فصلنامه علمی تخصصی مدیریت، حسابداری و اقتصاد ، دوره ۲ ، شماره ۲، تابستان ۱۳۹۷، ص ۳۷-۴۵
ISSN: 2588-3992
http://www.Tajournals.com

رویکرد عملی برآورد عرضه و تقاضا در بازار با استفاده از متغییر های نسبی، ضربی و رگرسیونی و نمونه گیری دو مرحله ای

سید جمال الدین موسوی مشهدی[۱] و وحید رضامیرابی[۲]
۱ دانشجوی دکتری مدیریت بازاریابی
۲ استاد مدیریت بازاریابی دانشگاه آزاد واحد تهران مرکز

چکیده

در برآورد و پیش بینی بازار در بسیاری از مواقع سعی می کنیم تا با استفاده از یک متغییر کمکی، میزان اطلاع فیشر را در خصوص پارامتر مجهول مورد علاقه در جامعه افزایش دهیم تا دقت برآورد آن پارامتر افزایش یابد. فرض اساسی در این حالت اینست که اطلاعات جامعه در خصوص متغییر کمکی(X) در اختیار است. لیکن در مسائل عملی معمولا اینگونه نیست و اطلاعات مورد اشاره، خود نیز مجهول هستند. روش متداول در اینگونه موارد استفاده از نمونه گیری دو مرحله ای می باشد. در این روش با استفاده از یک نمونه اولیه، پارامتر مجهول متغییر کمکی(X) برآورد می شود. سپس با بهره گیری از نمونه ثانویه، برآورد پارامتر مجهول متغییر وابسته (Y) با استفاده از اطلاعات بدست آمده متغییر کمکی(X)، صورت می پذیرد. در این حالت بسته به نوع رابطه همبستگی بین متغییر کمکی و وابسته، از برآوردگرهای رگرسیونی، ضربی و نسبتی استفاده می شود.
آماردانان برای حل این مشکل از تکنیک دیگری نیز استفاده می کنند. گاهی ضمن استفاده از روش نمونه گیری دو مرحله ای، متغییر کمکی دیگری را نیز در حل مسئله دخالت می دهند(که در مقاله ای دیگر بررسی خواهیم نمود). در نمونه اولیه، با استفاده از متغییر کمکی ثانویه (Z)، پارامتر مجهول متغییر کمکی اولیه (X) برآورد می شود. سپس در نمونه ثانویه با بهره گیری از اطلاعات تولید شده، به برآورد پارامتر مجهول متغییر وابسته (Y) اقدام می کنند. در این تکنیک همبستگی بین متغییر کمکی ثانویه (Z) و متغییر کمکی اولیه (X) قوی می باشد. در حالیکه همبستگی بین متغییر وابسته (Y) و متغییر کمکی ثانویه (Z) بسیار ضعیف است.
با کمک این تکنیک برآوردگرهای مختلفی به دنیای مدیریت از جمله بازرگانی و بویژه دربازاریابی و پیش بینی بازار ،پیش بینی رفتار مصرف کننده، سنجش نگرش خریدار، جایگاه محصول و ... معرفی شدند که مدیریت را در پیش بینی عرضه و تقاضا کمک نموده و در تدوین استراتژی های مناسب یاری می دهند. تفاوت این برآوردگرها در روش استفاده از اطلاعات متغییر(های) کمکی در برآورد پارامترهای مجهول متغییر کمکی اولیه (X) و متغییر وابسته (Y) می باشد. در این مقاله با استفاده از نمونه گیری دو مرحله ای، ضمن بررسی شرایط استفاده از هر برآوردگر و ارائه نمونه، به مقایسه این برآوردگرها نیز می پردازیم.

واژگان کلیدی: کارایی متقاطع- بهینگی پاراتو- کنترل وزن- واحد تصمیم گیرنده .

10.21859/account-......

فصلنامه علمی تخصصی مدیریت، حسابداری و اقتصاد ، دوره ۲ ، شماره ۲، تابستان ۱۳۹۷، ص ۳۷-۴۵
ISSN: 2588-3992
http://www.Tajournals.com

مقدمه

همواره تلاش می کنیم که به سود بیشتری دست یابیم و جایگاه ممتازتری را در بازار به خود اختصاص دهیم. تولید کالا و خدمات در زمان مناسب، با حجم مناسب، شیوه توزیع مناسب،قیمت مناسب و ... از مهمترین موضوعات مورد علاقه مدیریت بازاریابی می باشد تا با تدوین به موقع استراتژی و برنامه عملیاتی، خود را در بین رقبا حفظ و به سطح بالاتری و ارتقاء دهد. از این رو، زمان بعنوان مهمترین عامل در تصمیم گیری نقش کلیدی دارد. از سوی دیگر، تصمیم گیری بر پایه اطلاعات درست می تواند نتایج درست را بهمراه داشته باشد. تصمیم گیری از روی اطلاعات ناقص و نادرست نه تنها باعث بروز اشتباه خواهد شد بلکه فرصت ها و زمان را به یک تهدید تبدیل می کند. از این رو پیش بینی عوامل بازار و پیامدهای تصمیمات مدیریت از اهمیت ویژه ای برخوردار می باشد.

استفاده از متغییر های کمکی (که البته فرض بر آن است که اطلاعات آنان در اختیار است) بهترین ابزار برای پیش بینی عوامل بازار می باشد. با استفاده از رابطه بین متغییر کمکی و متغییر وابسته به پیش بینی متغییر وابسته از طریق اطلاعات متغییر کمکی می پردازیم. از این رو، داشتن متغییر کمکی که بیشترین اطلاعات فیشر را نسبت به متغییر وابسته در خود دارد، هدف مهم تلقی می شود. بسته به اینکه رابطه بین متغییرکمکی مثبت یا منفی باشد و این رابطه چقدر قوی و یا ضعیف باشد، از برآوردگر های نسبتی، ضربی و یا رگرسیونی استفاده می کنیم. پارامترهای متغییر کمکی که انتظار داریم مشخص باشد عبارتند از:

میانگین جامعه
$$\mu_X = \overline{X} = \frac{1}{N}\sum_{m=1}^{N} X_m$$

واریانس جامعه
$$S_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2$$

ضریب تغییرات
$$C_X = \frac{S_X}{\overline{X}}$$

برآوردگر نسبتی:

هدف اصلی در استفاده از اطلاعات کمکی، برآورد پارامترهای مربوط به صفت اصلی Y در جامعه به کمک نمونه تصادفی ساده بدون جایگذاری $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ می باشد. اکنون فرض کنید $\overline{X} = \frac{1}{N}\sum_{m=1}^{N} X_m$ و $\overline{Y} = \frac{1}{N}\sum_{m=1}^{N} Y_m$ ، در این صورت پارامتر R بصورت زیر تعریف می شود:

نسبت دو صفت در جامعه
$$R = \frac{\sum_{m=1}^{N} Y_m}{\sum_{m=1}^{N} X_m} = \frac{Y}{X} = \frac{\overline{Y}}{\overline{X}}$$

بطور مثال اگر Y_i مصرف برنج و X_i تعداد افراد خانوارها باشد، آنگاه R متوسط مصرف برنج هر فرد در جامعه می باشد. اکنون برآوردگرهای زیر را تعریف می کنیم:

برآورد نسبت دو صفت در جامعه
$$\hat{R} = \frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i} = \frac{y}{x} = \frac{\overline{y}}{\overline{x}}$$

10.21859/account-......

فصلنامه علمی تخصصی مدیریت، حسابداری و اقتصاد ، دوره ۲ ، شماره ۲، تابستان ۱۳۹۷، ص ۳۷-۴۵
ISSN: 2588-3992
http://www.Tajournals.com

برآورد نسبتی مقدار کل جامعه $y_R = RX = N\bar{y}_R$

برآورد نسبتی میانگین جامعه $\bar{y}_R = \frac{\bar{y}}{\bar{x}}\bar{X} = \hat{R}\bar{X}$

بدیهی است که $\hat{R}$ متغییری تصادفی است که با تغییر نمونه های به حجم n تغییر می کند.اگر مقادیر هر دو صفت مثبت باشد بطور تقریبی می توان گفت که وقتی در نمونه گیری تصادفی ساده، برآوردگر نسبتی بر برآوردگر معمولی برتری دارد که:

$C > \frac{1}{2}$(و یا بطور معادل وقتی که: $\rho > \frac{C_X}{2C_Y}$)

بنابراین گاهی حتی وقتی ارتباط کامل بین دو صفت X و Y برقرار باشد بازهم برآوردگر نسبتی خوب نخواهد بود. با بررسی دقیق تر مشخص می گردد که حتی اگر همبستگی کامل بین دو متغییر برقرار باشد که داشته باشیم : $Y_i = a + bX_i \quad i = 1,\cdots,N$ هنگامی برآوردگر نسبتی نامناسب است که مقدار a بزرگ باشد و این بدین معنی است که خط رگرسیون از مبداء بسیار دور است.

برآوردگر ضربی:

برای حالتی که همبستگی بین X و Y منفی است، مورتی Murthy روش برآورد دیگری را که به روش ضربی مشهور است پیشنهاد نمود. در واقع در این روش میانگین جامعه با استفاده از برآوردگر :

$$\boxed{\bar{y}_P = \frac{\bar{y}}{\bar{X}}\bar{x}}$$

برآورد می شود. چون در برآوردگر نسبتی از نسبت $\frac{\bar{y}}{\bar{x}}$ استفاده می شد آن را برآوردگر نسبتی می نامیدیم، ولی اکنون از حاصلضرب $\bar{y}.\bar{x}$ استفاده می کنیم و لذا آن را برآوردگر ضربی می نامیم.چنانچه حجم نمونه ، مقدار بزرگی باشد اریبی برآوردگر ضربی ناچیز می گردد. بنابراین بطور تقریبی می توان گفت: $MSE(\bar{y}_P) \cong Var(\bar{y}_P)$. اما در مقام مقایسه با برآوردگر رگرسیونی ، اگر مقادیر هر دو صفت مثبت و دارای همبستگی منفی باشند در اینصورت با در نظر گرفتن تقریب فوق ، $Var(\bar{y}_{RG}) \le Var(\bar{y}_P)$ و نامساوی به تساوی تبدیل می شود اگر خط رگرسیون جامعه از مبداء بگذرد. و در مقام مقایسه با نمونه گیری تصادفی ساده، وقتی برآوردگرهای ضربی بر برآوردگرهای معمولی برتری دارد که: $C < -\frac{1}{2}$(و یا بطور معادل وقتی که: $\rho < -\frac{C_X}{2C_Y}$) و همانند برآوردگر نسبتی، اگر همبستگی کامل بین دو صفت X و Y برقرار باشد هنگامی برآوردگر ضربی نامناسب است که مقدار a بزرگ باشد و این بدین معنی است که خط رگرسیون از مبداء بسیار دور است.

برآوردگر رگرسیونی:

اگر خط رگرسیون Y نسبت به X را برای کل افراد جامعه بصورت $Y = a + bX$ فرض نماییم در واقع این خط مکان هندسی نقاطی به طول $X_m = x_i$ و عرض $Y_m = E(y_i | X_m = x_i)$ که در آن $m = 1,\cdots,N$ می باشد. بنابراین می توان نوشت:

$$y_i = a + bx_i + \varepsilon_i \qquad i = 1,\cdots,N$$

توزیع متغییر تصادفی ε ، با فرض بزرگ بودن N، نرمال با میانگین صفر و واریانس σ^2 فرض می شود. در معادله بالا ضرایب a و b مجهولند. که این ضرایب با دو روش حداقل کردن توان دوم خطا و روش ماکسیمم درستنمایی برآورد می شوند. برآورد

10.21859/account-......

فصلنامه علمی تخصصی مدیریت، حسابداری و اقتصاد ، دوره ۲ ، شماره ۲، تابستان ۱۳۹۷، ص ۳۷-۴۵
ISSN: 2588-3992
http://www.Tajournals.com

این ضرایب به روش کمترین مربعات به صورت ($\hat{a}=\bar{y}-\hat{b}\bar{x}\ \ \&\ \ \hat{b}=\frac{s_{xy}}{s_x^2}$)بوده و برآوردگر رگرسیونی میانگین جامعه بصورت

$\bar{y}_{RG}=\bar{y}+\frac{s_{xy}}{s_x^2}(\bar{X}-\bar{x})$ و واریانس $Var(\bar{y}_{RG})\cong\frac{1-f}{n}S_Y^2\left(1-\rho^2\right)$ که $\rho=\frac{S_{XY}}{S_X S_Y}$ می باشد.

رابطه بین واریانس برآوردگر تصادفی ساده و ورگرسیونی ($Var(\bar{y}_{RG})\le Var(\bar{y})$) و نامساوی به تساوی تبدیل می شود هرگاه $\rho=0$. همچنین رابطه بین واریانس برآوردگر نسبتی با رگرسیونی $Var(\bar{y}_{RG})\le Var(\bar{y}_R)$ می باشد و نامساوی به مساوی تبدیل می شود اگر خط رگرسیون جامعه از مبداء بگذرد. و همانگونه که بیان شد، رابطه بین برآوردگر ضربی با برآوردگر رگرسیونی $Var(\bar{y}_{RG})\le Var(\bar{y}_P)$ ،

و نامساوی به تساوی تبدیل می شود اگر خط رگرسیون جامعه از مبداء بگذرد.

واقعیت عملی

در تمامی این برآوردگرها فرض اساسی بر در اختیار بودن اطلاعات سرشماری X می باشد.اما در عمل ممکن است اطلاعات سرشماری متغییر X در اختیار نباشد. در این حالت روش دیگری مورد بررسی قرار گرفته است. استراتژی مورد نظر در این حالت بهره گیری از نمونه گیری دو مرحله ای است که در هر مرحله اساس نمونه گیری بر نمونه گیری تصادفی ساده بنا شده است. در این استراتژی ما سعی می کنیم که به حل مشکل به شیوه های مختلف بپردازیم. دو روش در این استراتژی مورد توجه و ارزیابی می باشد. روش اول استفاده از یک متغییر کمکی و روش دوم استفاده از دو متغییر کمکی است.

در روش اول، ابتدا نمونه ای به حجم n' از جامعه انتخاب می کنیم. با استفاده از اطلاعات این نمونه پارامترهای جامعه X را برآورد می کنیم. سپس نمونه ای دیگر به حجم $n<n'$ از جامعه انتخاب می کنیم و با این نمونه ثانویه، به کمک روش های رگرسیونی، نسبتی و ضربی به برآورد پارامترهای متغییر اصلی (Y) می پردازیم. لیکن در روش دوم، از متغییر دیگری که به آن متغییر کمکی ثانویه خواهیم گفت، استفاده می کنیم. این متغییر دارای همبستگی شدید با متغییر کمکی اولیه می باشد. لیکن همبستگی متغییر کمکی ثانویه با متغییر اصلی بسیار ضعیف است. در این روش، اطلاعات Z (متغییر کمکی دوم) در اختیار است و با استفاده از نمونه اولیه به برآورد پارامترهای X می پردازیم. در نمونه دوم نیز با بدست آوردن برآورد پارامترهای X در مرحله قبل و اطلاعات متغییر Z ، به برآورد پارامترهای متغییر اصلی خواهیم پرداخت.

نمونه گیری دومرحله ای با استفاده از یک متغییر کمکی

همانطور که گفته شد، در این حالت ابتدا از جامعه مورد بررسی، U، نمونه ای تصادفی به حجم n' انتخاب می کنیم. به کمک این نمونه، میانگین متغییر X جامعه را برآورد می کنیم. این نمونه را $s'\subset U$ می نامیم. برآوردگر مورد استفاده با توجه به تعاریف نظریه نمونه گیری و خاصیت ناریبی بصورت $\hat{\mu}_X=\bar{x}'$ می باشد. با برآورد میانگین جامعه متغییر کمکی X، اکنون شرایط برآورد میانگین جامعه متغییر اصلی را فراهم می سازیم. برای این منظور نمونه ای دیگر به حجم n از نمونه اولیه انتخاب می کنیم و آن را $s\subset s'$ می نامیم. اکنون نمونه ثانویه شرایط استفاده از برآوردگرهای نسبتی، رگرسیونی و ضربی را دارد. برای سادگی در عملیات محاسباتی متغییرهای زیر را تعریف می کنیم: $\theta_1=(\frac{1}{n}-\frac{1}{N})\ \ \theta_2=(\frac{1}{n'}-\frac{1}{N})\ \ \theta_3=(\frac{1}{n}-\frac{1}{n'})$

برآوردگر رگرسیون دومرحله ای:

همانگونه که بیان شد، در برآورد رگرسیونی فرض مبتنی بر این است که میانگین متغییر کمکی در جامعه یعنی $\bar{X}_N$ معلوم است. اما می دانیم در بسیاری از کاربردها، ممکن است که این میانگین مشخص نباشد. اگر چنین نباشد یعنی $\bar{X}_N$ مشخص نباشد، در این حالت ابتدا یک نمونه تصادفی به حجم n' می گیریم و بر اساس این نمونه استخراج شده $\bar{X}_N$ را برآورد می کنیم که

10.21859/account-......

فصلنامه علمی تخصصی مدیریت، حسابداری و اقتصاد ، دوره ۲ ، شماره ۲، تابستان ۱۳۹۷، ص ۳۷-۴۵
ISSN: 2588-3992
http://www.Tajournals.com

$\hat{\mu}_X = \bar{x}'$ برآوردگر آن می باشد. سپس نمونه ای دیگر به حجم n از نمونه اول استخراج می کنیم و مقادیر متناظر صفت کمکی را مشخص می کنیم. لازم به تذکر است که در این روش برای کاهش هزینه ها نمونه به حجم n را از نمونه اول به حجم n' انتخاب می کنیم.بنابراین برآوردگر رگرسیونی در این حالت:

$$t_1 = \bar{y} + b_{yx}(\bar{x}' - \bar{x})$$

می باشد که یک برآوردگر ناریب است.

مثال: می خواهیم میانگین میزان تولید محصول چغندر قند را برآورد کنیم. متغییر کمکی مورد استفاده، مساحت زمین زیر کشت این محصول است. اما مساحت کل زمین های زیر کشت در اختیار نمی باشد. بنابراین در این پژوهش، با استفاده از نمونه گیری دو مرحله ای، مساحت زیر کشت را برآورد می کنیم و سپس میزان محصول را برآورد می کنیم. بنابراین داریم:

Y میزان تولید محصول چغندر قند

X مساحت زیر کشت چغندر قند

داده های اولیه میزان تولید چغندر قند در سال ۱۳۸۳(نمونه اولیه – ۷۷ مشاهده)

میزان تولید	مساحت زیر	مساحت مزرعه	میزان تولید	مساحت زیر	مساحت مزرعه	میزان تولید	مساحت زیر	مساحت مزرعه
263	18	50	90	2	8	90	3	7
154	7	100	90	3	6	90	2	8
20	1	20	95	3	8	90	3	6
200	5	30	100	3	8	95	3	8
50	1	4	50	2	5	100	3	8
350	9	30	70	1	5	50	2	5
400	10	16	50	2	5	70	1	5
70	1	5	50	2	5	50	2	5
90	3	8	52	2	5	50	2	5
50	2	5	70	1	5	52	2	5
70	1	5	90	3	8	120	4	10
120	4	10	50	2	5	90	3	7
90	3	7	70	1	5	90	2	8
90	2	8	263	18	50	90	3	6
90	3	6	154	7	100	95	3	8
95	3	8	20	1	20	100	3	8
100	3	8	200	5	30	50	2	5
50	2	5	350	9	30	70	1	5
70	1	5	400	10	16	50	2	5
50	2	5	70	1	5	50	2	5
50	2	5	90	3	8	52	2	5
52	2	5	50	2	5	70	1	5
120	4	10	70	1	5	90	3	8
90	3	7	120	4	10	50	2	5
50	2	5	50	2	5			
200	5	30	120	4	10			

لگاریتم داده های میزان تولید چغندر قند در سال ۱۳۸۳ (نمونه اولیه – ۷۷ مشاهده)

10.21859/account-......

فصلنامه علمی تخصصی مدیریت، حسابداری و اقتصاد ، دوره ۲ ، شماره ۲، تابستان ۱۳۹۷، ص ۳۷-۴۵
ISSN: 2588-3992
http://www.Tajournals.com

میزان تولید	مساحت زیر کشت	مساحت مزرعه	میزان تولید	مساحت زیر کشت	مساحت مزرعه	میزان تولید	مساحت زیر کشت	مساحت مزرعه
5.57	2.89	3.91	4.61	1.10	2.08	3.91	.69	1.61
5.04	1.95	4.61	3.91	.69	1.61	4.25	.00	1.61
3.00	.00	3.00	4.25	.00	1.61	3.91	.69	1.61
5.30	1.61	3.40	3.91	.69	1.61	3.91	.69	1.61
3.91	.00	1.39	3.91	.69	1.61	3.95	.69	1.61
5.86	2.20	3.40	3.95	.69	1.61	4.79	1.39	2.30
5.99	2.30	2.77	4.25	.00	1.61	4.50	1.10	1.95
4.25	.00	1.61	4.50	1.10	2.08	4.50	.69	2.08
4.50	1.10	2.08	3.91	.69	1.61	4.50	1.10	1.79
3.91	.69	1.61	4.25	.00	1.61	4.55	1.10	2.08
4.25	.00	1.61	5.57	2.89	3.91	4.61	1.10	2.08
4.79	1.39	2.30	5.04	1.95	4.61	3.91	.69	1.61
4.50	1.10	1.95	3.00	.00	3.00	4.25	.00	1.61
4.50	.69	2.08	5.30	1.61	3.40	3.91	.69	1.61
4.50	1.10	1.79	5.86	2.20	3.40	3.91	.69	1.61
4.55	1.10	2.08	5.99	2.30	2.77	3.95	.69	1.61
4.61	1.10	2.08	4.25	.00	1.61	4.25	.00	1.61
3.91	.69	1.61	4.50	1.10	2.08	4.50	1.10	2.08
4.25	.00	1.61	3.91	.69	1.61	3.91	.69	1.61
3.91	.69	1.61	4.25	.00	1.61	4.25	.00	1.61
3.91	.69	1.61	4.79	1.39	2.30	4.79	1.39	2.30
3.95	.69	1.61	4.50	1.10	1.95	5.30	1.61	3.40
4.79	1.39	2.30	4.50	.69	2.08	3.91	.69	1.61
4.50	1.10	1.95	4.50	1.10	1.79	3.91	.69	1.61
4.50	.69	2.08	4.55	1.10	2.08			
4.50	1.10	1.79	4.61	1.10	2.08			

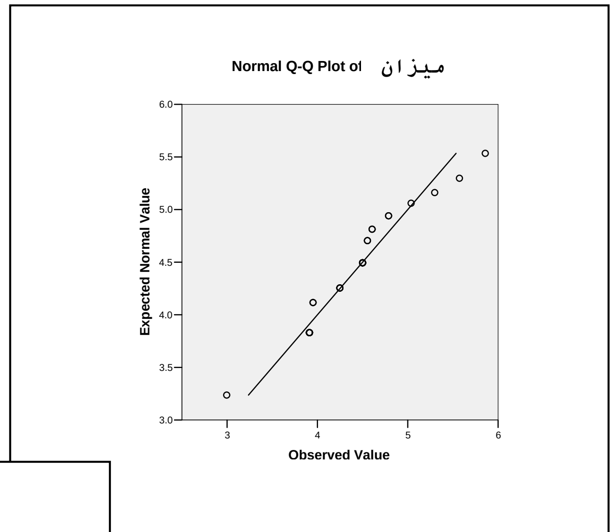
بعد از بررسی داده ها و تست Q-Qplot (نمودار صفحه بعد)، از داده های تغییر یافته که لگاریتم داده های اولیه می باشند، استفاده می کنیم. بنابراین با بهره گیری از برآوردگر رگرسیونی دو مرحله ای داریم:

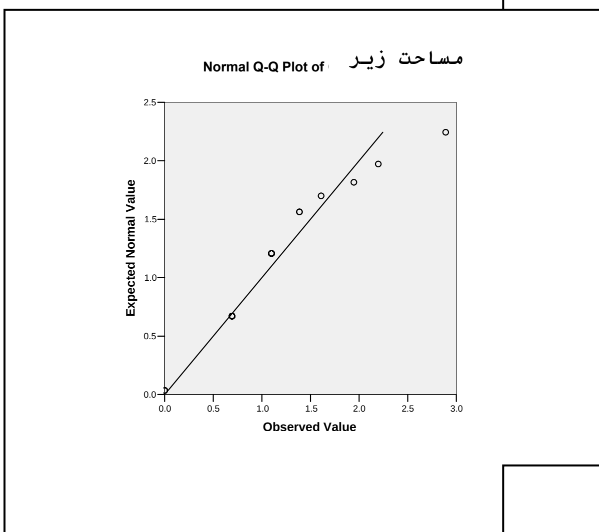
$\bar{y}$	S_y^2	$\bar{x}'$	$\bar{x}$	S_x^2	C_y	C_x	θ_1	θ_2	θ_3
4.3858	0.309	0.9146	0.9257	.407	0.126	0.691	0.024	0.006	0.018
b_{yx}	ρ_{yx}								
0.707	0.811								

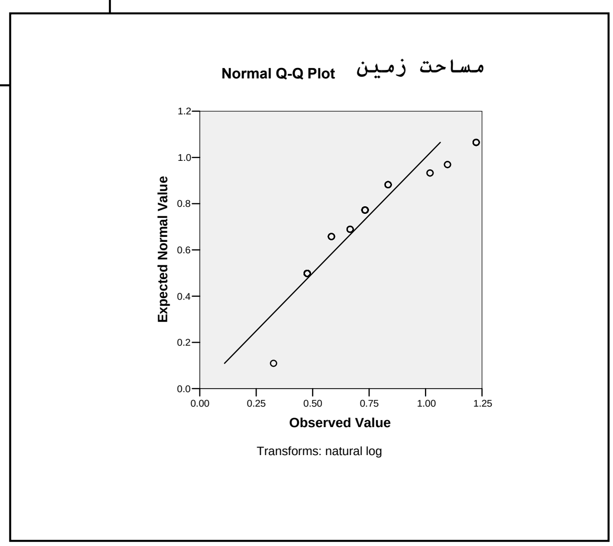
لازم به توضیح است که حجم جامعه ۱۴۴، نمونه اولیه ۷۷ و نمونه ثانویه ۳۲ در نظر گرفته شده است. بنابراین می توان به سادگی مقدار عددی t_1 را بدست آورد. لذا داریم:

10.21859/account-......

ISSN: 2588-3992
http://www.Tajournals.com







$$t_1 = \bar{y} + b_{yx}(\bar{x}' - \bar{x}) = 4.3858 + 0.707\,(0.9146 - 0.9257)$$

$$\Rightarrow t_1 = 4.3779$$

10.21859/account-......

ISSN: 2588-3992
http://www.Tajournals.com

اکنون میانگین مربع خطا برای t_1 به شرح زیر می باشد:

$$\hat{V}(t_1)=(4.3779)^2(0.126)^2\left[0.024(1-0.811^2)+0.006\times 0.811^2\right]=0.0037$$

کارایی این برآوردگر نسبت به تصادفی ساده بصورت زیر است:

$$\frac{V(\bar{y})}{MSE(t_1)}\times 100=\frac{0.0075}{0.0037}\times 100=202.7$$

در این مثال ، بجای بهره از برآوردگر رگرسیونی از برآوردگر نسبتی استفاده می کنیم. لیکن لازم به یادآوری است که، هنگامی برآوردگر نسبتی نامناسب است که مقدار a بزرگ باشد و این بدین معنی است که خط رگرسیون از مبداء بسیار دور است. به بیان دیگر هنگامی برآوردگر نسبتی نامناسب است که خط رگرسیون از مبداء مختصات دور باشد. لیکن برآوردگر رگرسیونی در صورت وجود ارتباط خطی شدید همواره مناسب است و به فاصله خط رگرسیون از مبداء بستگی ندارند. اکنون با توجه به آنچه در بخش قبل بیان شد داریم:

$\bar{y}$	S_y^2	$\bar{x}'$	$\bar{x}$	S_x^2	C_y	C_x	θ_1	θ_2	θ_3
4.3858	0.309	0.9146	0.9257	.407	0.126	0.691	0.024	0.006	0.018
b_{yx}	ρ_{yx}								
0.707	0.811								

بنابراین داریم:

$$t_2=(4.3858)\frac{0.9146}{0.9257}=4.3332$$

و میانگین مربع خطای آن بصورت:

$$MSE(t_2)=(4.3332)^2\left[0.006\times 0.126^2+0.018\left(0.126^2-2\times 0.811\times 0.691\times 0.126+0.691^2\right)\right]$$

$$\Rightarrow MSE(t_2)=0.120$$

کارایی این برآوردگر نسبت به تصادفی ساده بصورت زیر است:

$$\frac{V(\bar{y})}{MSE(t_2)}\times 100=\frac{0.0075}{0.1208}\times 100=6.208$$

لازم به توضیح است که به دلیل بزرگ بودن فاصله خط رگرسیون از مبداء مختصات(بزرگ بودن a)، این برآوردگر کارایی چندانی از خود نسبت به برآوردگر تصادفی ساده نشان نمی دهد.

10.21859/account-......

فصلنامه علمی تخصصی مدیریت، حسابداری و اقتصاد ، دوره ۲ ، شماره ۲، تابستان ۱۳۹۷، ص ۳۷-۴۵
ISSN: 2588-3992
http://www.Tajournals.com

مراجع:

[1] Dulal Chandra Roy (2003). A Regression-type estimator in two-phase sampling using two auxiliary variables. Pak J. Statist. Vol 19(3). PP. 281-290.

[2] Singh, P. and Ruiz Espejo, M. (2003). On Linear regression and ratio – Product estimation of a finite population mean. Royal statistical Society, 52, part 1, PP.59-67

[3]دکتر عمیدی : نظریه نمونه گیری و کاربردهای آن

10.21859/account-......