



Comparison of Efficiency in Generalized Poisson Regression Model and the Standard Poisson Regression Model in Analyzing Fertility Behavior among Women, Kashan, 2012

Hossein Falahzadeh (Ph.D)¹, Zahra Jahanarae (MS.c)², Mohssen Askarishahi (Ph.D)³, Mohammad Lotfi (Ph.D)⁴

1. Professor, Department of Biostatistics and Epidemiology, School of Public Health, Yazd Shahid Sadoughi Univ Med Sci Yazd, Iran.

2. Corresponding Author: MSc Student in Biostatistics, School of Public Health, Yazd Shahid Sadoughi Univ Med Sci Yazd, Iran.

Email: jahanarae@gmail.com Tel: 09132595708

3. Assistant Professor, Department of Biostatistics and Epidemiology, School of Public Health, Yazd Shahid Sadoughi Univ Med Sci Yazd, Iran.

4. Professor, Department of Biostatistics and Epidemiology, School of Public Health, Yazd Shahid Sadoughi Univ Med Sci - Yazd, Iran.

Abstract

Introduction: Different statistical methods can be used to analyze fertility data. When the response variable is discrete, Poisson model is applied. If the condition does not hold for the Poisson model, its generalized model will be applied. The goal of this study was to compare the efficiency of generalized Poisson regression model with the standard Poisson regression model in estimating the coefficient of effective factors in the number of children.

Methods: This is a cross-sectional study carried out on a population of married women with the age range of 15-49 years old in Kashan, Iran. The cluster sampling method was used for data collection. Clusters consisted of the urban blocks determined by the municipality. The total number of 10 clusters each of 30 households was selected according to the health center's framework. The necessary data were then collected through a self-made questionnaire and direct interviews with women under study. Further, the data analysis was performed by using the standard and generalized Poisson regression models through the R software.

Results: The average number of children for each woman was 1.45 with a variance of 1.073. A significant relationship was observed between the husband's age, number of unwanted pregnancies, and the average duration of breastfeeding with the present number of children in the two standard and generalized Poisson regression models ($p < 0.05$). The mean age of women participating in this study was 33.1 ± 7.57 years (from 25.53 years to 40.67) and the mean age of marriage was 20.09 ± 3.82 (from 16.27 years to 23.91), and the mean age of their husbands was 37.9 ± 8.4 years (from 29.5 years to 46.3). In the current study, the majority of women were in the age range of 30-35 years old with the median of 32 years old, however, most of men were in the age range of 35-40 years old with the median of 37 years old. While 236 of women did not have unwanted pregnancies, most participants of the present study had one unwanted pregnancy.

Conclusion: According to the achieved results, the generalized Poisson regression model is more suitable for data analysis and can estimate the coefficient of effective factors on the number of children more precisely.

Keywords: Fertility behavior, Standard Poisson model, Generalized poisson model, Number of children

Conflict of interest: The authors declared that there is no Conflict interests.



This Paper Should be Cited as: Hossein Falahzadeh (Ph.D), Zahra Jahanarae (MS.c), Mohssen Askarishahi (Ph.D), Mohammad Lotfi (Ph.D). Comparison of Efficiency in Generalized Poisson Regression Model and the Standard... J Tolooebehdasht Sci 2017; 15(1):1-9[Persian]



مقایسه کارایی مدل‌های رگرسیون پواسن تعمیم یافته با رگرسیون پواسن استاندارد در تحلیل رفتار باروری زنان شهر کاشان در سال ۱۳۹۱

نویسندگان: حسین فلاح زاده^۱، زهرا جهان آرای^۲، محسن
عسکری شاهی^۳، محمد حسن لطفی^۴

۱. استاد گروه آمار و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید صدوقی یزد
۲. نویسنده مسئول: دانشجوی کارشناسی ارشد آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید صدوقی یزد. تماس: ۰۹۱۰۳۱۰۷۰۳۰. Email: jahanarae@gmail.com
۳. استادیار گروه آمار و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید صدوقی یزد
۴. استاد گروه آمار و اپیدمیولوژی دانشکده بهداشت، دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید صدوقی یزد

چکیده

مقدمه: در تحلیل داده‌های باروری از روشهای مختلف آماری می‌توان استفاده نمود. زمانی که متغیر پاسخ گسسته باشد، مدل پواسن استفاده می‌شود. در صورتی که شرایط مدل پواسن برقرار نباشد، از تعمیم یافته آن استفاده می‌شود. هدف از این مطالعه، مقایسه‌ی برازندگی مدل تعمیم یافته نسبت به مدل پواسن معمولی در برآورد ضرائب عوامل موثر بر تعداد فرزندان موجود می‌باشد.

روش بررسی: این مطالعه مقطعی بوده و جامعه‌ی مورد مطالعه، زنان ازدواج کرده‌ی ۴۹-۱۵ شهر کاشان می‌باشند. روش انتخاب نمونه به صورت خوشه‌ای و خوشه‌های مورد نظر بلوک‌های شهری بودند که توسط شهرداری تعیین شده بود. تعداد ۱۰ خوشه، هر خوشه ۳۰ خانوار بر اساس چارچوب مرکز بهداشت جمع‌آوری گردید و اطلاعات لازم از طریق پرسشنامه‌ی خودساخته و یا مصاحبه‌ی مستقیم با زنان توسط مراجعه به درب منازل آنها تکمیل گردید. سپس تحلیل داده‌ها با استفاده از رگرسیون پواسن معمولی و تعمیم یافته توسط نرم افزار R انجام شد.

یافته‌ها: میانگین فرزندان هر زن ۱/۴۵ و واریانس آن ۱/۰۷۳ است. سن همسر، تعداد حاملگی ناخواسته و متوسط طول مدت شیردهی در دو مدل پواسن معمولی و پواسن استاندارد با تعداد فرزندان موجود رابطه معنی دار داشتند ($p < .05$). میانگین سنی زنان شرکت کننده در این مطالعه $33/1 \pm 7/57$ سال (۴۰/۶۷ تا ۲۵/۵۳ سال) و میانگین سنی ازدواج زنان $3/82 \pm 2/09$ سال (۲۳/۹۱ تا ۱۶/۲۷) (C.I, 5%: ۲۳/۹۱ تا ۱۶/۲۷) و میانگین سنی همسران $37/9 \pm 8/4$ سال (۴۶/۳ تا ۲۹/۵) (C.I, 5%: ۴۶/۳ تا ۲۹/۵) بود. همانطور که ملاحظه می‌شود بیشتر زنان در گروه سنی ۳۵-۳۰ سال و میانه‌ی سنی آنها ۳۲ سال و بیشتر مردان در گروه سنی ۴۰-۳۵ سال و میانه‌ی سنی آنها ۳۷ سال می‌باشند. در این مطالعه بیشتر زنان دارای حاملگی ناخواسته هستند و ۲۳۶ نفر از آنها حاملگی ناخواسته ندارند.

نتیجه گیری: نتایج مطالعه نشان داد که مدل تعمیم یافته نسبت به مدل پواسن معمولی جهت تحلیل داده‌ها مناسب‌تر می‌باشد و می‌تواند ضرائب عوامل موثر بر تعداد فرزندان را دقیق‌تر برآورد نماید.

واژه های کلیدی: رفتار باروری، مدل رگرسیون پواسن، مدل رگرسیون پواسن تعمیم یافته، تعداد فرزندان
این مقاله برگرفته از پایان نامه دانشجویی کارشناسی ارشد آمار زیستی دانشکده بهداشت دانشگاه علوم پزشکی شهید صدوقی یزد می‌باشد.

طلوع بهداشت

دو ماهنامه علمی پژوهشی

دانشکده بهداشت یزد

سال شانزدهم

شماره: اول

فروردین و اردیبهشت ۱۳۹۶

شماره مسلسل: ۶۱

تاریخ وصول: ۱۳۹۲/۹/۳

تاریخ پذیرش: ۱۳۹۲/۱۰/۱۴

افزایش بی‌رویه جمعیت و مشکلات ناشی از آن امروزه یکی از مهمترین معضلات جامعه بخصوص کشورهای در حال توسعه می‌باشد. یکی از اساسی‌ترین علل آن حاملگی‌های ناخواسته می‌باشد (۱). حاملگی‌های ناخواسته نه تنها باعث افزایش بعد خانوار و کاهش فاصله بین تولدها می‌شود، بلکه زمینه خطر سازی جهت مرگ و میر مادران و کودکان می‌باشد. زیرا حاملگی ناخواسته در صورت ادامه حاملگی باعث افزایش فشار روانی مادر گردیده، که این امر نیز منتهی به عدم مراقبت کافی از خود و جنین می‌گردد. کاهش حاملگی‌های ناخواسته عامل موثری در سلامت باروری محسوب شده و تلاش در این امر کیفیت زندگی مادران را ارتقاء می‌بخشد (۲). با توجه به اینکه دوران بارداری از حساس‌ترین دوره‌های زندگی زنان بوده و بیشتر مرگ‌هایی که در این دوران یا حین زایمان اتفاق می‌افتند، قابل پیشگیری هستند، ارائه مراقبت بهداشتی دوران بارداری و پس از آن یکی از راهکارهای اساسی در ارتقاء سلامت مادران و پیشگیری از مرگ و میر آنان می‌باشد. در مراقبت‌های روتین از مادران باردار مداخلات پایش وزن‌گیری، آموزش تغذیه، شیوه زندگی، تامین مکمل‌های آهن، ویتامین‌ها، اسید فولیک و غربالگری بیماری‌هایی چون دیابت، بیماری‌های عفونی، علائم پره اکلامپسی برای بیش از سه دهه انجام می‌گیرد (۳). یکی از فرضیاتی که در مورد باروری در هر جامعه‌ای مطرح است، مسئله بقاء فرزندان است. هرچه مرگ و میر نوزادان و کودکان بالاتر باشد، زوجین تمایل بیشتری به داشتن فرزند دارند تا از این طریق جایگزینی برای فرزندان از دست داده‌ی خود داشته باشند (۴). این موضوع توسط برخی از جمعیت‌شناسان تحت



عنوان عامل جانشینی مطرح شده است. بنابراین تحلیل داده‌های مربوط به رفتار باروری از اهمیت خاصی برخوردار است. برای تحلیل این داده‌ها می‌توان از روش‌های مختلف آماری استفاده کرد. داده‌های مربوط به رفتار باروری با تعداد فرزندهای زنده متولد شده به عنوان متغیر پاسخ، از نوع شمارشی می‌باشند، بنابراین برای مدل‌بندی این داده‌ها باید از مدل‌های مربوط به داده‌های شمارشی استفاده نمود. مدل‌بندی آماری یکی از روش‌های تبیین تغییرات در داده‌های مشاهده شده از طریق معادلات ریاضی می‌باشد که با استفاده از آن می‌توان به چگونگی تغییرات داده‌ها پی‌برد (۵). مدل‌های داده‌های شمارشی به طور گسترده در مطالعات تجربی استفاده می‌شود. مدل‌های معمول برای داده‌های شمارشی عبارتند از: مدل پواسن، مدل دوجمله‌ای و مدل دوجمله‌ای منفی (۶، ۷). شرط استفاده از مدل پواسن معادل بودن میانگین و واریانس است. اگر این شرط برقرار نباشد، برآورد در مدل رگرسیون پواسن سازگار است اما کارا نیست. بنابراین استنباط بر اساس برآورد خطای معیار اعتبار زیادی ندارد. بنابراین برای مقابله با این مشکل از مدل پواسن تعمیم‌یافته استفاده می‌کنیم (۹-۶). در واقع conway و maxwell توزیع پواسن را با اضافه کردن پارامتری به مدل توسعه دادند که این پارامتر پراکندگی می‌باشد. (بیش پراکندگی و کم پراکندگی). مدل پواسن تعمیم‌یافته رقیب خوبی برای مدل دو جمله‌ای منفی است که داده‌های شمارشی آن بیش پراکنده هستند (۱۰) در سال‌های اخیر توجه زیادی روی آنالیز داده‌های شمارشی شده که نشان می‌دهد بیش پراکندگی رابطه‌ای با توزیع پواسن دارد. در مواردی که داده‌ها دارای بیش پراکندگی می‌باشند توزیع دوجمله‌ای منفی و توزیع پواسن آمیخته به طور



$S=3$ و خطای برآورد $1/4$ و سطح اطمینان 95% . تعداد 200 نمونه مورد نیاز است. با احتساب ضریب نمونه گیری خوشه‌ای به نمونه گیری تصادفی ساده تعداد نمونه در $1/4$ ضرب می‌گردد بنابراین تعداد 300 نفر مورد بررسی قرار گرفته است. خوشه‌های مورد نظر بلوکهای شهری می‌باشد که توسط شهرداری تعیین می‌شود. تعداد 10 خوشه، هر خوشه 30 خانوار بر اساس چارچوب مرکز بهداشت جمع‌آوری می‌شود و اطلاعات لازم از طریق پرسشنامه‌ی خودساخته و با مصاحبه‌ی مستقیم با زنان توسط مراجعه به درب منازل آنها تکمیل می‌گردد. این پرسشنامه شامل اطلاعات دموگرافیک (سن زن، سن شوهر، تحصیلات زن، تحصیلات شوهر، سن ازدواج زن، شغل همسر، وضعیت اقتصادی و اجتماعی و تعداد فرزندان) می‌باشد. اطلاعات داده‌های جمع‌آوری شده از طریق نرم‌افزار R تحلیل شدند و برای تجزیه و تحلیل داده‌ها از روش‌های آماری رگرسیون پواسنی استفاده شده و $p < 0.05$ معنی‌داری در نظر گرفته می‌شود. در تئوری آمار و احتمال، توزیع پواسون (یا قانون پواسون اعداد کوچک) یک توزیع احتمالی گسسته است که احتمال اینکه یک حادثه به تعداد مشخصی در فاصله‌ی زمانی یا مکانی ثابتی رخ دهد را شرح می‌دهد، به شرط اینکه این حوادث با نرخ میانگین مشخصی و مستقل از زمان آخرین حادثه رخ دهند. اگر Y_i متغیرهای تصادفی مستقل با شرایط بالا باشد، در اینصورت Y_i دارای توزیع پواسن است. میانگین و واریانس توزیع پواسن با هم برابرند یعنی $E(Y_i) = \text{var}(Y_i) = \mu_i$ و $\mu_i = \exp(\beta x_i)$ که x_i بردار K بعدی از متغیرهای برونزا می‌باشد و β بردار $1 \times K$ بعدی پارامترهاست. در صورتیکه تعداد پاسخ‌ها تحت تاثیر زمان باشند آنگاه $\mu_i = t_i \exp(\beta x_i)$

گسترده مورد استفاده قرار می‌گیرد (۱۱). گاهی تعداد صفرها در نمونه‌های شمارشی بیشتر از مقدار مورد انتظار تحت مدل رگرسیونی است. در این موارد، استفاده از مدل با انباشتگی صفر و مدل هاردل توصیه شده است که این مدل‌ها می‌توانند هم بر بیش پراکندگی و هم بر کم پراکندگی مقابله کنند (۷). گاهی مدل‌سازی باروری زنان با استفاده از مدل داده‌های شمارشی با انباشتگی انجام می‌شود. داده‌های شمارشی به طور معمول دارای توزیع نرمال نمی‌باشند و چوله به راست یا چپ هستند. بنابراین روشهای آماری مبتنی بر توزیع نرمال برای تحلیل چنین داده‌هایی مناسب نیست، در این حالت استفاده از مدل‌های تعمیم‌یافته مناسب‌تر است (۱۲). هدف اصلی این مطالعات ارائه کردن یک مدل داده‌ی شمارشی قابل انعطاف است که به طور قابل قبولی الگوهای باروری مشاهده شده را شرح دهد (۱۳). مدل با انباشتگی صفر می‌تواند در جاهای دیگری به غیر از باروری مثل بیمه بکار رود. در این مطالعه، در نظر داریم که ضمن معرفی مدل‌های رگرسیونی مربوط به داده‌های شمارشی، کاربرد برخی از آنها را بر روی تحلیل رفتارهای باروری جهت برآورد ضرایب برخی عوامل مؤثر فردی و اجتماعی بر تعداد فرزندان بررسی و نتایج به دست آمده از مدل‌های مختلف را با یکدیگر مقایسه نماییم.

روش بررسی

این مطالعه مقطعی بر روی 300 نفر از زنان متاهل $15-49$ ساله شهر کاشان که در سن باروری قرار داشتند و با استفاده از نمونه گیری خوشه‌ای انتخاب شدند، انجام گردید. با توجه به اینکه متغیر هدف در این مطالعه تعداد فرزندان موجود است و با توجه به مطالعات مشابه، مقدار حداکثر انحراف معیار تعداد فرزندان



متغیرها را کاهش می‌دهد و باعث خطای نپذیرفتن آن متغیر در مدل می‌گردد (۶). اگر Y_i دارای توزیع پواسن تعمیم‌یافته با پراکندگی باشد، در اینصورت $\text{var}(Y_i) = \mu_i / (1 + \alpha\mu_i)^2$ و $E(Y_i) = \mu_i$ که $\mu_i = t_i \exp(\beta x_i)$ تعداد فرزندان هر زن را نشان می‌دهد و α پارامتر پراکندگی می‌باشد. برای برآورد ضرائب پارامترهای مدل‌های پواسنی از روش درست‌نمایی ماکزیمم استفاده می‌شود. نیکویی برازش مدل‌ها در توزیع با استفاده از آماره لگاریتم درست‌نمایی آزمون می‌گردد.

یافته‌ها

در این مطالعه داده‌های ۳۰۰ زن در شهر کاشان مورد بررسی قرار گرفت. میانگین فرزندان ۱/۴۵ و واریانس آن ۱/۰۷ بود. همچنین مد توزیع فراوانی تعداد فرزندان ۱ و میانه تعداد فرزندان ۱ بدست آمد. حدود ۴٪ بیش از ۳ فرزند و درصد قابل توجهی از زنان دارای ۱ یا ۲ فرزند می‌باشند (جدول ۱).

که t_i متغیر زمان و مدت زمان مواجهه فرد است. در تحلیل داده‌های باروری، تعداد فرزندان به شدت تحت تاثیر طول دوره زناشویی فرد می‌باشد. طبیعی است که هر چه طول دوره زناشویی بیشتر باشد تعداد فرزندان بیشتر است بنابراین این متغیر باید به عنوان متغیر زمان (offset) در مدل‌ها وارد شود. اگر فرضیات توزیع پواسن برقرار نباشد واریانس این توزیع با میانگین آن برابر نیست که در اینصورت از پواسن تعمیم‌یافته استفاده می‌کنیم. اگر واریانس بزرگتر از میانگین باشد، پواسن تعمیم‌یافته، داده‌های شمارشی را با پراکندگی زیاد ارائه می‌دهد که منجر به کم برآوردی خطای معیار شده و مقدار آماره‌ی آزمون هر کدام از ضرائب متغیرها را افزایش می‌دهد و باعث خطای پذیرش آن متغیر در مدل می‌گردد. اگر واریانس کوچکتر از میانگین باشد، پواسن تعمیم‌یافته داده‌های شمارشی را با پراکندگی کم ارائه می‌دهد که منجر به بیش برآوردی خطای معیار شده و مقدار آماره‌ی آزمون هر کدام از ضرائب

جدول ۱: توزیع فراوانی تعداد فرزندان

تعداد فرزند	فراوانی	فراوانی نسبی	درصد فراوانی نسبی	فراوانی تجمعی	درصد فراوانی تجمعی
۰	۵۵	۰/۱۸۳۳۳	۱۸/۳۳	۵۵	۱۸/۳۳
۱	۱۰۹	۰/۳۶۳۳۳	۳۶/۳۳	۱۶۴	۵۴/۶۶
۲	۹۷	۰/۳۲۳۳۳	۳۲/۳۳	۲۶۱	۸۷
۳	۲۷	۰/۰۹	۹	۲۸۸	۹۶
۴	۱۰	۰/۰۳۳۳۳	۳/۳۳	۲۹۸	۹۹/۳۳
۵	۲	۰/۰۰۶۶۶	۰/۶۶۶	۳۰۰	۱۰۰



(جدول ۳). عوامل موثر بر تعداد فرزندان در دو مدل پواسن معمولی و تعمیم یافته مورد مقایسه قرار گرفته است (جدول ۴). در این جدول برآورد ضرائب هر کدام از متغیرهای مستقل همراه با خطای استاندارد آن و p آن آورده شده است. در توزیع پواسن تعمیم یافته مقدار α نشان دهنده‌ی میزان پراکندگی می‌باشد و متغیر طول دوره زناشویی به عنوان offset است. برای ارزیابی مناسب بودن مدل، مقدار لگاریتم درست نمایی برای مدل‌ها بیان گردید. مقدار کمتر این آماره از نظر قدر مطلق دلالت بر بهتر بودن مدل برازشی می‌کند. بر اساس جدول ۴ مدل پواسن تعمیم یافته مناسب‌تر از مدل پواسن معمولی است که بر اساس نتایج دو مدل متغیرهای سن همسر، تعداد حاملگی ناخواسته و متوسط شیردهی معنی دار شدند.

جدول ۲: توزیع فراوانی ویژگی‌های جمعیت شناختی زنان و همسران

در این مطالعه توزیع فراوانی ویژگی‌های جمعیت شناختی زنان و همسران آنها آورده شده است. میانگین سنی زنان شرکت کننده در این مطالعه $33/1 \pm 7/57$ سال (۲۵/۵۳ تا ۴۰/۶۷ سال) و میانگین سنی همسران $37/9 \pm 8/4$ سال (۲۹/۵ تا ۴۶/۳۱: C.I,5%) بود. همانطور که ملاحظه می‌شود بیشتر زنان در گروه سنی ۲۵-۳۵ سال و میانه‌ی سنی آنها ۳۲ سال و بیشتر مردان در گروه سنی ۳۰-۴۰ سال و میانه‌ی سنی آنها ۳۷ سال می‌باشد. میانگین سنی ازدواج زنان $20/09 \pm 3/82$ سال (۱۶/۲۷ تا ۲۳/۹۱: C.I,5%) بود و در این مطالعه بیشتر زنان دارای ۱ حاملگی ناخواسته هستند و ۲۳۶ نفر از آنها حاملگی ناخواسته ندارند (جدول ۲).

همچنین میانگین و انحراف معیار متغیرهای طول دوره زناشویی، سن زن و سن ازدواج زن نیز آورده شده است

نام متغیر	بر حسب متغیر	تعداد	درصد
تعداد حاملگی ناخواسته	۰	۲۳۶	۷۸/۷
	۱	۵۳	۱۷/۷
	۲	۱۰	۳/۳
سن همسر	۳	۱	۰/۳
	۲۰-۲۹	۵۲	۱۷/۳۳
	۳۰-۳۹	۱۲۴	۴۱/۳۳
	۴۰-۴۹	۹۱	۳۰/۳۳
	۵۰-۵۹	۳۰	۱۰
	۶۰ به بالا	۳	۱
سن زن	۱۵-۲۴	۴۳	۱۴/۳۳
	۲۵-۳۴	۱۴۱	۴۷/۰
	۳۵-۴۴	۸۴	۲۸/۰
سن ازدواج زن	۱۰-۱۹	۱۴۱	۴۷/۰
	۲۰-۲۹	۱۵۴	۵۱/۳۳
	۳۰-۳۹	۵	۱/۶۷
	۱-۹	۱۳۰	۴۳/۳۳
	۱۰-۱۹	۹۸	۳۲/۶۶
طول دوره زناشویی	۲۰-۲۹	۵۹	۱۹/۶۶
	۳۰ به بالاتر	۱۳	۴/۳۳



جدول ۳: میانگین و انحراف معیار متغیرهای طول دوره زناشویی، سن زن و سن ازدواج زن

تعداد فرزند	طول دوره زناشویی	سن زن	سن ازدواج زن
۰	$3/87 \pm 4/58$	$26/03 \pm 4/99$	$22/16 \pm 3/40$
۱	$9/32 \pm 5/53$	$30/47 \pm 5/61$	$21 \pm 3/79$
۲	$16/94 \pm 6/23$	$36/43 \pm 6/35$	$19/31 \pm 3/32$
۳	$23/92 \pm 5/15$	$41/40 \pm 4/52$	$17/62 \pm 2/31$
۴	$29/2 \pm 3/77$	$44 \pm 4/32$	$14/8 \pm 2/1$
۵	$26 \pm 2/83$	$42 \pm 7/07$	$16 \pm 4/24$

جدول ۴: تعیین عوامل موثر بر تعداد فرزندان با استفاده از روشهای رگرسیون پواسن معمولی و پواسن تعمیم یافته

متغیرها	پواسن معمولی		پواسن تعمیم یافته	
	برآورد	خطای استاندارد	برآورد	خطای استاندارد
مقدار ثابت	-2/0337	/487	-2/030	/503
تحصیلات زن	/016	/0513	/017	/054
سن همسر	-/0169	/007	-/017	/007
تحصیلات همسر	/024	/051	/0299	/053
وضعیت اقتصادی	-/0277	/027	-/008	/028
سن ازدواج زن	/011	/017	/005	/018
تعداد حاملگی	/2196	/077	/198	/0816
ناخواستہ				
شغل زن	-/105	/127	-/087	/129
مهندس	-/110	/219	-/105	/222
پزشک	-/355	/520	-/354	/529
معلم	/0199	/234	/0145	/245
آزاد	/047	/114	/047	/119
متوسط فاصله موالید	/017	/025	-/015	/026
متوسط شیردهی	/015	/007	/0195	/007
α			-/033	/0543
شاخص نیکویی		Log-likelihood=349/29		Log-likelihood=345/6855
برازش				



بحث و نتیجه گیری

نتایج مطالعه نشان داد که میانگین و واریانس متغیر پاسخ معادل نمی‌باشد و واریانس کمتر از میانگین است و میانگین بیشتر از نما است بنابراین داده‌ها چوله به راست هستند. در این مطالعه اکثر زنان (۱۰۹ نفر) دارای ۱ فرزند هستند و فقط ۲ نفر از آنان (۰/۶۶٪) دارای ۵ فرزند هستند. نتایج جدول فوق نشان می‌دهد که اکثر زنان (۰/۴۷٪) در گروه سنی ۲۵-۳۵ سال (۰/۴۳/۳۳٪) آنها دارای طول دوره‌ی زناشویی ۱-۱۰ سال و سن ازدواج ۰/۵۱/۳۳٪ آنها در فاصله‌ی ۲۰-۳۰ سال است. در این مطالعه بیشتر زنان دارای ۱ حاملگی ناخواسته هستند و ۲۳۶ نفر از آنها حاملگی ناخواسته ندارند و اکثر مردان (۰/۴۱/۳۳٪) در گروه سنی ۳۰-۴۰ سال هستند. همچنین معنی‌دار بودن مقدار α در مدل پواسن تعمیم یافته نشان دهنده‌ی این است که مدل پواسن معمولی بر داده‌ها به درستی برازش نمی‌گردد و مقدار کمتر آماره آزمون لگاریتم درستنمایی در مدل پواسن تعمیم یافته دلالت بر این امر دارد. بنابراین بکار بردن مدل توزیع پواسن تعمیم یافته مناسب‌تر است. همان‌طور که در جدول ۳ مشاهده می‌شود با افزایش سن زن و طول دوره زناشویی، تعداد فرزندان افزایش می‌یابد که این امری بدیهی است چرا که زنان مسن‌تر فرصت کافی برای داشتن فرزندان بیشتر را دارند. بنابراین در تحلیل دو مدل متغیر طول دوره زناشویی به عنوان متغیر زمان (offset) در مدل‌ها وارد شده است ولی از وارد کردن سن زن به دلیل همبستگی زیاد آن با طول دوره زناشویی صرف نظر گردید.

در این مطالعه مقدار α منفی است که نشانگر کم پراکندگی در داده‌هاست. آقای wang و دوستانش در سال ۱۹۹۷ کم پراکندگی را در داده‌های باروری آلمان با استفاده از مدل شمارشی تعمیم یافته پیدا کردند و به این نتیجه رسیدند که مدل پواسن تعمیم یافته نسبت به مدل پواسن استاندارد (معمولی) برای داده‌های باروری ارائه شده تناسب بیشتری دارد (۱۰). در برازش دو مدل پواسن و پواسن تعمیم یافته متغیرهای سن همسر، تعداد حاملگی ناخواسته و متوسط طول مدت شیردهی با تعداد فرزندان رابطه‌ی معنی‌دار دارند. در نتیجه این عوامل بر تعداد فرزندان موثر شدند. در این مطالعه مشخص شد که مدل رگرسیون پواسن معمولی به طور مناسبی به داده‌ها برازش نمی‌گردد و باید از مدل مناسب‌تری استفاده شود. بنابراین پیشنهاد می‌شود در صورتیکه داده‌ها شمارشی باشند و شرایط مدل پواسن برقرار نباشد از تعمیم یافته آن و یا دوجمله‌ای منفی استفاده گردد.

تشکر و قدردانی

این مقاله بر گرفته از رساله دوره کارشناسی ارشد آمار زیستی می‌باشد. از کارکنان مرکز بهداشت شهر کاشان تشکر و سپاس‌گزاری را دارم.

تضاد منافع

نویسندگان این مقاله اعلام می‌دارند که هیچ گونه تضاد منافی وجود ندارد.

References

- 1- Khalili M SA, Pornajaf A, Shohani B. Factors influencing unintended pregnancy among patients referred to health centers of Ilam. Journal of Medical Sciences. 2002-2003;12.[Persian]



- 2- Azizy A, Amirian F, Pashae T AM. Study unwanted pregnancy and its association with health-related quality of life in women Bardarshhrstan Salas (Kermanshah Province). 2006.[Persian]
- 3- Bakhshian F, Jabbari H. Effectiveness of health care Aznzn Bardard. Journal of Nursing and Midwifery. 2009:43-54. [Persian]
- 4- Abasishovazy M, Makdonald P, Chavoshi MK. View of women about reproductive behavior using qualitative methods in Yazd. Letter Social Science .2002:169-203. [Persian]
- 5- Zare NSM, Vosooghi MGH. Performance comparison of generalized Poisson regression and multilevel with Poisson in reproductive behavior of rural women in the province. . Journal of Babol University of Medical Sciences.2010 :35-40.[Persian]
- 6- Coggins BL. Comparing Models for Fitting Count Data. College of Arts and Sciences 2009.
- 7- Chou N, Steenhard D. A Flexible Count Data Regression Model Using SAS® PROC NL MIXED.250-2009.
- 8- Winkelmann R. Recent developments in count data modeling: Theory and application. Journal of Economic Surveys 1995.
- 9- Ismail N, Jemain AZ. Handling Overdispersion with Negative Binomial and Generalized Poisson Regression Models. Casualty Actuarial Society Forum 2007.
- 10-Wang W, Famoye F. Modeling household fertility decisions with generalized Poisson regression. Journal of Population Economics. 1997;10:273-83.
- 11- Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. Journal of the American Statistical Association .1993:9-25.
- 12- Marshall A. Length of stay-based patient flow models: Recent developments and future direction. Health Care Manage Science. 2005:213-20.
- 13- Melkersson M, Rooth DO. Modeling female fertility using inflated count data models. Journal of Population Economics. 2000;13:189-203.