

انتخاب میان مدل‌های خطی و خطی لگاریتمی با اتکاء به روش‌های اقتصادسنجی

(بررسی موردی منحنی انگل)

محمدرضا منجذب*

برای تخمین مدل‌های اقتصادسنجی همواره محقق با این سؤال روبرو است که کدام یک از مدل‌های خطی و خطی لگاریتمی ترجیحاً بهتر است و باید مورد برآورد و تخمین قرار گیرد. مقاله حاضر با معرفی آزمون MWD به این سؤال مهم پاسخ می‌دهد و در بررسی موردی با معرفی منحنی انگل در مورد برخی از کالاهای اساسی در ایران مدل‌های مرجع را در مورد هر یک انتخاب می‌نماید.

در بررسیهای آماری در قالب مدل‌های اقتصادسنجی، همواره این سؤال برای محقق مطرح است که آیا تخمین مدلها باید به صورت خطی و یا خطی لگاریتمی صورت گیرد؟ بطور معمول محققان در تخمین مدلها به صورت سلیقه‌ای عمل می‌کنند و اگر چنانچه به دنبال تفسیر ضرایب برآورد شده به عنوان مشتق جزئی باشند، مدل را بصورت خطی مورد برآورد قرار می‌دهند. و اگر چنانچه هدف از برآورد مدل، رسیدن به کشش‌ها باشد مدل را به صورت خطی لگاریتمی مورد برآورد قرار می‌دهند.

در مدل‌های تخمینی، کمتر به این نکته توجه شده است که بین مدل‌های خطی و خطی لگاریتمی به استناد تجزیه و تحلیل‌های آماری در قالب اقتصادسنجی کدام مدل برتر و بهتر است و می‌توان از آن در برآورد مدل استفاده شود و این نکته بدون توجه به نظر اولیه محقق باید صورت گیرد.

هدف از نگارش این مقاله معرفی روش و آزمون MWD¹ در گزینش میان دو نوع مدل خطی و خطی لگاریتمی است، که در این راستا با استفاده از داده‌های آماری هزینه و بودجه خانوار شهری و روستایی ۱۳۷۸ مرکز آمار ایران، منحنی انگل برخی از کالاهای خوراکی مورد برآورد قرار می‌گیرد و از این روش در انتخاب مدل بهتر استفاده خواهد شد.

معرفی روش آزمون MWD

انتخاب میان دو مدل رگرسیون خطی (متغیر وابسته تابعی خطی از متغیرهای توضیحی) با یک مدل رگرسیون خطی لگاریتمی (لگاریتم متغیر وابسته تابعی از لگاریتم متغیرهای توضیحی) با استفاده از آزمون مطرح شده توسط مکینون^(۱)، وایت^(۲) و دیویدسون^(۳) که جهت اختصار به آن MWD test می‌گوئیم صورت گرفته

1- MacKinnon

2- White

3- Davidson

و یکی از دو مدل در این میان انتخاب می‌شود.

برای ارائه این آزمون، فرض صفر در مقابل فرض یک؛ به صورت ذیل مطرح

می‌گردد:

H_0 : مدل خطی (Y تابعی خطی از متغیرهای توضیحی X)

H_1 : مدل خطی لگاریتمی ($\ln Y$ تابعی خطی لگاریتمی از متغیرهای توضیحی $\ln X$ ها)

که بطور معمول H_0 و H_1 بیانگر فرض صفر و فرض مقابل هستند. آزمون

MWD شامل مراحل ذیل است:

مرحله اول: تخمین مدل خطی و برازش مقادیر \hat{Y}

مرحله دوم: تخمین مدل خطی لگاریتمی و برازش مقادیر $\hat{\ln Y}$

مرحله سوم: محاسبه $Z_1 = (\ln \hat{Y} - \hat{\ln Y})$

مرحله چهارم: تخمین Y روی Xها و Z_1 . اگر ضریب Z_1 در این مدل بطور

آماري توسط آماره t معنادار باشد در این صورت فرض H_0 رد می‌شود.^(۱)

مرحله پنجم: محاسبه $Z_2 = \text{Antilog } \hat{\ln Y} - \hat{Y}$

مرحله ششم: تخمین $\ln Y$ روی $\ln X$ ها و Z_2 . که اگر ضریب Z_2 در این مدل

معنادار باشد در این صورت فرض H_1 رد می‌شود.^(۲)

معرفی انواع منحنی انگل

در این بررسی، به استناد داده‌های آماری مقطعی منحنی انگل مورد برآورد

قرار خواهد گرفت. لذا انواع منحنی‌های انگل قابل برآورد به شرح ذیل معرفی

می‌گردند:^(۳)

$$\ln Y_i = \alpha_0 + \alpha_1 \ln X_i + \alpha_2 \ln N_i + U_i \quad (1)$$

$$\ln (Y_i/N_i) = \beta_0 + \beta_1 \ln (X_i/N_i) + U_i \quad (2)$$

که در این رابطه Y_i و X_i ، میزان مصرف و درآمد و N_i اندازه خانوار می باشد.
 تفسیر ضریب α_1 مربوط به کشش درآمدی کالای مورد نظر است و نیز α_2 در صورتی که مقدار آن کمتر از یک گردد، بیانگر صرفه جوئی های ناشی از مقیاس خواهد بود. بدین معنا که با افزایش بعد خانوار، میزان مصرف هر نفر بطور متوسط کاهش می یابد.

داده های مدل

اطلاعات استفاده شده در این بررسی، داده های مربوط به اطلاعات هزینه و بودجه خانوار می باشد که آخرین اطلاعات آماری گردآوری شده توسط مرکز آمار ایران و مربوط به سال ۱۳۷۸ می باشد که به صورت مجزا و برای خانوارهای شهری و روستایی گردآوری شده است. به لحاظ اینکه داده های مربوط به درآمد خانوارها به صورت مجزا ارائه نشده، به استناد تئوری های موجود می توان از مجموع هزینه های خوراکی و غیرخوراکی خانوار استفاده نمود.

نتایج مدل نان روستایی

از میان داده های مقطعی موجود به عنوان نمونه منحنی انگل مربوط به انواع نان، گوشت، قند و شکر و نوشابه در مورد دو گروه خانوارهای شهری و روستایی مورد بررسی و برآورد قرار می گیرد، که در این مدلها، نان شهری و روستایی بطور مشروح و سایر مدلها، بطور مختصر گزارش می شود.

ابتدا برای تخمین منحنی انگل نان روستایی و شناسایی خطی یا خطی لگاریتمی مدل (۳) که هزینه انواع نان روستایی، تابعی خطی از بعد خانوار و درآمد خانوار می باشد، مورد برآورد قرار گرفت و نتیجه تخمین به این شرح می باشد.

$$Brdr = 12430 + 0/012 \text{ Incr} + 9737/6 \text{ Size}$$

$\begin{matrix} \text{t:} & 0/62 & 3/3 & 1/6 \end{matrix}$

$$\bar{R}^2 = 0/97 \quad DW = 1/16 \quad F = 131$$

متغیرهای $Brdr$ ، $Incr$ و $Size$ به ترتیب مصرف نان خانوار روستایی، درآمد و بعد یا اندازه خانوار روستایی در سال ۱۳۷۸ می‌باشند.
 آزمون واریانس ناهمسانی نشان می‌دهد که معادله فوق دارای چنین مشکلی نمی‌باشد:

$$\text{Arch Test } F=1/84, Pr=0/32$$

سپس مدل (۳) مورد برازش قرار می‌گیرد و مقادیر برازش شده $Brdr$ را با حرف B نشان می‌دهیم.

پس از آن مدل خطی (۳) را به صورت خطی لگاریتمی و به شکل مدل (۴) مورد برآورد قرار می‌دهیم که نتیجه تخمین به این شرح است:

$$Lbrdr = \frac{3}{6} + \frac{0}{37} Lincr + \frac{0}{28} Lsize \quad (4)$$

$$\bar{R}^2 = 0/99 \quad DW = 1/35 \quad F = 477$$

متغیرهای مدل (۴)، عیناً مربوط به مدل (۳) بوده و فقط حرف L علامت لگاریتم است. لذا مدل فوق یک مدل تمام لگاریتم یا خطی لگاریتمی است. آزمون واریانس ناهمسانی نشان می‌دهد که معادله فوق دارای چنین مشکلی نمی‌باشد.

$$\text{Arch Test } F=1/26 \quad Pr=0/30$$

مقایسه دو مدل (۳) و (۴) از نظر آماری صحیح نمی‌باشد، زیرا متغیرهای وابسته $Brdr$ و $Lbrdr$ در سمت چپ معادلات فوق یک شکل نمی‌باشند و لذا برای مقایسه این دو مدل باید از روش MWD استفاده شود.

بنابراین مدل (۴) مورد برازش قرار گرفته و مقادیر برازش شده $Lbrdr$ را با

LBI نمایش می‌دهیم. سپس مقادیر Z_1 به این صورت مورد محاسبه قرار می‌گیرد:

$$Z_1 = LBH - LB \quad (5)$$

سپس Z_1 در معادله (۳) به عنوان یک متغیر توضیحی وارد شده و مدل (۳) به شکل ذیل، مجدداً مورد برآورد قرار می‌گیرد:

$$Brdr = 600948 + 0/019 Incr - 1/76 Size - 353295 Z_1 \quad (6)$$

$\begin{matrix} t: & & & & \\ & 1/08 & 2/4 & -0/45 & -1/06 \end{matrix}$

$$\bar{R}^2 = 0/97 \quad DW = 1/36 \quad F = 89$$

چنانچه ملاحظه می‌شود ضریب Z_1 با احتمال ۰/۳۳ معنادار نمی‌باشد و لذا فرض H_0 رد نمی‌شود. دلیل این امر آن است که Z_1 تأثیری در مدل (۳) به عنوان متغیر توضیحی ندارد. بنابراین نگارش مدل به صورت خطی تأیید می‌شود.

برای بررسی مدل خطی لگاریتمی ابتدا مقادیر BH را به صورت زیر مورد محاسبه قرار می‌دهیم:

$$BH = \text{Antilog}(LBH) \quad (7)$$

$$Z_2 = BH - B \quad (8)$$

سپس با ورود Z_2 به مدل (۴)، مدل (۹) به این شرح تخمین زده می‌شود:

$$Lbrdr = 4/7 + 0/42 Lincr + 0/53 Lsize - 9/8 \times 10^{-16} Z_2 \quad (9)$$

$$\bar{R}^2 = 0/99 \quad DW = 2/5 \quad F = 382$$

ضریب Z_2 به احتمال ۰/۱۷ معنادار نمی‌باشد و لذا فرض H_0 نیز رد نمی‌شود. دلیل این امر آن است که Z_2 تأثیری در مدل (۴) به عنوان متغیر توضیحی ندارد. بنابراین نگارش مدل به صورت خطی لگاریتمی تأیید می‌شود.

اما با مقایسه دو مدل (۶) و (۹) چون احتمال Z_2 کمتر است و لذا مدل خطی لگاریتمی مدل برگزیده است، بنابراین مدل (۴) ملاک قضاوت منحنی انگل در مورد نان

روستایی ۱۳۷۸ قرار می‌گیرد.

نتایج مدل (۴) بیانگر این است که:

اولاً: کشش درآمدی نان روستایی ۰/۴۹ و کمتر از یک می‌باشد، لذا نان یک کالای ضروری می‌باشد.

ثانیاً: ضریب بعد خانوار معادل ۰/۳۹ می‌باشد و لذا صرفه‌جوئی‌های ناشی از مقیاس در خانوارهای روستایی وجود دارد و با افزایش بعد خانوار؛ میزان مصرف هر فرد بطور متوسط در خانوار کاهش یافته‌است.

نتایج مدل نان شهری

ابتدا برای تخمین منحنی انگل نان شهری و شناسایی خطی یا خطی لگاریتمی مدل (۱۰): که هزینه انواع نان شهری، تابعی خطی از بعد خانوار و درآمد خانوار می‌باشد، مورد برآورد قرار گرفت و نتیجه این تخمین به شرح ذیل می‌باشد:

$$Brdc = ۲۳۷۴۴/۵ - ۰/۰۰۲۷Incc + ۶۵۲۵۷ Size \quad (۱۰)$$

t: ۱/۰۷ -۰/۹۸۷ ۱۵/۸۷

$$\bar{R}^2 = ۰/۹۸ \quad DW = ۱/۵۳ \quad F = ۲۸۶/۸$$

متغیرهای Brdc, Incc و Size به ترتیب مصرف نان خانوار شهری، درآمد،

اندازه یا بعد خانوار شهری در سال ۱۳۷۸ می‌باشند. آزمون واریانس ناهمسانی نشان می‌دهد که معادله فوق دارای چنین مشکلی نمی‌باشد.

Arch Test : F=۱/۲۵

سپس مدل (۱۰) مورد برازش قرار می‌گیرد و مقادیر برازش شده Brdc را با حرف B نشان می‌دهیم.

مدل (۱۰) بصورت خطی لگاریتمی بصورت مدل (۱۱) مورد تخمین قرار

می‌گیرد:

$$LBrdc = \frac{18}{\sqrt{1}} - \frac{0.43}{-2/\sqrt{1}} Lincc + \frac{1}{16/8} Lsize \quad (11)$$

$$\bar{R}^2 = 0.99 \quad DW = 1/3 \quad F = 522$$

L علامت لگاریتم بوده و لذا مدل (۱۱) یک مدل خطی لگاریتمی است. آزمون واریانس ناهمسانی نشان می‌دهد که معادله فوق دارای چنین مشکلی نمی‌باشد.

$$\text{Arch Test : } F = 0.93 \quad Pr = 0.37$$

مقایسه دو مدل (۱۰) و (۱۱) از نظر آماری صحیح نمی‌باشد، زیرا متغیرهای وابسته LBrdc و Brdc در سمت چپ معادلات فوق یک شکل یا با مقادیری یکسان نمی‌باشند و لذا برای مقایسه این دو مدل از روش MWD باید استفاده شود. بنابراین مدل (۱۱) مورد برازش قرار گرفته و مقادیر برازش شده Lbrdc را با LBH نمایش می‌دهیم. سپس مقادیر Z_1 به صورت ذیل مورد محاسبه قرار می‌گیرد:

$$Z_1 = LBH - LB$$

و به همین ترتیب طبق روش قبلی در مورد نان روستایی، آزمون‌ها ادامه می‌یابد و نتایج آنها به این شکل است که ضریب Z_1 به احتمال ۰/۳۳ معنادار نمی‌باشد و لذا H_0 قبول می‌شود. همچنین ضریب Z_2 به احتمال ۰/۱۷ معنادار نمی‌باشد و فرض H_1 نیز پذیرفته می‌شود. بنابراین چون مدل دوم به احتمال کمتری رد نشده است؛ لذا مدل خطی لگاریتمی مدل برگزیده است؛ هر چند که در نوع؛ مدل می‌تواند مورد پذیرش محقق باشد.

تفسیر مدل (۱۱) به عنوان مدل اصلی:

اولاً؛ کشش درآمدی نان شهری ۰/۴۳- و کمتر از یک می‌باشد. ثانیاً؛ ضریب بعد خانوار ۱/۰۲ بزرگتر از یک است و بنابراین صرفه‌جویی‌های ناشی از مقیاس در خانوارهای شهری وجود نداشته است.

در مورد قند و شکر (روستایی) فرض H_0 قبول شد ولی فرض H_1 تأیید نگردید و لذا نگارش مدل به صورت خطی مورد تأیید قرار گرفت و مدل به صورت ذیل برآورد گردید:

$$Sugrr = -2949/9 + 0/007 Incr + 32888 Size \quad (12)$$

t: -0/12 1/55 4/31

$$\bar{R}^2 = 0/98 \quad DW = 2/16 \quad F = 194$$

مدل فوق دچار مشکل واریانس ناهمسانی نمی باشد.

در مورد قند و شکر شهری فرض H_0 رد شد ولی فرض H_1 تأیید گردید و لذا نگارش مدل به صورت خطی لگاریتمی، مورد تأیید قرار گرفت و مدل به صورت ذیل برآورد گردید:

$$Lsugrc = 30/5 - 1/2 Lincscs \quad (13)$$

t: 19/8 -11/9

$$\bar{R}^2 = 0/94 \quad DW = 1/34 \quad F = 142$$

مدل فوق دچار مشکل واریانس ناهمسانی نمی باشد.

بطوری که $Sugrr$ ، $Sugrc$ و $Inccs$ به ترتیب هزینه خانوار روستایی و شهری برای قند و شکر و نسبت درآمد به بعد خانوار شهری می باشند و L نیز علامت لگاریتم است.

هر دو کالای فوق دارای کنش درآمدی کوچکتر از یک می باشند. در مورد نوشابه (روستایی) فرض H_0 و H_1 هر دو تأیید می شوند ولی مدل خطی لگاریتمی بطور قویتری تأیید می شود و لذا به این صورت برآورد می گردد:

$$LBeverr = 30/8 - 1/4 Lincrs$$

t: 3/2 -2/1

$$\bar{R}^2 = 0/27 \quad DW = 0/92 \quad F = 4/4$$

مدل فوق دچار مشکل واریانس ناهمسانی نمی‌باشد.

در مورد مصرف نوشابه در خانوار شهری، هر دو فرض H_0 و H_1 تأیید گردید ولی چون فرض H_0 بطور قویتری تأیید گردید، مدل خطی آن مورد برآورد قرار گرفت و گزارش شد:

$$Beverc = -29115/7 + 0/004 Incc + 3223/2 Size \quad (15)$$

t: -2/4 5/1 2/7

$$\bar{R}^2 = 0/94 \quad DW = 1/2 \quad F = 69$$

در مورد مصرف گوشت در خانوار روستایی، فرض H_0 قبول می‌شود و لذا مدل خطی به صورت ذیل برآورد و گزارش می‌شود:

$$Meatr = 20194 + 0/094 Incc - 17195/8 Size \quad (16)$$

t: 0/64 16/3 -1/7

$$\bar{R}^2 = 0/99 \quad DW = 2/18 \quad F = 1229$$

مدل مصرف گوشت در خانوار شهری به دو صورت خطی و خطی لگاریتمی قابل نگارش است؛ زیرا هر دو فرض H_0 و H_1 تأیید گردیده‌است. ولی مدل خطی لگاریتمی بطور قویتری تأیید شده‌است و به این صورت مورد تخمین نیز قرار گرفته‌است:

$$Lmeatc = -2/8 + 0/998 Lincc + 0/17 Lsize \quad (17)$$

t: -2/2 12/8 5/7

$$\bar{R}^2 = 0/99 \quad DW = 1/84 \quad F = 819$$

محاسبه کشش درآمدی در مدل خطی

بعضی از محققان به دنبال برآورد کشش درآمدی کالاها هستند، و لذا دست به گزینش مدل خطی لگاریتمی و انتخاب آن می‌زنند؛ زیرا ضریب لگاریتم درآمد در مدل‌های فوق همگی به معنای کشش درآمدی هستند.

این انتخاب بر مبنای مدل علمی صورت نگرفته و لازم است میان دو نوع مدل خطی و خطی لگاریتمی، انتخاب و آزمون صورت گیرد. به عنوان نمونه در آزمون‌های فوق دو مدل (۱۵) و (۱۶) بیانگر این است که منحنی انگل نوشابه (شهری) و گوشت (روستایی) باید به صورت خطی برآورد گردند و سایر مدل‌ها به صورت خطی لگاریتمی. اکنون سؤال این است که کشش‌های درآمدی مدل‌های خطی فوق چگونه برآورد شوند؟ پاسخ این سؤال بسیار ساده است.

چنانچه می‌دانیم فرمول محاسباتی کشش درآمدی به شکل ذیل است:

$$ED = \frac{\Delta X}{\Delta Inc} \cdot \frac{Inc}{X} \quad (18)$$

به طوری که X میزان مصرف کالای مورد نظر و Inc درآمد خانوار، علامت Δ نیز به معنای تغییر در متغیر مورد نظر است. ضرایب تخمینی دو مدل (۱۵) و (۱۶) یعنی مدل‌های خطی؛ بیانگر مشتق جزئی می‌باشد. به عبارت دیگر ضریب Inc در مدل (۱۵) همان $\frac{\Delta Beverc}{\Delta Incc}$ و ضریب Inc در مدل (۱۶) همان $\frac{\Delta Meatr}{\Delta Incr}$ می‌باشد. بنابراین از فرمول محاسباتی کشش درآمدی (۱۸) اکنون برای دو کالای نوشابه شهری و گوشت روستایی؛ سمت چپ این فرمول بدست آمده‌است و برای محاسبه سمت راست، بر اساس اطلاعات هزینه و درآمد بودجه خانوارهای شهری و روستایی کافی است مقادیر متوسط درآمد (هزینه) و هزینه نوشابه شهری و هزینه گوشت روستایی

مقدار هزینه شهری برای کل هزینه‌های خوراکی و غیرخوراکی بطور متوسط برابر ۲۱،۷۹۳،۶۸۳ ریال، برای خانوار روستایی این هزینه معادل ۱۳،۶۳۷،۲۷۱ ریال، هزینه مصرفی متوسط یک خانوار شهری برای نوشابه ۶۸،۲۱۶ ریال، و یک خانوار روستایی برای گوشت ۱،۲۰۹،۷۵۷ ریال در سال ۱۳۷۸ بوده است و لذا با جای‌گذاری اطلاعات فوق کثشهای درآمدی دو کالای فوق بدست می‌آید:

$$۱/۲۸ = \frac{۲۱۷۹۳۶۸۳}{۶۸۲۱۶} (۰/۰۰۴) = \text{کثش درآمدی نوشابه شهری}$$

لذا کالای فوق یک کالای لوکس تلقی می‌شود و اعمال مالیات بر مصرف نوشابه شهری توجیه اقتصادی دارد. زیرا بیشتر این مالیات را خانوارهای شهری با درآمدهای بالاتر پرداخت می‌کنند.

$$۱/۰۶ = \frac{۱۳۶۳۷۲۷۱}{۱۲۰۹۷۵۷} (۰/۰۹۴) = \text{کثش درآمدی گوشت روستایی}$$

کالای فوق یک کالای نرمال تلقی می‌گردد و کثش درآمدی آن تقریباً معادل واحد است.

نتیجه اینکه؛ اگر چنانچه هدف محقق از برآورد منحنی انگل کالاهای مختلف، برآورد درآمدی آنها است ابتدا می‌توان، با آزمون MWD، مدل‌های خطی و خطی لگاریتمی را در مورد کالاهای مختلف مورد شناسایی قرار داد، سپس ضرایب لگاریتم درآمدی در سالهای لگاریتمی، که همان کثش درآمد خواهد بود، گزارش گردد. اما در مدل‌های خطی می‌توان از محاسبات فوق کثش درآمدی را بدست آورد.

نتیجه

باتوجه و استفاده از روش MWD، می‌توان میان مدل‌های خطی و خطی لگاریتمی دست به انتخاب زد. در این مبحث این روش در مورد کالاهایی نظیر؛ نان، گوشت، نوشابه، قند و شکر در مورد دو خانوار روستایی و شهری برای سال ۱۳۷۸ در قالب منحنی انگل استفاده شده است و تنها در دو مدل نوشابه برای خانوارهای شهری،

و گوشت برای خانوارهای روستایی، مدلهای به صورت خطی مورد تخمین قرار گرفتند و سایر مدلهای نیز به صورت خطی لگاریتمی برآورد گردیدند.

پی نوشتها:

- 1- J.Mackinnon & White & Davidson; "Tests for Model Specification in the Presence of Alternative Hypothesis; Some Further Results", *Journal of Econometrics*, Vol.21, (1983), pp.53-70.
- 2- D.N.Gujarati, "Basic Econometrics", 3th Edition, (Mc.Grow-Hill Book Incl:1995), pp.265-267.
- 3- S.J. Paris and H.S.Houthakker, *The Analysis of Family Budgets*, Cambridge University Press, (London:1971), Chap7.

- ۴ - بررسی هزینه - درآمد بودجه خانوار شهری. مرکز آمار ایران، ۱۳۷۸.
- ۵ - بررسی هزینه - درآمد بودجه خانوار روستایی، مرکز آمار ایران، ۱۳۷۸.