تصادفی بودن μ_i بستگی دارد. در صورتیکه μ_i ثابت باشد، برای تخمین از روش رگرسیون متغیر مجازی یا مدل کواریانس استفاده می شود. حال آنکه تصادفی بودن μ_i منجر به مدل اجزای خطا شده و روش تخمین آن روش حداقل مربعات تعمیم یافته (GLS) خواهد بود.

معادله برای i امین واحد مقطعی در شکل ماتریسی را به صورت زیر می‌توان نوشت:

(رابطه1-23)

$$\bar{y}_{it}=(B_1'+N_i)J+X_{si}B_s+e_i$$

که در آن $y_i=(y_{i1},y_{i2},.....,y_{it})'$ ، $e_i=(e_{i1},e_{i2},.....,e_{it})'$ و $J_T=(1,1,....,1)$ است که همگی آنها دارای ابعاد T.1 هستند. ماتریس X_{si} شامل مشاهداتی از متغیرهای مستقل (بدون عرض از مبدأ) برای i امین واحد مقطعی است. بعد این ماتریس T.K می‌باشد. همچنین فرض می‌شود که میانگین جملات اختلال برای هر واحد مقطعی طی زمان صفر است، یعنی:

$$\mathrm{E}(e_i)=0$$

ماتریس واریانس – کوواریانس هر واحد مقطعی آنها قطری:

(رابطه1-24)

$$\mathrm{E}(e_ie_i')=\delta_e{}^2I_T$$

و در بین واحدهای مختلف صفر است:

$$\mathrm{E}(e_ie_i')=0 \quad ، \quad i\neq j$$

چنانچه معادله1-23 در حالت mi برای تمام واحدهای مقطعی یعنی برای NT مشاهده بنویسم، خواهیم داشت:

$$y=\begin{bmatrix} I_N\otimes J_T & X_S\begin{pmatrix}B_1\\ B_S\end{pmatrix}\end{bmatrix} \qquad \text{(رابطه1-25)}$$

که

$$X'_S = (X'_{S1}, X'_{S2}, \ldots\ldots, X'_{ST})$$
$$e' = (e'_1, e'_2, \ldots\ldots\ldots\ldots\ldots\ldots, e'_T)$$
$$y' = (y'_{1,} y'_2, \ldots\ldots\ldots\ldots\ldots, y'_T)$$
$$B'_1 = (B'_{11}, B'_{12}, \ldots\ldots\ldots, B'_{1N})$$

روش براورد معادله1-23 حداقل مربعات معمولی (OLS) است که با توجه به صفر بودن میانگین بردار جملات اختلال، یک e قطری بودن ماتریس واریانس – کوواریانس، قضیه گووس – مارکوف در مورد این برآوردکننده‌ها صادق است. افزون بر آن، ماتریس $[I_N \otimes J_T \quad X_S]$ ماتریس غیرتصادفی است. فرمول برآوردکننده‌های OLS کوواریانس برای معادله 1-23 به ترتیب به شرح زیر است:

(رابطه1-26)

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} TI_N & (I_N \otimes J_T)' X_S \\ X'_S (I_N \otimes J_T) & X'_S X_S \end{bmatrix} \begin{bmatrix} (I_N \otimes J_T)' Y \\ X'_S \end{bmatrix}$$

(رابطه1-27)

$$\sum_{(b'_1, b'_2)} = \delta_e^{\,2} \begin{bmatrix} TI_N & (I_N \otimes J_T)' X_S \\ X'_S (I_N \otimes J_T) & X'_S X_S \end{bmatrix}^{-1}$$