

کاربرد مدل رگرسیون پواسنی تعمیم یافته در تحلیل داده‌های باروری زنان روستایی استان فارس

نجف زارع^۱، مهرباب صیادی^۲، الهام رضائیان فرد^۳، هاله قائم^۴

^۱ دانشجویار گروه آمار زیستی، دانشکده پزشکی، دانشگاه علوم پزشکی شیراز، ایران
^۲ مشاور و تحلیلگر آماری، کارشناس ارشد آمار زیستی، دانشگاه علوم پزشکی شیراز، ایران
^۳ کارشناس بهداشت خانواده، دانشگاه علوم پزشکی شیراز، ایران
^۴ مربی گروه اپیدمیولوژی، دانشکده بهداشت و تغذیه، دانشگاه علوم پزشکی شیراز، ایران
نویسنده رابط: مهرباب صیادی: دانشگاه علوم پزشکی شیراز، دانشکده پزشکی، گروه آمار زیستی. همراه: ۰۹۱۷۱۳۳۲۹۱۷، غایب: ۰۷۱۱-۲۳۴۷۳۱۵، پست الکترونیک: sayadi_me@yahoo.com
تاریخ دریافت: ۱۳۸۸/۴/۴؛ پذیرش: ۱۳۸۸/۱۱/۳

مقدمه و اهداف: مدل‌بندی آماری تغییرات مشاهده شده در داده‌ها را از طریق معادلات ریاضی تبیین می‌نماید. در حالتی که متغیر پاسخ گسسته باشد مدل پواسن مورد استفاده قرار می‌گیرد و در صورتی که شرایط مدل پواسن برقرار نباشد، بهتر است از تعمیم یافته آن استفاده کرد. لذا هدف از این مطالعه، تاکید و توجه به ساختار داده‌ها، معرفی مدل رگرسیون پواسنی تعمیم یافته و بکار بردن این مدل جهت برآورد ضرایب عوامل مؤثر بر تعداد فرزندان و مقایسه آن با مدل رگرسیون پواسن معمولی است. **روش کار:** در این مطالعه ضمن معرفی مدل پواسنی تعمیم یافته کاربرد آن در تحلیل داده‌های باروری بکار رفته است. این داده‌ها از یک نمونه ۱۰۱۹ نفری زنان روستایی استان فارس بصورت مقطعی و با استفاده از روش نمونه‌گیری طبقه‌بندی بدست آمد. متغیر تعداد فرزند ان زنده متولد شده یک زن به عنوان متغیر پاسخ شمارشی جهت کاربرد مدل در نظر گرفته شده است. **نتایج:** میانگین فرزندان هر زن $2/88 \pm 4/3$ بود. مقدار آماره آزمون، $\log\text{-likelihood}$ برای مدل پواسنی معمولی $1950/93$ - و برای مدل پواسنی تعمیم یافته $1946/93$ - بود. **نتیجه‌گیری:** نتایج نشان داد که داده‌ها Over Dispersion دارد. و بر اساس معیارهای انتخاب بهترین مدل، مدل پواسنی تعمیم یافته جهت تحلیل این داده‌ها مناسب است و می‌تواند ضرایب عوامل مؤثر بر تعداد فرزند را دقیق‌تر برآورد نماید. **واژگان کلیدی:** داده‌های باروری، توزیع پواسن، توزیع پواسن تعمیم یافته، تعداد فرزند

مقدمه

مدل بندی آماری یکی از روش‌های تبیین تغییرات در داده‌های مشاهده شده از طریق معادلات ریاضی است که با استفاده از آن می‌توان به چگونگی تغییرات داده‌ها پی برد. در حالتی که متغیر پاسخ (وابسته) گسسته و نامنفی است مدل پواسنی مورد استفاده قرار می‌گیرد (۱-۴).

در سال‌های اخیر مدل پواسنی برای پاسخ‌های شمارشی بطور فراوان مورد استفاده قرار گرفته است. King و Winkelmann مدل‌های شمارشی تعمیم یافته را برای توزیع‌های پواسن، دو جمله‌ای و دو جمله‌ای منفی توسعه دادند و نشان دادند که مفروضات مدل پواسن معمولی برای بعضی از داده‌های شمارشی محدودیت ایجاد می‌کند (۵، ۶). شرط اصلی استفاده از مدل پواسن، معادل بودن میانگین و واریانس متغیر پاسخ است اگر این شرط برقرار نباشد مدل توزیع پواسن تعمیم یافته مناسب خواهد بود (۷-۹).

داده‌های شمارشی معمولاً 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 تعمیم یافته مناسب‌تر است (۱۰). یکی از فرضیات مدل رگرسیون پواسن معمولی این است که احتمال اتفاق هر پیشامد در هر دوره، مستقل از هم است اما در بعضی داده‌های شمارشی وقوع اولیه یک پیشامد ممکن است باعث افزایش یا کاهش احتمال پیشامد در آینده شود. در داده‌های باروری تعداد فرزندان از یک زن به زن دیگر مستقل است. ولی ممکن است تعداد فرزندان یک زن مستقل از هم نباشد. یعنی زوجینی که در یک زمان معین این تعداد فرزند را کافی نمی‌دانند تصمیم به داشتن فرزند بعدی می‌گیرند. اگر فرض بالا برقرار نباشد استفاده از مدل پواسنی منجر به برآورد نا صحیحی از ضرایب رگرسیونی می‌گردد. توزیع‌های پواسنی تعمیم یافته در حالت‌های عملی و