● مجلهٔ دانشگاه علوم پزشکی بابل، دوره هشتم، شمارهٔ ۳، خرداد - تیر۱۳۸۵، صفحه ۴۳-۳۶
دریافت: ۸۳/۹/۱۷، پذیرش: ۸۵/۳/۱۰

مدل سازی مدت اقامت بیماران بستری شده در بیمارستان براساس توزیع پوآسن آمیخته

عبدالرضا رجائی فرد[۱*]، محمد رفیعی[۲]
۱- دانشیار گروه آمار زیستی دانشکده بهداشت، دانشگاه علوم پزشکی شیراز ۲- استادیار گروه بهداشت و پزشکی اجتماعی دانشگاه علوم پزشکی اراک

سابقه و هدف: مدل بندی یکی از اساسی ترین روشهای تبیین متغیرهای آماری است که با استفاده از آن می توان به چگونگی توزیع متغیر پاسخ مورد نظر پی برد. جهت تحلیل داده هائی مانند مدت اقامت بیماران بستری شده در بیمارستان، شرایط نرمال بودن داده ها و همگنی واریانس های خطا برقرار نیست، بنابراین یا باید از روشهای ناپارامتریک و یا تصحیح کننده هائی همچون تغییر متغیر لگاریتمی متغیر پاسخ برای بدست آوردن خصوصیات پارامتریک استفاده کرد. تجربه نشان داده است در این مواقع توزیع های آمیخته تا حدود زیادی می تواند نیکوئی برازش مدل را بهبود بخشد. لذا هدف از این مطالعه معرفی مدل آمیخته پوآسن و بکار بردن مدل های رگرسیونی پوآسن آمیخته برای تبیین مقادیر مدت زمان اقامت بیماران در بیمارستان و بدست آوردن عوامل مؤثر بر این مدت زمان می باشد، همچنین مقایسه این مدل ها با مدل های رایج رگرسیونی در این گونه داده ها می باشد.

مواد و روشها: در این مطالعه پس از معرفی مدل های آمیخته پوآسنی و رگرسیون حاصل از آنها، این مدل ها جهت داده های مدت اقامت بیماران در دو بخش داخلی و جراحی بیمارستان ولیعصر شهر اراک بکار برده شده است. متغیرهائی همانند سن, وضعیت تأهل، محل تولد و محل زندگی به عنوان متغیرهای مستقل و متغیر مدت زمان اقامت بیماران در بیمارستان به عنوان متغیر پاسخ شمارشی برای کاربرد مدل در نظر گرفته شدند.

یافته ها: نتایج مطالعه نشان داد با توجه به مقدار آمارهLog-likelihood حاصل از مدل آمیخته پوآسنی و پراکندگی زیاد متغیر پاسخ، در بخش جراحی مدل آمیخته پوآسنی مدل مناسبی جهت تبیین مقادیر مدت اقامت بر اساس متغیرهای دیگر می باشد و در بخش داخلی متغیر پاسخ پراکندگی زیاد ندارد، بنابر این مدل پوآسن آمیخته به خوبی نمی تواند مقادیر مدت اقامت بیماران در این بخش را تبیین کند.

نتیجه گیری: با توجه به ناهمگنی مقادیر مدت اقامت بیماران در بخش جراحی و مقدار آمارهLog-likelihood حاصل از بکار بردن مدل, مدل آمیخته پوآسنی یک مدل مناسب جهت تبیین مقادیر مدت اقامت بیماران می باشد، در صورتیکه از مدل های معمولی استفاده شود مقدار آماره Log-likelihood بیشتر خواهد شد و عوامل معنادار کمتری در مدل موجود خواهند بود. کاربرد این مدل ها در مواردی که متغیر پاسخ شمارشی پراکندگی زیاد دارد, پیشنهاد می گردد.

واژه های کلیدی: *رگرسیون پوآسن آمیخته، مدت اقامت بیماران در بیمارستان، رگرسیون پوآسنی، رگرسیون خطی ساده.*

مقدمه

در مواقعی که داده های مربوط به متغیر پاسخ, متغیر شمارشی باشد برای بیان توزیع آن مناسبتر است که از توزیع هائی مانند توزیع پوآسن یا دوجمله ای منفی استفاده شود(۳و۲و۱). جهت داده های شمارشی اگر مقادیر میانگین و واریانس حدوداً نزدیک به

Archive of SID
www.SID.ir

هم باشند، توزیع پوآسن و اگر مقدار واریانس از مقدار میانگین بیشتر باشد (پراکندگی زیاد متغیر پاسخ) مدل توزیع دوجمله ای منفی جهت داده های متغیر پاسخ مناسب خواهد بود(۷و۶و۵و۴). این توزیع ها بنا به خاصیتهای ویژه خود برای داده هائی که دارای توزیع نرمال نمی باشند و دارای چولگی به راست یا چپ هستند, مورد استفاده قرار می گیرند(۱۰و۹و۸). پس از در نظر گرفتن توزیع پوآسن یا دوجمله ای منفی جهت متغیر پاسخ برای تبیین متغیر پاسخ بر اساس متغیر های مستقل دیگر از رگرسیون پوآسنی یا دوجمله ای منفی استفاده خواهد شد. بحثی که در سالهای اخیر مطرح شده است، این است که در خیلی از مواقع در توزیع های پوآسن, توزیع با مؤلفه هائی بیشتر از یک مؤلفه در نظر گرفته می شود و مسئله توزیع های آمیخته مطرح می گردد. در این توزیع ها نیازی به اینکه توزیع باید چند نما داشته باشد نمی باشد، بلکه یک توزیع چوله به راست با یک نما نیز می تواند بعنوان یک توزیع آمیخته با مؤلفه هائی بیشتر از یک مؤلفه مطرح گردد. همچنین توزیع آمیخته پوآسن یک توزیع مناسب برای داده های شمارشی چوله می باشدکه با استفاده از آزمونهای آماری دارای توزیع پوآسن معمولی نمی باشند (۳). توزیع های آمیخته پوآسنی یک حالت تعمیم داده شده از توزیع های پوآسن یک مؤلفه ای می باشند که در خیلی از مواقع از توزیع پوآسن مناسب تر خواهند بود. توزیع های آمیخته در حالت های عملی و کاربردی بسیار مطرح و کاربرد دارند و بیشتر در خصوص داده هائی که پراکندگی زیادی دارند مورد استفاده قرار می گیرند (۱۱). از طرفی در مواردی که داده ها دارای مشاهدات دور افتاده باشند، توزیع های آمیخته بعنوان ابزاری مناسب برای تبیین داده ها مطرح می شوند، در این خصوص دو دیدگاه وجود دارد، دیدگاه اول به کل توزیع بصورت دو توزیع, یکی جهت دادهای اصلی غیر دور افتاده و دیگری برای داده های دور افتاده، می پردازد در دیدگاه دوم با استفاده از تصحیح کننده های آماری مقادیر دور افتاده حذف گشته تا توزیع داده ها متعادل و همگن شود. البته باید در نظر داشت که در مواردی که هدف بررسی ارتباط بین متغیر وابسته و دیگر متغیرهای مستقل می باشد، حذف مقادیر دور افتاده می تواند نتایج گمراه کننده ای را ارائه دهد ولی بکار بردن توزیع آمیخته می تواند این ارتباط را بخوبی نشان دهد(۱۲). برای مدل بندی مقادیر مدت اقامت بیماران، مطالعه ای در ایران که بیانگر کاربرد مدل های پیشرفته و رایج جهت تبیین آنها باشد، تابحال صورت نگرفته است و این مدل بندی می تواند جهت سیستم های بهداشتی بسیار مؤثر باشد. طول مدت بستری شدن در بیمارستان اغلب به عنوان معیار بازدهی و کارائی خدمات بیمارستانی مورد استفاده قرار می گیرد و به عنوان نماینده و معرف معقول مصرف ذخائر بهداشتی تلقی می گردد و از آن برای مقاصد مختلفی مانند مدیریت خدمات بهداشتی, کنترل کیفیت خدمات بیمارستانی، مفید بودن خدمات بیمارستانی و طرح ریزی نیازمندیهای بیمارستانی مورد استفاده قرار می گیرد (۱۸-۱۳). هدف از بررسی حاضر بیان مختصری از پوآسن آمیخته و به کارگیری این مدل ها جهت تبیین مدت اقامت بیماران در دو بخش داخلی و جراحی بیمارستان ولیعصر اراک می باشد، همچنین مقایسه ای بین مناسبت این مدل ها با مدل های رایج و معمولی (رگرسیون خطی) در مورد تبیین مقادیر مدت اقامت بیماران بر اساس برخی عوامل دموگرافیک صورت گرفته است.

مواد و روشها

توزیع پوآسن آمیخته:

ساختار یک توزیع آمیخته به شرح زیر می باشد:

$$f(y)=\alpha_1 g_1(y;\lambda_1)+\alpha_2 g_2(y,\lambda_2)+\ldots+\alpha_c g_c(y;\lambda_c);\sum_{j=1}^{c}\alpha_j=1$$

که α_j بیانگر نسبت j امین مولفه توزیع و c تعداد مؤلفه ها و $g_j(y;\lambda_j)$ تابع توزیع احتمال j امین مؤلفه با پارامتر λ_j می باشد(۱۲). تحلیل این مدل به این صورت است که توزیع تجربی داده ها با در نظر گرفتن هر مؤلفه بعنوان وزن, به چند مؤلفه شکسته می شود. پارامترهای مدل (α,λ) می باشند که با استفاده از روشهائی مانند برآورد ماکزیمم درستنمائی یا الگوریتم EM برآورد می گردند. بیان کامل این مدل ها در مقالات Everitte و Hand موجود می باشد(۱۱). جهت بدست آوردن مؤلفه های یک توزیع آمیخته باید $g_j(y;\lambda_j)$ را معین کرد. آنالیز توزیع آمیخته تعیین

Archive of SID

www.SID.ir

می کند که آیا می توان داده های متغیر پاسخ را بعنوان یک توزیع معین با چند مؤلفه بیان نمود یا خیر؟ در صورتیکه $g_j(y;\lambda_j)$ بعنوان توزیع پوآسن در نظر گرفته شود, مدل حاصل یک مدل آمیخته پوآسنی خواهد بود و بصورت زیر نشان داده می شود:

$$;\sum_{j=1}^{c}\alpha_j=1$$

$$f(y)=\alpha_1\frac{e^{-\lambda 1}\lambda 1^y}{y!}+\alpha_2\frac{e^{-\lambda 2}\lambda 2^y}{y!}+...+\alpha_c\frac{e^{-\lambda c}\lambda c^y}{y!};y=0,1,2,3,...$$

ضرائب λ,α ضرائب مدل می باشند که بایستی برآورد شوند. در این مطالعه این ضرائب با استفاده از نرم افزار C.A.MAN برآورد شده اند. برای آشنائی با چگونگی شکل توزیع پوآسن دو مؤلفه ای، اشکال این توزیع با چند پارامتر مختلف در شکل ۱ نشان داده شده است.

در کاربرد عملی توزیع آمیخته پوآسنی اصولاً دو مؤلفه ای یا در نهایت از ۳ مؤلفه ای استفاده می شود و بیان c مؤلفه ای صرفاً یک بیان تئوریک می باشد, در این مقاله توزیع آمیخته پوآسنی دو مؤلفه ای برای داده های مدت اقامت بیماران در نظر گرفته شده است، توزیع پوآسن آمیخته دو مؤلفه ای یا بالاتر در سالهای اخیر مورد توجه زیادی قرار گرفته است، توسط این مدل ها می توان بسیاری از داده های شمارشی که عدم حتمیت دارند (نادر هستند) را مورد پوشش قرار داد. این مدل هم در مطالعات پزشکی و هم در مطالعات مهندسی و اجتماعی جدیداً بسیار حائز اهمیت شده است. مدل آمیخته پوآسن یک حالت عمومی شده توزیع پوآسن است و در مواقعی که مشاهدات پراکندگی زیاد دارند مورد استفاده قرار می گیرند(۴). اگر y بیانگر متغیرتصادفی باشد که دارای توزیع پوآسن آمیخته است, زمانی توزیع پوآسن دو مؤلفه ای بکار می رود که

داده ها ناشی از دو گروه مختلف باشند مانند مدت اقامت بیماران که مؤلفه اول (α_1) بیانگر نسبت توزیع اکثر بیمارانی که حدوداً مدت اقامت آنها نزدیک به هم می باشد و مؤلفه دوم (α_2) بیانگر نسبت توزیع بیمارانی می باشد که دارای مدت اقامت کم یا زیاد (نقاط دور افتاده) می باشند. و توزیع بصورت مقابل خواهد بود:

$$f(y)=\alpha_1\frac{e^{-\lambda 1}\lambda 1^y}{y!}+(1-\alpha_1)\frac{e^{-\lambda 2}\lambda 2^y}{y!};y=0,1,2,3,....$$

مدل رگرسیون پوآسن آمیخته: از آنجائیکه برای مدل رگرسیونی مقدار در نظر گرفته می شود, بنابر این :

$$Log(\lambda_i)=\beta_{j0}+\beta_{j1}x_1+...+\beta_{jp}x_p;\ \text{i=1,2,...,c}$$

بیان می شود که در آن β_{jk} بیانگر k امین ضریب رگرسیونی برای متغیر مستقل x_k می باشد. علامت مثبت ضریب بیانگر ارتباط مستقیم عامل بر روی متغیر پاسخ شمارشی خواهد بود و در مورد متغیرهای کیفی با رفرانس قرار دادن یک طبقه، ارتباط بدست خواهد آمد(۸). این مدل رگرسیونی با استفاده از برنامه نوشته شده در نرم افزارS-Plus جهت تحلیل داده های مورد نظر بکار گرفته شده است. نیکوئی برازش مؤلفه ها در توزیع با استفاده از آماره Log-liklihood و $-2Log\lambda$ مورد بررسی قرار گرفته اند. در تمام حالات فرضیه صفر توزیع پوآسن یک مؤلفه ای در نظر گرفته شده است. بهبودی مدل با مؤلفه های بیشتر در کاهش مقدار $-2Log\lambda$ بوده است.

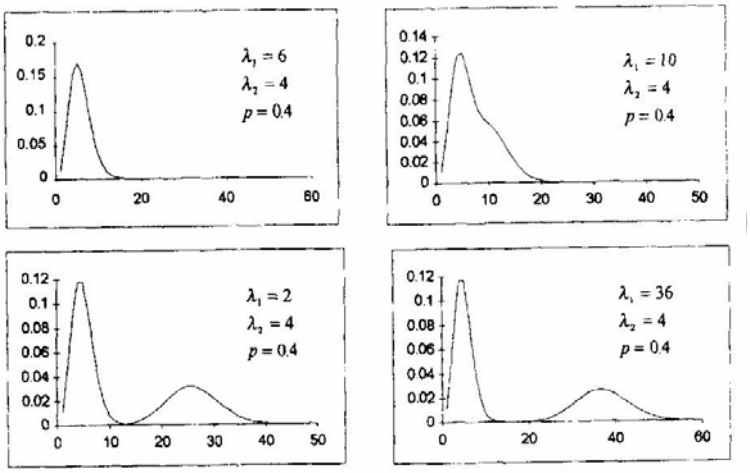
کاربرد: در این مقاله متغیر پاسخ مدت اقامت بیماران در بیمارستان در نظر گرفته شده است, علت این کاربرد این است که مدت اقامت بیماران بعنوان پیشامدهائی که در زمان خاص تعریف شده است می باشد و می تواند دارای توزیع پوآسن باشد(۱۲).

داده های مورد بررسی از یک مطالعه مشاهده ای و مقطعی که بر روی دو سری نمونه ۱۳۲۳ نفری از بخش جراحی و ۸۴۰ نفر از بخش داخلی از بیمارستان ولیعصر(عج) اراک در سال ۱۳۸۱ بصورت تصادفی سیستماتیک انتخاب گردیده بودند، مورد بررسی قرار گرفتند. بیمارستان ولیعصر(عج) تنها بیمارستان آموزشی وابسته به دانشگاه علوم پزشکی اراک می باشد که مراجعین به آن شامل تمام طبقات اجتماعی– اقتصادی می باشد. حجم نمونه ها بر اساس یک مطالعه مقدماتی از مقادیر مدت اقامت بیماران و استفاده از فرمول $n=\frac{Z^2_{\alpha/2}S^2}{d^2}$ بوده است، مقادیر S^2 بر اساس نمونه مقدماتی تصادفی ۱۰۰ نفر از هر بخش بوده است و حجم نمونه ها جدا گانه بدست آمده اند. نمونه ها بصورت تصادفی سیستماتیک از دو بخش انتخاب شدند و برای آنها متغیر مدت زمان اقامت در بیمارستان به عنوان متغیر پاسخ و متغیرهای سن, وضعیت تأهل, محل زندگی و

Archive of SID

www.SID.ir

محل تولد به عنوان متغیرهای مستقل تعریف شدند. مدت زمان اقامت بیماران تعداد روزهایی که بیمار از پذیرش تا ترخیص در بیمارستان گذرانده بود تعریف گردید (مقادیر بصورت صحیح گرد شده اند) بنابر این بصورت داده های شمارشی در نظر گرفته شده است.



شکل ۱. اشکال توزیع پوآسن دو مؤلفه ای با پارامترهای متفاوت

از آنجائیکه مقادیر مدت اقامت بیماران در بیمارستان دارای توزیع نامتقارن و چوله به راست می باشد(۲۱-۱۹)، در گذشته جهت تعدیل و تصحیح این موضوع در مطالعات مختلف ابتدا تغییر متغیر لگاریتمی جهت تصحیح چولگی متغیر مدت اقامت بیماران به کار گرفته شده و سپس رگرسیون چند گانه جهت بدست آوردن ارتباط این متغیر با متغیرهای مستقل دیگربکار می رفته است(۲۲). این فرض یک حالت غیر واقعی از همگنی و تقارن در داده های متغیر پاسخ ایجاد می کند. مطالعات زیادی در خصوص روشهای جدا کردن مقادیر دور افتاده از بقیه نقاط ارائه شده است که اکثراً بصورت منطقی توجیه پذیر نیستند, علت این امر نیز در بیان مشاهده ای است که داده های بعد از آن دور افتاده می باشند(۲۳).

داده ها ابتدابدون درنظرگرفتن بخش و محدویتهای دیگر جهت معین کردن مؤلفه های توزیع پوآسن به کارگرفته شده وسپس بخش بیمارستانی نیز مورد آنالیز قرار گرفته است. جهت تحلیل رگرسیون خطی از نرم افزارSAS و برنامه های نوشته شده در آن استفاده گردید.

یافته ها

کاربرد مدل بیان شده در این مقاله برروی نمونه های انتخاب شده از دو بخش جراحی و داخلی بیمارستان ولیعصر (عج) شهر اراک صورت گرفته است. متوسط مدت زمان بستری بیماران در بخش داخلی۴/۰۱ ± ۴/۲ روز و در بخش جراحی متوسط این متغیر ۶/۱۲ ± ۴/۹۷ روز بوده است. با توجه به پراکندگی بیشتر مقادیر مدت اقامت بیماران در بخش جراحی مدل آمیخته پوآسنی مدل مناسبی جهت تبیین مقادیر مدت اقامت بیماران در این بخش خواهد بود. جدول ۱ بیانگر آنالیز توزیع آمیخته جهت مدت زمان اقامت بیماران در دو بخش مختلف می باشد.

این جدول با استفاده از نرم افزار C.A.MAN بدست آمده است و نشان می دهد که جهت بخش داخلی مدل پوآسن یک مؤلفه ای با توجه به مقدار (Log-likelihood) مقادیر مدت اقامت بیماران را بهتر تشریح می کند. مقدار Log-likelihood جهت بخش داخلی در مدل پوآسن یک مؤلفه ای برابر ۲۷۶۶/۰۶۳- می باشد در صورتیکه وقتی مدل پوآسن دو مؤلفه ای بکار گرفته شودمقدار Log-likelihood برابر ۲۷۹۶/۲۱- و در سه مؤلفه ای ۲۸۰۱/۲۹-

Archive of SID

www.SID.ir

می باشد. بنابراین برای مقادیر مدت اقامت بیماران در بخش داخلی مدل پوآسن یک مؤلفه ای مناسب تر از دو یا سه مؤلفه می باشد. همچنین جهت بخش جراحی توزیع مدت اقامت بیماران با استفاده از مدل پوآسن دو مؤلفه ای بهتر از یک یا سه مؤلفه ای است. $-2\ln\lambda = 18.724, P-Value < 0.01$ در مدل سه مؤلفه ای مؤلفه ای مناسب تر بوده است، بنابراین رگرسیون بکار رفته نیز فقط جهت یک مؤلفه ای صورت گرفته است. با توجه به این جدول می بینیم که برای مدل با مؤلفه های بیشتر موارد معنادار تغییر زیادی در مقدار Log-likelihood ایجاد نمی شود.

جدول ۲ نتایج رگرسیون پوآسن آمیخته را جهت دو بخش نشان می دهد, با توجه به اینکه برای بخش داخلی مدل پوآسن یک بیشتری بدست آمده است. همچنین با توجه به جدول ۳ می بینیم که در رگرسیون پوآسن آمیخته موارد فاکتورهای معنادار خیلی بیشتر از رگرسیون خطی با تغییر متغیر لگاریتمی است.

جدول ۱. آنالیز پوآسن آمیخته جهت مدت زمان اقامت بیماران در دو بخش داخلی و جراحی در سال ۱۳۸۰

بخش	مؤلفه	مؤلفه اول			مؤلفه دوم			مؤلفه سوم		
		برآورد پارامتر	خطای استاندارد	*نسبت	برآورد پارامتر	خطای استاندارد	نسبت	برآورد پارامتر	خطای استاندارد	نسبت
جراحی	اولین مؤلفه	۴/۲۶۳	۰/۴۳۳	۱۰۰	۴/۰۶۱	۰/۲۰۳	۹۸/۲۵	۳/۹۲۶	۰/۳۰۷	۳۷/۵
	دومین مؤلفه	_____	_____	_____	۱۸/۹۲۱	۰/۹۸۵	۱/۷۵	۳/۹۸۴	۱/۱۶۰	۶۲/۱
	سومین مؤلفه	-۲۷۷۸/۹۸		_	___	_____	_____	۲۲/۸۷۱	۰/۷۶۸	۰/۴۱
	Log-likelihood				۲۴۲۱/۶۴			۲۴۲۰/۷۲	___	_____
داخلی	اولین مؤلفه	۶/۴۲۱	۰/۹۷۵	۱۰۰	۵/۲۹۱	۴/۲۹۱	۷۹/۲۱	۵/۰۴۱	۳/۲۱	۶۹
	دومین مؤلفه	_____	_____	___	۲۴/۷۱	۱۵/۶۵۲	۲۰/۷۹	۱۴/۷۱	۸/۴۹۱	۲۹/۰
	سومین مؤلفه	-۲۷۶۶/۰۶	_	_____	_____	_____	_____	۲۹/۷۲	۲۱/۲۰	۲/۰
	Log-likelihood						۳۸۷۲/۱۹	۲۸۰۱/۲۹	___	_____

*- منظور از نسبت، درصدی می باشد که داده ها تحت تبیین مؤلفه های توزیع می باشند(۲۸).

جدول ۲. نتایج آنالیز ضرائب رگرسیون پوآسنی آمیخته جهت بیماران بستری شده در دو بخش داخلی و جراحی در سال ۱۳۸۰

Archive of SID

www.SID.ir

بخش		بخش داخلی			بخش جراحی		
عامل	طبقه رفرانس	در کل	خطای استاندار ضرائب	درکل	خطای استاندار ضرائب	دومؤلفه ای	خطای استاندار ضرائب
عرض از مبدأ	-	**۰/۲۶۵	۰/۰۳۲	**۲/۵۷۱	۰/۵۴۴	** ۳/۷۲۱	۰/۸۲۸
سن بیماران	-	۰/۰۰۱۸	۰/۰۸۷	*-۰/۰۰۷	۰/۰۰۴	** *۰/۰۵۲۱	۰/۰۰۸
وضعیت تأهل	مجرد	- ۰/۰۰۲	۰/۰۰۸	-۰/۰۴۱	۰/۰۷۴	*۰/۰۵۴۹	۰/۰۲۹
محل تولد	شهر	-۰/۰۶۲	۰/۴۱۱	**-۰/۲۱۵	۰/۰۹۶	-۰ /۰۱۷۸	۰/۰۲۸
محل زندگی	شهر	**۰/۱۲۸	۰/۰۱۵	۰/۰۷۱	۰/۰۸۳	*-۰/۴۷۰	۰/۲۷۳

p<۰/۰۵ * p<۰/۰۱ ** p<۰/۰۰۱ ***

Archive of SID

www.SID.ir

جدول ۳. مقایسه نتایج رگرسیون خطی با رگرسیون پوآسن آمیخته در بخش جراحی در سال ۱۳۸۰

عامل	طبقه رفرانس	رگرسیون خطی (باتغییر لگاریتمی)		رگرسیون پوآسنی رگرسیون پوآسن		رگرسیون پوآسن آمیخته یک مؤلفه ای		رگرسیون پوآسن آمیخته دومؤلفه ای	
		ضریب	خطای استاندار	ضریب	خطای استاندار	ضریب	خطای استاندار	ضریب	خطای استاندار
عرض از مبدأ	ـــــ	۲/۲۶۱ **	۰/۰۹۵	۲/۵۷۱ **	۰/۴۲۴	۳/۷۲۱ ***	۰/۰۸۷	***۳/۷۲۱	۰/۱۲۱
سن بیماران	ـــــــ	۰/۰۲۱	۰/۰۸۹	-۰/۰۰۷ *	۰/۰۰۴	۰/۰۰۲	۰/۰۵۵	-۰/۰۵۲۱ ***	۰/۰۰۰
وضعیت تأهل	مجرد	۰/۰۰۹	۰/۰۱۶	-۰/۰۰۲	۰/۰۷۵	۰/۰۱۲	۰/۰۲۱	۰/۸۶۶ ***	۰/۰۴۴
محل تولد	شهر	۰/۰۳۲	۰/۰۹۲	-۰/۰۶۸	۰/۲۱۹	-۰/۲۱۵ **	۰/۰۴۲	-۰/۱۵۲ **	۰/۰۱۶
محل زندگی	شهر	-۰/۰۱	۰/۰۴۷	-۰/۰۷۱	۰/۶۲۹	۰/۱۲۶	۰/۲۹۶	۰/۴۷ *	۰/۲۱۰
-۳۷۲۴ /۵		آماره Deviance		-۱۲۸۷/۲۵-				-۸۲۷/۴	

p<۰/۰۵ * p<۰/۰۱ ** p<۰/۰۰۱ ***

بحث و نتیجه گیری

با توجه به مقادیر Log-likehood و جداول ۱و۲ می بینیم که آنالیز توزیع آمیخته اغلب روش مناسبی برای آنالیز مقادیر مدت اقامت بیماران و بدست آوردن برخی فاکتورهای مرتبط می باشد. روش های رگرسیون سنتی (خطی) بر پایه فرضیات خاصی (نرمال بودن و همگنی واریانس ها) می باشند که در خصوص مقادیر مدت اقامت بیماران در بیمارستان اصلاً بر قرار نیستند، حتی زمانی که تغییر لگاریتمی در خصوص متغیر فوق نیز صورت گیرد، هنوز نرمال بودن و همگنی واریانس مشاهدات پذیرفته نمی شود.

با توجه به جدول ۱ مدل پوآسن یک مؤلفه ای مدل مناسبی جهت بیماران بستری در بخش داخلی و مدل دو مؤلفه ای جهت داده های مربوط به مدت زمان اقامت بیماران در بخش جراحی بوده است. شاید در بخش جراحی پراکندگی مدت اقامت بیماران بیشتر از بخش داخلی بوده است. با توجه به جدول ۲ که بیانگر ضرائب رگرسیون پوآسن آمیخته جهت دو بخش جراحی و داخلی می باشد, می بینیم که در بخش داخلی عرض از مبدأ خط رگرسیونی و فاکتورهای سن بیماران و محل زندگی عوامل معنادار بر روی مدت زمان اقامت بیماران در این بخش دارند. همچنین سن اثر مستقیم بر روی مدت اقامت و محل زندگی دارد (اگر شهر را به عنوان رفرانس در نظر بگیریم) مدت زمان اقامت بیماران روستائی بیشتر از شهری می باشد در مطالعات انجام شده نیز این مطلب تأئید شده است(۱۲). همین موارد در مطالعه دیگری که در کشور استرالیا بر روی مدت اقامت بیماران با استفاده از توزیع آمیخته نرمال صورت گرفته است، تأئید گردیده است(۱۱). با توجه به همین جدول در بخش جراحی می بینیم که چون مدل دو مؤلفه ای پوآسن برای این بخش مناسب تر بوده است، در جدول فوق نیز فقط ضرائب دو مؤلفه ای در نظر گرفته شده است. عرض از مبدأ خط رگرسیون و عوامل سن، وضعیت تأهل، محل تولد و محل زندگی همگی معنادار می باشند. با توجه به علامت و طبقه رفرانس می بینیم که سن اثر مستقیم بر روی مدت اقامت بیماران در این بخش و وضعیت تأهل با رفرانس قرار دادن طبقه مجرد اثر عکس بر روی مدت اقامت دارد. به این معنا که افراد متأهل مدت زمان بیشتری را در این بخش اقامت داشته اند.

همینطور محل تولد با رفرانس قرار دادن شهر اثر معکوس بر طول مدت اقامت بیماران در این بخش دارد، بنابراین بیمارانی که

Archive of SID

www.SID.ir

محل زندگی آنها شهر نمی باشد، مدت زمان اقامت بیشتری در این بخش داشته اند. در خصوص محل زندگی نیز دقیقاً مفهوم بالا بر قرار می باشد، در مطالعه ای که در استرالیا انجام شده نیز مورد فوق بر قرار است(۱۲).

با در نظر گرفتن دو جدول ۱و ۲ می توانیم نتیجه بگیریم که ضرائب معنادار در رگرسیون با دو مؤلفه، بیشتر از یک مؤلفه بوده است. در نهایت با توجه به جدول ۳ که بیانگر مقایسه بین رگرسیون پوآسن آمیخته و رگرسیون خطی ساده با تغییر متغیر لگاریتمی می باشد، می بینیم که در رگرسیون خطی فقط عرض از مبدأ خط رگرسیونی معنادار است در صورتیکه در رگرسیون دو مؤلفه ای تمام عوامل سن, وضعیت تأهل، محل زندگی و محل تولد معنادارند و از آنجائیکه مناسبت مدل دو مؤلفه ای جهت این بخش بررسی شده و تأئید شده است، لذا این عوامل ظاهری و ناشی از خطا نبوده اند. در مطالعات انجام شده جهت مدل بندی اقامت بیماران سن را به عنوان عامل معنادار اعلام نکرده اند، در اکثر این مطالعات روش تجزیه و تحلیل همان رگرسیون خطی با تغییر لگاریتمی مقادیر متغیر پاسخ بوده است(۲۵و۲۴). در مطالعات انجام شده برای تحلیل مدت اقامت بیماران در بیمارستان از مدل های دیگری همانند مدل های خطی تعمیم یافته نیز استفاده گردیده است، در این مطالعات قید گردیده است که ناهمگنی داده ها تأثیر بسزائی در عدم اعتبار مدلهای بیان شده داشته است و در انتهای این مطالعات به پیشنهاد بکارگیری مدلهای آمیخته جهت مدل بندی مدت اقامت پرداخته اند(۲۹-۲۶).

****

References

1. Skrondal A, Rabe S. Some applications of generalized linear latent and mixed models in epidemiology: repeated measures, measurement error and multilevel modeling. Norsk Epidemiology 2003; 13(2): 265-78.

2. Wang K, Kelvin KWY, Lee AH. A zero-inflated poisson mixed model to analyze diagnosis related groups with majority of same-day hospital stays. Comput Methods Programs Biomed 2002; 68(3): 195-203.

3. Zelterman D. Discrete distributions: applications in health. New york: John Wiley 2004; pp: 141-61.

4. Xio W, Xio M. A mixed poisson model and its application to attribute testing data. J Microelectron Reliab 1996; 36(2): 133-40.

5. Ng SK, Yau KKW, Lee AH. Modeling inpatient length of stay by hierarchical mixture regression via the EM algorithm. Math Comput Model 2003; 37(3-4): 365-75.

6. Karlis D, Xekalaki E. Mixed poisson distributions. Int Stat Re 2005; 73(1): 35-58.

7. Jie QG, Trivedi PK. Flexible parametric models for long-tailed paitent count distributions. Oxford Bulletin Econ Stat 2002; 64(63): 63-82.

8. Schattmann P, Dietz E, Bohning D. Covariate adjusted mixture models and disease mapping with the program Dismapwin, Stat Med 1996; 15(7-9): 919-26.

9. Marazzi A, Paccaud F, Rueux C, et al. Fitting the distributions of length of stay by parametric models. Med Care 1998; 36(6): 915– 27.

10. Marshall A. Length of stay-based patient flow models: Recent developments and future directions. Health Care Manage Sci 2005; 8: 213–20.

11. Everitte BS, Hand DJ. Finite mixture distributions, London, Chapman and Hall 1981; pp: 95-140.

Archive of SID

www.SID.ir

12. Xio J, Lee A, Vemuri S. Mixture distribution analysis of length of hospital stay for efficient funding. J Socio- Econ Plan Sci 1999; 33(1): 39-59.

13. Solomon GL. Length of the hospital stay for mothers and newborns. N Engl J Med 1996; 334(17): 1134-9.

14. Zernikow B, Holtmannspotter K, Michel E, et al. Predicting length-of-stay in preterm neonates. Eur J Pediat 1999; 158(1): 59–62.

15. Malkin JD, Keeler E, Broder MS, et al. Postpartum length of stay and newborn health: A cost-effectiveness analysis. Pediatrics 2003; 111 (4 Pt 1):e316–22.

16. Bianco A, Pileggi C, Trani F, Angelillo IF. Appropriateness of admissions and days of stay in pediatric wards of Italy. Pediatrics 2003; 112 (1 Pt 1): 124–8.

17. Srivastava R, Homer CJ. Length of stay for common pediatric conditions: teaching versus nonteaching hospitals. Pediatrics 2003; 112(2): 278–81.

18. Silber JH, Rosenbaum PR, Even-Shoshan O, et al. Length of stay, conditional length of stay, and prolonged stay in pediatric asthma. Health Serv Res 2003; 38(3): 867–86.

19. Shachtman RH, Snapinn SM, Quade D, et al. A method for constructing case-mix indexes, with application to hospital length of stay. Health Serv Res 1986; 20 (6 Pt 1): 737-62.

20. Silberbach M, Shumaker D, Menashe V, et al. Predicting hospital charge and length of stay for congenital heart disease surgery. Am J Cardiol 1996; 72(12): 985-63.

21. Wolf MW, Roubin GS, Schweiger M. Length of hospital stay and complications after percutaneous transluminal coronary angioplasty. Circulation 1995; 92: 311-19.

22. Lave JR, Frank RG. Factors affecting medical patients length of stay in psychiatric units. Health Care Financ Rev 1998; 10(2): 57-66.

23. Fetter RB. The new ICD-9-cm diagnosis related groups classification scheme. HCFA contract report. Health Care Financ Administr, Washington DC, 1984; pp: 10-22.

24. Epstein AM, Stern RS, Tognetti J. The association patients socio-economic characteristics with the length of hospital stay and hospital charges within diagnostic related groups. N Engl J Med 1988; 318(24): 1579-85.

25. Marissa A, Packard F, Ruffle C, et al. Fitting the distributions of length of stay by parametric models. Med Care 1998; 36(6): 915-27.

26. McLachlan G, Peel D. Finite mixture models. New York, John Wiley 2004; pp: 51-2.

27. Leung KM, Elashoff RM, Ress KS, et al. Hospital-and patient-related characteristics determining maternity length of stay: A hierarchical linear model approach. Am J Public Health 1998; 88(3): 377-81.

28. Rental R, Kiess MJ, DesHarnais S, Reutter K. Applications for risk-adjusted outcome measures. Qual Assur in Health Care 1999; 3: 283-92.

Archive of SID

www.SID.ir

30. Lee AH, Ng ASK, Yau KKW. Determinants of maternity length of stay: A gamma mixture risk- adjusted model. Health Care Manage Sci 2001; 4 (4): 249-55.

* آدرس نویسنده مسئول: شیراز، دانشکده بهداشت، گروه آمار زیستی، تلفن: ۰۷۱۱-۲۲۵۲۱۰۰۱-۸.

rajaeefard@sums.ac.ir

Archive of SID

www.SID.ir