مجله تخصصی اپیدمیولوژی ایران؛ ۱۳۹۳، دوره ۱۰، شماره ۱: صفحات ۲۴-۱۵.

مقاله پژوهشی

به کارگیری مدل‌های رگرسیون لجستیک ترتیبی در مطالعات کیفیت زندگی

زهرا اسدالهی[۱]، پیمان جعفری[۲]، محسن رضاییان[۳]

[۱] کارشناس ارشد و مربی آمار زیستی، گروه پزشکی اجتماعی، دانشکده پزشکی، دانشگاه علوم پزشکی رفسنجان، رفسنجان، ایران

[۲] دانشیار آمار زیستی، گروه آمار زیستی، دانشکده پزشکی، دانشگاه علوم پزشکی شیراز، شیراز، ایران

[۳] استاد اپیدمیولوژی، گروه پزشکی اجتماعی، دانشکده پزشکی، دانشگاه علوم پزشکی رفسنجان، رفسنجان، ایران

نویسنده رابط: زهرا اسداللهی، نشانی: رفسنجان، میدان انقلاب، دانشکده پزشکی،گروه پزشکی اجتماعی، تلفن: ۵۲۳۴۰۰۳- ۰۳۹۱، پست الکترونیک : asadollahi.zahra@gmail.com

تاریخ دریافت: ۹۱/۱۱/۱؛ پذیرش: ۹۲/۷/۶

مقدمه و اهداف: با توجه به افزایش گرایش به سنجش کیفیت زندگی در سال‌های اخیر و حجم گسترده پرسشنامه‌های کیفیت زندگی، تعیین روش مناسب به منظور تجزیه و تحلیل داده‌های مربوط به این مطالعات ضروری به نظر می‌رسد. هدف مطالعه حاضر، معرفی مدل‌های رگرسیون لجستیک ترتیبی به منظور تجزیه و تحلیل داده‌های کیفیت زندگی می‌باشد.

روش کار: داده‌های این مطالعه مقطعی برگرفته از اطلاعات مطالعه‌ای در زمینه بررسی کیفیت زندگی ۹۳۸ دانش آموز انجام شده است. برای تحلیل داده‌ها نتایج دو مدل رگرسیون لجستیک دوحالتی و مدل‌های رگرسیون لجستیک ترتیبی با هم مقایسه شده است.

نتایج: نتایج آزمون نیکویی برازش نشان داد که هر سه مدل به خوبی برازش شده بود. بر اساس برازش مدل‌های رگرسیون لجستیک ترتیبی سه متغیر از متغیرهای توضیحی از نظر آماری با پاسخ مرتبط بودند در حالیکه بر اساس برازش مدل رگرسیون لجستیک دوحالتی که پس از ترکیب دو طبقهٔ متغیر پاسخ تنها دو متغیر معنی‌دار بود. بنابراین باید تا حد امکان از ترکیب کردن طبقه‌های متغیر پاسخ خودداری نمود، زیرا این امر منجر به از دست رفتن اطلاعات می‌شود.

نتیجه‌گیری: به طور کلی می‌توان گفت برای تجزیه و تحلیل داده‌های کیفیت زندگی به دلیل ماهیت متغیر پاسخ، مدل‌های رگرسیون لجستیک ترتیبی با توجه به برآورد پارامترهای کمتر و سهولت تفسیر نتایج، مناسب می‌باشند.

واژگان کلیدی: کیفیت زندگی، مدل‌های رگرسیون لجستیک ترتیبی، مدل رگرسیون لجستیک Binary

مقدمه

سال‌های متمادی کوشش سازمان‌های مرتبط با سلامت برای افزایش طول عمر متمرکز بوده است. پس از موفقیت در این زمینه، در کنار افزایش کمی طول عمر، کیفیت زندگی مطرح شد. زیرا بسیاری از افراد مبتلا به بیماری‌های مزمن، امروزه بیش از گذشته عمر می‌کنند و با بهبود روش‌های تشخیصی و درمانی بیماری‌ها، به ویژه در مورد بیماری‌های مزمن، افزایش طول عمر نه تنها در جمعیت سالم که در بیماران نیز رخ داده است (۱). پس از این پیشرفت چشم‌گیر، نقش مرگ‌ومیر در جمعیت‌ها به عنوان یک شاخص توسعه کم‌رنگ‌تر شد و امروزه به افزایش کیفی طول عمر یا کیفیت زندگی نیز توجه می‌شود. اندازه‌گیری کیفیت زندگی و توجه به آن در کنار درمان و مراقبت‌های سلامت، اهمیت جنبه‌های اجتماعی و روانی را در کنار جنبه‌های جسمانی سلامت مطرح کرده است (۲).

با توجه به اهمیت کیفیت زندگی که از جنبه‌های مهم سلامت است، پزشکان و محققان تلاش کرده‌اند که آن‌را به یک اندازه کمی تبدیل کنند، تا مقایسه کیفیت زندگی جمعیت‌های مختلف امکان‌پذیر شود (۳). بدین منظور پرسش‌نامه‌های بسیاری طراحی شده است، اما پرسش‌نامه‌های کیفیت زندگی سازمان جهانی بهداشت بیش‌تر مورد توجه هستند. در این میان، پرسش‌نامه کیفیت زندگی سازمان بهداشت جهانی[۱] برای سنجش کلی کیفیت زندگی (سلامت جسمانی، سلامت روانی، روابط اجتماعی و سلامت محیط) و پرسش‌نامه[۲] عمومی برای سنجش کیفیت زندگی وابسته به سلامت بیش‌ترین کاربرد را دارند. اما آن‌چه که به تازگی در حوزه علوم پزشکی مورد اقبال واقع شده است، بررسی کیفیت زندگی در گروه‌های خاصی از بیماران می‌باشد (۴). نکته دیگری که از جنبه آماری باید مورد توجه قرار گیرد، ماهیت پاسخ سؤالات در پرسش‌نامه‌های کیفیت زندگی است (۵). در مطالعات کیفیت زندگی، متغیر پاسخ بر اساس یک مقیاس ترتیبی مانند

[۱] World Health Organization Quality Of Life (WHOQOL)

[۲] Short-form health survey with only 36 questions (SF-36)

[Downloaded from irje.tums.ac.ir on 2024-08-01]

(ضعیف، متوسط، خوب، خیلی خوب و عالی) اندازه‌گیری می‌شود، اما در هنگام تجزیه و تحلیل معمولاً این متغیر رتبه‌ای به یک متغیر اسمی تبدیل شده و از رگرسیون لجستیک دوتایی[۱] استفاده می‌شود. اگر چه چنین روش‌هایی نادرست نیستند، اما این امر منجر به از دست رفتن اطلاعات به علت نادیده گرفتن برخی از طبقه‌های پاسخ می‌شود و توان آماری به طور قابل توجهی کاهش می‌یابد (۶).

به طور کلی، در بررسی تأثیر مجموعه‌ای از متغیرهای مستقل بر متغیر پاسخ، انتخاب مدل تا حد زیادی با توجه به مقیاس پاسخ تعیین می‌شود. بنابراین با توجه به ماهیت پاسخ سؤالات در پرسش‌نامه‌های کیفیت زندگی که بر اساس یک مقیاس رتبه‌ای اندازه‌گیری می‌شود، تعیین روش مناسب به منظور تجزیه و تحلیل داده‌های مربوط به این مطالعات ضروری به نظر می‌رسد (۶). روش‌های آماری مانند مدل‌های رگرسیون لجستیک ترتیبی[۲] برای تجزیه و تحلیل پاسخ‌های ترتیبی پیشنهاد شده است (۱۵-۱۴،۳)، اما استفاده از آن‌ها نه تنها، در مطالعات کیفیت زندگی، بلکه در تمامی رشته‌های اپیدمیولوژی و علوم پزشکی و بهداشتی بسیار کم مورد استفاده قرار گرفته است، اما بر اساس مطالعات انجام شده در سایر کشورها، ثابت شده مدل‌های رگرسیون لجستیک ترتیبی برای تجزیه و تحلیل داده‌های کیفیت زندگی مناسب هستند (۱۵). در مطالعات گذشته مربوط به کیفیت زندگی که در ایران و بیشتر کشورها صورت گرفته است از تجزیه و تحلیل اطلاعات به کمک مدل مرسوم رگرسیون لجستیک دوتایی استفاده شده است (۲۰-۱۶). هدف پژوهش حاضر، معرفی مدل‌های رگرسیون لجستیک ترتیبی به منظور تجزیه و تحلیل داده‌های کیفیت زندگی با ارایه مثال کاربردی می‌باشد. هدف از به کار بردن مدل‌های رگرسیون لجستیک ترتیبی، کشف مدل‌های با پارامتر کم‌تر، تفسیر نتایج بهتر و به حداکثر رساندن کیفیت پیش‌بینی متغیر وابسته، در روابط بین متغیر پاسخ و مجموعه متغیرهای توضیحی است.

روش کار

داده‌های این مطالعه مقطعی برگرفته از اطلاعات مطالعه‌ای در زمینه بررسی کیفیت زندگی ۹۳۸ دانش‌آموز در شهر شیراز در سال ۱۳۸۸ با استفاده از پرسش‌نامه اختصاصی (pedsQL)[۳] انجام شد. جزییات بیش‌تر در خصوص روش مطالعه، روش‌های نمونه‌گیری در مقاله منتشر شده قابل دسترسی است (۲۱). انجام آنالیزهای بیش‌تر روی این داده‌ها با کسب اجازه از نویسندگان محترم این مقاله بوده است. متغیرهای مختلفی مانند جنسیت، سطح تحصیلات والدین، مقطع تحصیلی کودکان، سابقه‌ی بیماری، تعداد فرزندان، سن والدین و کودکان ثبت شده بود. ابتدا همه‌ی متغیرها در برازش مدل‌ها وارد شدند، اما تنها متغیرهای سطح تحصیلات مادر، مقطع تحصیلی و سابقه‌ی بیماری به‌طور معنی‌داری با سطح کیفیت زندگی مرتبط بودند. بنابراین تنها این متغیرها، به‌عنوان متغیرهای توضیحی در برازش مدل‌ها وارد شدند. متغیر سطح تحصیلات مادر به صورت سه سطح تحصیلات ابتدایی تا راهنمایی، دبیرستان تا دیپلم و فوق‌دیپلم و بالاتر ثبت شده بود. متغیر مقطع تحصیلی به صورت سه مقطع مدرسه ابتدایی، راهنمایی و دبیرستان ثبت شده بود. دانش‌آموزان دارای سابقه‌ی بیماری و عدم سابقه‌ی بیماری مشخص شده بودند.

هم‌چنین پرسش‌نامه (pedsQL) شامل ۴۶ سؤال بود که ۴ بعد متفاوت از سلامت شامل عملکرد اجتماعی، عاطفی، جسمانی و عملکرد در مدرسه دانش‌آموزان را اندازه گرفته بود. نمره‌ی کلی کیفیت زندگی دانش‌آموزان که از میانگین نمرات خام ۴۶ سؤال پرسش‌نامه محاسبه شده بود مقداری بین ۵-۱ داشت و نمره‌ی بیش‌تر نشان دهنده‌ی کیفیت زندگی بهتر بود. در این تحقیق، روش‌های آمار توصیفی برای تعیین میانگین و انحراف معیار متغیرهای کمی و فراوانی (درصد) متغیر های کیفی به کار گرفته شد. برای تحلیل داده‌ها نیز از رگرسیون لجستیک دوتایی و مدل‌های رگرسیون لجستیک ترتیبی استفاده شد. تمامی مدل‌ها به کمک نرم‌افزار SAS نسخه ی ۹/۲ به داده‌ها برازش داده شد. مقدار احتمال کم‌تر از ۰/۰۵ معنی‌دار در نظر گرفته شد.

در زیر مدل‌های مختلف رگرسیون لجستیک ترتیبی و کاربرد هر کدام از آن‌ها بیان شده است:

به طور کلی مدل‌های رگرسیون لجستیک ترتیبی بر اساس احتمال‌های تجمعی[۴] به دست می‌آیند (۲۲).

مدل شانس متناسب[۵]

مدل شانس متناسب، معروف به مدل لوجیت تجمعی می‌باشد. متغیر پاسخ در اصل پیوسته است و سپس طبقه‌بندی می‌شود (۷،۶). این مدل احتمال پاسخ کم‌تر یا مساوی طبقه داده شده را با احتمال پاسخ بزرگ‌تر از این طبقه مقایسه می‌کند (۱۵).

[۱] Binary
[۲] Ordinal logistic regression models
[۳] Pediatric Quality Of Life
[۴] Cumulative Logit
[۵] Proportional odds model – POM

[Downloaded from irje.tums.ac.ir on 2024-08-01]

$$\lambda_j(\underline{x}) = \ln\left\{\frac{\Pr(Y=1\mid\underline{x})+\ldots+\Pr(Y=j\mid\underline{x})}{\Pr(Y=j+1\mid\underline{x})+\ldots+\Pr(Y=k\mid\underline{x})}\right\} = \ln\left\{\frac{\sum_{1}^{j}\Pr(Y=j\mid\underline{x})}{\sum\Pr(Y=j\mid\underline{x})}\right\}$$

$$\lambda_j(\underline{x}) = \alpha_j + (\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p), j = 1,\ldots,k-1$$

در این مدل بردار ضرایب رگرسیون β، به j بستگی ندارد و این بدین معنی است که این مدل فرض می‌کند رابطه بین xi و y مستقل از طبقه‌ی متغیر پاسخ می‌باشد (۷). این فرض یکسانی لگاریتم OR در میان k طبقه پاسخ را، فرض شانس متناسب می‌نامند (۱۳). بنابراین در روند ساخت مدل، باید اعتبار این فرض را بررسی کرد (۱۰). این مدل هم‌چنین خاصیت عدم تغییر در انتخاب طبقه‌های پاسخ را نشان می‌دهد. این ویژگی بدین معنی است که اگر مدل شانس متناسب برای هر مقیاس متغیر پاسخ برقرار باشد، آن‌گاه مدل بازای ترکیب رسته‌های پاسخ نیز برقرار است (۱۵).

به منظور برازش مدل شانس متناسب در این پژوهش، نمره کلی کیفیت زندگی دانش‌آموزان که از میانگین نمرات خام ۴۶ سؤال پرسش‌نامه محاسبه شده بود و مقداری بین ۵-۱ داشت؛ در یک مقیاس ترتیبی جدید این نمرات به سه طبقه به صورت زیر تقسیم‌بندی شده‌اند: نمره‌های بین ۲/۵-۱ نشان دهنده کیفیت زندگی پایین (کد ۱) می‌باشد، نمره‌های بین ۴-۲/۵ بیان کننده کیفیت زندگی متوسط (کد ۲) و نمره‌های بین ۵-۴ بیان کننده کیفیت زندگی خوب (کد ۳) است. این تقسیم‌بندی بر اساس نظر پژوهش‌گر انجام شده است.

مدل لوجیت پیوسته- نسبت[1]

فین برگ مدل لوجیت پیوسته- نسبت را با طبقه‌های پاسخ ترتیبی، به صورت دنباله‌ای در نظر گرفته است. در این نوع لوجیت پیوسته- نسبت، هر رسته با گروه‌بندی رسته‌های بالاتر از سطوح مقیاس پاسخ مقایسه می‌شود (۲۳). مدل به صورت زیر بیان شده است:

$$\lambda_j(\underline{x}) = \ln\left\{\frac{\Pr(Y=j\mid\underline{x})}{\Pr(Y=j+1\mid\underline{x})+\ldots+\Pr(Y=k\mid\underline{x})}\right\} = \ln\left\{\frac{\Pr(Y=j\mid\underline{x})}{\sum_{j+1}^{k}\Pr(Y=j\mid\underline{x})}\right\}$$

$$\lambda_j(\underline{x}) = \alpha_j + (\beta_{j1}x_1 + \beta_{j2}x_2 + \ldots + \beta_{jp}x_p), j = 1,\ldots,k-1$$

برای هر رسته (j=1,..k)، این مدل دارای ثابت‌های مختلف و ضرایب متفاوتی برای هر مقایسه است. بیش‌ترین کاربرد این مدل زمانی است که رسته‌های خاصی از متغیر پاسخ مورد توجه هستند (۱۵،۶). البته در برخی منابع بیش‌ترین کاربرد این مدل را مربوط به زمانی می‌دانند که طبقه‌های متغیر پاسخ به‌صورت طبقه‌های متوالی باشد. بدین معنی که برای رسیدن به طبقه‌های بالاتر حتماً باید طبقه‌های پایین‌تر پشت سر گذاشته شود (۲۲،۲۴). خاصیت عدم تغییر در انتخاب رسته‌های پاسخ برای این مدل برقرار نیست (۱۱).

مدل شانس متناسب جزیی[2]

مدل شانس متناسب جزیی یک تعمیم مدل شانس متناسب است. در شرایطی، تنها برای برخی از متغیرهای موجود در مدل، فرض شانس متناسب برقرار می‌باشد (۲۵). همان‌طور که در فرمول مدل شانس متناسب جزیی نشان داده شده است، بدون از دست دادن اصل کلی، فرض می‌کنیم که برای q متغیر اول فرض شانس متناسب برقرار نباشد (۷). برای یک متغیر که در آن فرض شانس متناسب برقرار نیست به‌عنوان مثال x۱ ضریب آن توسط ضریب γj۱ افزایش می‌یابد، که اثر مربوط با هر لوجیت است، بنابراین ضریب متغیر x۱ (β۱ + γj۱) است (۱۵).

$$\lambda_j(\underline{x}) = \ln\left\{\frac{\Pr(Y=1\mid\underline{X})+\ldots+\Pr(Y=j\mid\underline{x})}{\Pr(Y=j+1\mid\underline{x})+\ldots+\Pr(Y=k\mid\underline{x})}\right\}$$

$$\lambda_j(\underline{x}) = \alpha_j + [(\beta_1+\gamma_{j1})x_1 + \ldots + (\beta_q+\gamma_{jq})x_q + (\beta_{q+1}x_{q+1}) + \ldots + (\beta_p x_p)], j = 1,\ldots,k-1$$

روش انجام کار به این صورت است که برای وارد کردن داده‌ها در نرم‌افزار SAS ابتدا با استفاده از جداول توافقی SPSS فراوانی‌ها بر اساس متغیرهای توضیحی مورد نظر و طبقه‌ی کیفیت زندگی به‌عنوان متغیر پاسخ ترتیبی محاسبه و سپس به صورت ماتریسی وارد نرم‌افزار SAS شده است.

ابتدا فرض شانس متناسب برای هر متغیر مستقل با استفاده از آزمون Score بررسی شد. مناسبت هر مدل با استفاده از آماره‌های پیرسون (QP) و نسبت درست‌نمایی (QL) مورد بررسی قرار گرفت. هم‌چنین برآورد پارامترها، آزمون معنی‌داری متغیرهای توضیحی و نسبت‌های شانس نیز بررسی شد.

یافته‌ها

از ۹۳۸ دانش‌آموز مورد بررسی، ۵۹۴ دانش‌آموز (۶۳/۳ درصد)، در طبقه کیفیت زندگی متوسط قرار داشتند.

برازش مدل شانس متناسب

[1] Continuation Ratio Logistic Model – CRM

[2] Partial Proportional Odds Model - PPOM

[Downloaded from irje.tums.ac.ir on 2024-08-01]

ابتدا فرض شانس متناسب بررسی شد و با توجه به نتایج آزمون Score در جدول شماره ۱ فرض شانس متناسب برای همه متغیرها پذیرفته شده است و نیازی به برازش مدل شانس متناسب جزئی وجود نداشت. جدول شماره ۱ هم‌چنین، برآورد پارامترها، آزمون معنی‌داری متغیرهای توضیحی و نسبت‌های شانس را نیز نشان می‌دهد. با توجه به نتایج آزمون والد و مقایسه‌ی آن با سطح معنی‌داری ۰/۰۵ نشان می‌دهد که همه‌ی متغیرها معنی‌دار می‌باشد. برآوردهای نسبت شانس محاسبه شده برای هر یک از متغیرهای معنی‌دار را می‌توان به‌صورت زیر تفسیر کرد:

شانس قرار گرفتن در طبقه‌های پایین‌تر کیفیت زندگی، در بچه‌هایی که دارای سابقه بیماری هستند، نسبت به سایر بچه‌ها ۱/۷۶ برابر می‌باشد. هم‌چنین شانس قرار گرفتن در طبقه‌های پایین‌تر کیفیت زندگی، در بچه‌های دبیرستانی نسبت به کودکان ابتدایی ۲/۱۳ برابر می‌باشد و بر اساس متغیر سطح تحصیلات مادر شانس قرار گرفتن در طبقه‌های پایین‌تر کیفیت زندگی، در بچه‌هایی که سطح تحصیلات مادرشان ابتدایی و یا راهنمایی است تقریباً ۱/۸۸ برابر بچه‌هایی است که تحصیلات مادرشان فوق‌دیپلم و بالاتر می‌باشد. هم‌چنین این شانس برای بچه‌هایی که سطح تحصیلات مادرشان دبیرستان و یا دیپلم است ۱/۵۱ برابر بچه‌هایی است که تحصیلات مادر آن‌ها فوق‌دیپلم و بالاتر می‌باشد. آزمون نیکویی برازش مدل شانس متناسب در جدول ۲ نشان داده شده است.

برازش مدل رگرسیون لجستیک دوتایی

به منظور برازش مدل رگرسیون لجستیک دوتایی، ابتدا متغیر پاسخ به صورت یک متغیر پاسخ دو حالتی تبدیل شد. بدین منظور دو طبقه‌ی کیفیت زندگی متوسط و خوب با هم ترکیب شده است. بنابراین در این مدل متغیر پاسخ دارای دو طبقه‌ی کیفیت زندگی پایین و خوب بود. برای برازش این مدل در نرم‌افزار SAS از PROC LOGISTIC استفاده شده است. جدول شماره ۳، برآورد پارامترها، آزمون معنی‌داری متغیرهای توضیحی و نسبت‌های شانس را نشان می‌دهد. با توجه به نتایج آزمون والد و مقایسه‌ی آن با سطح معنی‌داری ۰/۰۵ نشان می‌دهد که بر اساس برازش مدل رگرسیون لجستیک دوتایی، تنها متغیرهای سابقه‌ی بیماری و مقطع تحصیلی معنی‌دار می‌باشد. برآوردهای نسبت شانس محاسبه شده برای هر یک از متغیرهای معنی‌دار را می‌توان به صورت زیر تفسیر کرد: شانس قرار گرفتن در طبقه‌ی پایین کیفیت زندگی، برای بچه‌های دارای سابقه بیماری، نسبت به سایر بچه‌ها ۱/۶۴ برابر می‌باشد. شانس قرار گرفتن در طبقه‌ی پایین کیفیت زندگی، برای بچه‌های راهنمایی، نسبت به کودکان ابتدایی تقریباً ۱/۲۶ برابر می‌باشد. هم‌چنین این شانس برای بچه‌های دبیرستانی نسبت به کودکان ابتدایی ۲/۰۳ برابر می‌باشد. بر اساس برازش این مدل متغیر سطح تحصیلات مادر معنی‌دار نمی‌باشد.

برازش مدل لوجیت‌های پیوسته- نسبت

برای برازش این مدل به داده‌های کیفیت زندگی دانش‌آموزان، همانند مدل شانس متناسب، متغیر پاسخ به سه طبقه‌ی کیفیت زندگی پایین، متوسط و خوب تقسیم شده است. این مدل برای هر طبقه‌ی متغیر پاسخ، ضرایب متفاوتی را در نظر می‌گیرد. برای برازش این مدل در نرم‌افزار SAS از Proc Catmod استفاده شده است. مدل لوجیت پیوسته نسبت برای مدل‌سازی، احتمال قرار گرفتن در طبقهٔ کیفیت زندگی پایین نسبت به طبقه‌های بالاتر از $\log(p_1/p_2+p_3)$ استفاده می‌کند. جدول شماره ۴ برآورد پارامترها، آزمون معنی‌داری متغیرهای توضیحی و نسبت‌های شانس را بر اساس برازش اولین لوجیت پیوسته- نسبت نشان می‌دهد. با توجه به نتایج آزمون والد، بر اساس برازش اولین لوجیت پیوسته- نسبت، متغیرهای سابقه‌ی بیماری، مقطع تحصیلی و سطح تحصیلات مادر معنی‌دار می‌باشد. برآوردهای نسبت شانس محاسبه شده برای هر یک از متغیرهای معنی‌دار را می‌توان به‌صورت زیر تفسیر کرد:

شانس قرار گرفتن در طبقه‌ی کیفیت زندگی پایین، برای بچه‌های دارای سابقه بیماری نسبت به سایر بچه‌ها تقریباً ۱/۸۴ برابر می‌باشد. شانس قرار گرفتن در طبقه‌ی کیفیت زندگی پایین، برای بچه‌های دبیرستانی، نسبت به کودکان ابتدایی تقریباً ۲/۲۳ برابر می‌باشد. شانس قرار گرفتن در طبقه‌ی کیفیت زندگی پایین، برای بچه‌هایی که سطح تحصیلات مادرشان ابتدایی و یا راهنمایی است؛ نسبت به بچه‌هایی که تحصیلات مادرشان فوق‌دیپلم و بالاتر است، تقریباً ۲/۸۲ برابر می‌باشد.

لوجیت‌های پیوسته- نسبت برای مدل‌سازی، احتمال قرار گرفتن در طبقه‌ی کیفیت زندگی متوسط نسبت به طبقه‌ی کیفیت زندگی خوب، با حذف طبقه‌ی کیفیت زندگی پایین از $\log(p_2/p_3)$ استفاده می‌کند. جدول شماره ۵ برآورد پارامترها، آزمون معنی‌داری متغیرهای توضیحی و نسبت‌های شانس را بر اساس برازش دومین لوجیت پیوسته- نسبت نشان می‌دهد. با توجه به نتایج آزمون والد و مقایسه‌ی آن با سطح معنی‌داری

[Downloaded from irje.tums.ac.ir on 2024-08-01]

۰/۰۵ نشان می‌دهد که بر اساس برازش دومین لوجیت پیوسته- نسبت، تنها متغیر مقطع تحصیلی معنی‌دار می باشد.

برآوردهای نسبت شانس محاسبه شده برای هر یک از سطوح متغیر معنی‌دار را می‌توان به‌صورت زیر تفسیر کرد:

شانس قرار گرفتن در طبقه‌ی کیفیت زندگی متوسط، برای بچه‌های دبیرستانی نسبت به کودکان ابتدایی تقریباً ۱/۸۳ برابر می‌باشد.

جدول شماره ۶ مقدار QL اولین و دومین لوجیت به ترتیب برابر ۶۵/۶۹ با ۳۹ درجه آزادی و ۳۴/۷۴ با ۳۷ درجه آزادی را نشان می‌دهد. برازش مجموع آن‌ها دارای آماره نیکویی برازش QL=۱۰۰/۴۳ با درجه آزادی ۷۶ می‌باشد که با مقایسه‌ی آن با توزیع مربع کای، کفایت مدل تأیید می‌شود.

جدول شماره ۱- بر آورد پارامترها و نسبت‌های شانس بر اساس برازش مدل شانس متناسب

متغیرها	سطوح	برآوردها (انحراف معیار)	آزمون والد (احتمال معنی‌داری)	نسبت شانس (فاصله اطمینان)	آزمون Score (احتمال معنی‌داری)
سابقه بیماری	عدم سابقه بیماری	-	-	۱	۰/۶۸ (P = ۰/۴۰)
	سابقه بیماری	۰/۵۷ (۰/۲۰)	۷/۶۹ (p=۰/۰۰۵)	۱/۷۶ (۰/۳۷-۰/۸۴)	
مقطع تحصیلی	ابتدایی	-	-	۱	۱/۱۶ (P = ۰/۵۵)
	راهنمایی	۰/۲۴ (۰/۱۸)	۱/۸۳ (p=۰/۱۷)	۱/۲۷ (۰/۸۹-۱/۸۱)	
	دبیرستان	۰/۷۵ (۰/۱۷)	۱۹/۷۳ (p=۰/۰۰۰۱)	۲/۱۳ (۱/۵۲-۲/۹۸)	
سطح تحصیلات مادر	ابتدایی تا راهنمایی	۰/۶۱ (۰/۲۴)	۶/۵۹ (p=۰/۰۱)	۱/۸۸ (۱/۱۵-۲/۹۷)	۰/۵۵ (P = ۰/۴)
	دبیرستان تا دیپلم	۰/۴۱ (۰/۲۰)	۳/۹۸ (p=۰/۰۴)	۱/۵۱ (۱/۰۰۸-۲/۲۶)	
	فوق‌دیپلم و بالاتر	-	-	۱	

بر اساس نتایج آزمون نیکویی برازش، کفایت مدل شانس متناسب بر اساس معیار دیویانس (Deviance) تأیید می‌شود (p =۰/۰۶).

جدول شماره ۲- آزمون نیکویی برازش مدل شانس متناسب

معیار	مقدار (value)	درجه آزادی (DF)	$value/DF$	Pr > ChiSq
Deviance	۱۰۶/۰۷۲۹	۸۵	۱/۲۴۷۹	۰/۰۶۰۶

[Downloaded from irje.tums.ac.ir on 2024-08-01]

جدول شماره ۳- بر آورد پارامترها و نسبت‌های شانس بر اساس برازش مدل رگرسیون لجستیک

متغیر ها	سطوح	برآوردها (انحراف معیار)	آزمون والد (احتمال معنی‌داری)	نسبت شانس (فاصله اطمینان)
سابقه بیماری	عدم سابقه بیماری	-	-	۱
	سابقه بیماری	۰/۵۰ (۰/۲۷)	۳/۳۲ (۰/۰۶)	۱/۶۴ (۰/۳۴-۱/۰۳)
مقطع تحصیلی	ابتدایی	-	-	۱
	راهنمایی	-۰/۷۱ (۰/۲۰)	۱۲/۲۰ (۰/۰۰۱)	۱/۲۶ (۰/۸۴-۱/۸۹)
	دبیرستان	-۰/۴۸ (۰/۲۰)	۵/۷۴ (۰/۰۱)	۲/۰۳ (۱/۳۶-۳/۰۴)
سطح تحصیلات مادر	ابتدایی تا راهنمایی	۰/۴۷ (۰/۲۹)	۲/۶۰ (۰/۱۰)	۱/۶۰ (۰/۹۰-۲/۸۶)
	دبیرستان تا دیپلم	۰/۲۸ (۰/۲۴)	۱/۳۷ (۰/۲۴)	۱/۳۳ (۰/۸۲-۲/۱۶)
	فوق‌دیپلم و بالاتر	-	-	۱

بر اساس نتایج آزمون نیکویی برازش، کفایت مدل رگرسیون لجستیک دوتایی بر اساس معیار دیویانس (Deviance) تأیید می‌شود (P =۰/۶۲۱).

جدول شماره ٤ - برآورد پارامترها و نسبت‌های شانس بر اساس برازش اولین لوجیت پیوسته- نسبت

متغیرها	سطوح	برآوردها (انحراف معیار)	آزمون والد (احتمال معنی‌داری)	نسبت شانس (فاصله اطمینان)
سابقه بیماری	عدم سابقه بیماری	-	-	۱
	سابقه بیماری	۰/۶۱ (۰/۲۵)	۵/۹۸ (۰/۰۱)	۱/۸۴ (۱/۱۲-۳/۰۱)
مقطع تحصیلی	ابتدایی	-	-	۱
	راهنمایی	-۰/۰۸ (۰/۱۴)	۰/۳۰ (۰/۵۸)	۱/۳۲ (۰/۷۵-۲/۳۰)
	دبیرستان	۰/۴۴ (۰/۱۲)	۱۲/۱۸ (۰/۰۰۱)	۲/۲۳ (۱/۳۷-۳/۶۵)
سطح تحصیلات مادر	ابتدایی تا راهنمایی	۰/۴۲ (۰/۱۷)	۶/۰۳ (۰/۰۱)	۲/۸۲ (۱/۳۷-۵/۸۰)
	دبیرستان تا دیپلم	۰/۱۸ (۰/۱۵)	۱/۴۹ (۰/۲۲)	۲/۲۲ (۱/۱۴-۴/۳۱)
	فوق‌دیپلم و بالاتر	-	-	۱

[Downloaded from irje.tums.ac.ir on 2024-08-01]

جدول شماره ۵ -بر آورد پارامترها و نسبت‌های شانس بر اساس برازش دومین لوجیت پیوسته- نسبت

متغیرها	متغیرها	برآوردها (انحراف معیار)	آزمون والد (احتمال معنی‌داری)	نسبت شانس (فاصله اطمینان)
سابقه بیماری	عدم سابقه بیماری	-	-	۱
	سابقه بیماری	۰/۳۸ (۰/۲۸)	۱/۷۷ (۰/۱۸)	۱/۴۶ (۰/۸۳-۲/۵۷)
مقطع تحصیلی	ابتدایی	-	-	۱
	راهنمایی	-۰/۰۶ (۰/۱۲)	۰/۲۴ (۰/۶۱)	۱/۲۳ (۰/۸۱-۱/۸۸)
	دبیرستان	۰/۳۳ (۰/۱۱)	۷/۸۷ (۰/۰۰۵)	۱/۸۳ (۱/۲۱-۲/۷۸)
سطح تحصیلات مادر	ابتدایی تا راهنمایی	۰/۱۴ (۰/۱۵)	۰/۸۵ (۰/۳۵)	۱/۳۶ (۰/۷۶-۲/۴۵)
	دبیرستان تا دیپلم	۰/۰۲ (۰/۱۲)	۰/۰۳ (۰/۸۵)	۱/۲۱ (۰/۷۳-۱/۹۸)
	فوق‌دیپلم و بالاتر	-	-	۱

جدول شماره ٦- خروجی آزمون نیکویی برازش اولین و دومین لوجیت پیوسته- نسبت

معیار	مقدار (value)	درجه آزادی (DF)	$value/DF$	Pr > ChiSq
Deviance اولین لوجیت	۶۵/۹۶۰۳	۳۹	۱/۶۸۴۴	۰/۰۰۴۸
Deviance دومین لوجیت	۳۴/۷۴۰۳	۳۷	۰/۹۳۸۹	۰/۵۷۵۴

بحث

هدف اصلی این مطالعه، بررسی روش‌های مناسب تجزیه و تحلیل داده‌های کیفیت زندگی می‌باشد. با توجه به افزایش گرایش به سنجش کیفیت زندگی در سال‌های اخیر و حجم گسترده پرسش‌نامه‌های کیفیت زندگی، این نکته مشخص می‌شود که طراحی این پرسش‌نامه‌ها به دانش جدیدی در مطالعات بالینی تبدیل شده است، به نحوی که علوم متفاوتی از جمله ریاضیات و آمار به‌عنوان بخش اصلی این مطالعات معرفی شده است (۲۶). با وجود فراوانی این مطالعات، هنوز در مورد مفهوم، سنجش و تجزیه و تحلیل داده‌های مربوط به آن انسجامی وجود ندارد (۱۵).

مطالعه حاضر بیش‌تر به بررسی و مقایسه نتایج مدل‌های رگرسیون لجستیک ترتیبی با نتایج برازش مدل رگرسیون لجستیک دوتایی و تفسیر نتایج آن، در تجزیه و تحلیل داده‌های مربوط به کیفیت زندگی پرداخته است.

مقایسه‌ی بین نتایج برازش مدل‌های شانس متناسب و مدل رگرسیون لجستیک دوتایی و مدل پیوسته- نسبت بر داده‌های کیفیت زندگی دانش‌آموزان انجام شد. نتایج آزمون نیکویی برازش نشان داد که هر سه مدل به خوبی برازش شده بود. بر اساس برازش مدل‌های شانس متناسب و اولین لوجیت پیوسته- نسبت سه متغیر سابقه‌ی بیماری، مقطع تحصیلی و سطح تحصیلات مادر از نظر آماری با پاسخ مرتبط بودند. بر اساس برازش دومین لوجیت پیوسته- نسبت، تنها متغیر مقطع تحصیلی معنی‌دار بود، که این نتیجه با توجه به اختلاف کم و نزدیکی بین دو طبقه‌ی کیفیت زندگی متوسط و خوب می‌تواند توضیح داده شود. هم‌چنین در این پژوهش از مدل رگرسیون لجستیک دوتایی برای بررسی عوامل مؤثر بر کیفیت زندگی دانش‌آموزان استفاده شد. بر اساس نتایج حاصل از این مدل روی پاسخ دو حالتی

[Downloaded from irje.tums.ac.ir on 2024-08-01]

(کیفیت زندگی پایین و خوب)، تنها ارتباط بین متغیرهای سابقه بیماری و مقطع تحصیلی با کیفیت زندگی تأیید شد و ارتباط بین متغیر سطح تحصیلات مادر با کیفیت زندگی در این مدل رد شد و این نقطه ضعف این مدل محسوب می‌شود. بنابراین باید تا حد امکان از ترکیب کردن طبقه‌های متغیر پاسخ خودداری نمود، زیرا این امر منجر به از دست رفتن اطلاعات به علت نادیده گرفتن برخی از طبقه‌های پاسخ می‌شود؛ همان‌طور که در این مثال پس از ترکیب دو طبقه‌ی کیفیت زندگی پایین و متوسط به داده‌ها، متغیر سطح تحصیلات مادر معنی‌دار نشد.

بنابراین با توجه به نتایج فوق در مورد انتخاب مدل مناسب از بین مدل‌های رگرسیون لجستیک ترتیبی و رگرسیون لجستیک دوتایی می‌توان نتیجه گرفت که مدل‌های رگرسیون لجستیک ترتیبی برآوردهای قابل اطمینان‌تری را برای تجزیه و تحلیل داده‌های ترتیبی فراهم می‌کند.

در مورد انتخاب مناسب‌ترین مدل رگرسیون لجستیک ترتیبی بین دو مدل شانس متناسب و پیوسته- نسبت باید گفت که هر چند که برازش مدل پیوسته- نسبت مناسب نشان داد، اما کاربرد بیش‌تر این مدل مربوط به زمانی است که طبقه‌ی خاصی از متغیر پاسخ مورد توجه باشد و یا در مواردی است که طبقه‌های متغیر پاسخ به‌صورت متوالی باشند و برای رسیدن به طبقه‌های بالاتر متغیر پاسخ، باید حتماً طبقه‌های پایین‌تر را پشت سر گذاشت که در این مطالعه مورد ندارد و توصیه نمی‌شود.

بنابراین با توجه به کاربرد مدل شانس متناسب هنگامی که متغیر پاسخ در اصل پیوسته است و سپس طبقه‌بندی می‌شود و با توجه به این‌که در این مطالعه نمره کلی کیفیت زندگی دانش‌آموزان که از میانگین نمرات خام ۴۶ سؤال پرسش‌نامه محاسبه شده بود مقداری بین ۵-۱ داشت که در یک مقیاس ترتیبی جدید این نمرات به سه طبقه تقسیم‌بندی شدند و هم‌چنین با توجه به برقراری شرط شانس متناسب برای همه‌ی متغیرهای توضیحی موجود در مطالعه، مدل شانس متناسب به علت برآورد پارامترهای کم‌تر نسبت به سایر مدل‌های رگرسیون ترتیبی و سهولت تفسیر نتایج، به‌عنوان مناسب‌ترین مدل در این مطالعه معرفی می‌گردد که نتایج به دست آمده با مطالعات Mery Natali Silva Abreu و Ananth CV هم‌خوانی دارد (۶و۱۵).

هم‌چنین در صورت عدم برقراری شرط شانس متناسب برای برخی از متغیرهای توضیحی موجود در مطالعه، مناسب‌ترین جایگزین برای مدل شانس متناسب، مدل شانس متناسب جزیی بود. مدل شانس متناسب جزیی برای برخی از متغیرها تنها یک نسبت شانس برای همه طبقه‌های متغیر پاسخ و برای بعضی از متغیرهای توضیحی در هر طبقه‌ی متغیر پاسخ یک نسبت شانس برآورد می‌کند.

نتیجه‌گیری

به طور کلی می‌توان گفت برای تجزیه و تحلیل داده‌های کیفیت زندگی به دلیل ماهیت متغیر پاسخ، دو مدل شانس متناسب و شانس متناسب جزیی با توجه به برآورد پارامترهای کم‌تر و سهولت تفسیر نتایج، مناسب و به‌کارگیری آن‌ها برای تجزیه و تحلیل داده‌های این مطالعات توصیه می‌شود.

تشکر و قدردانی

بدین‌وسیله از خانم زهرا سلطانی که داده‌های این پژوهش را در اختیار گروه تحقیق قرار دادند؛ صمیمانه تشکر و قدردانی می‌شود. هم‌چنین از اساتید محترم دانشگاه علوم پزشکی شیراز که در مراحلی از این طرح ما را کمک نمودند، تشکر و قدردانی می‌شود.

منابع

1. Seidl EM, Zannon CM. Quality of life and health: conceptual and methodological issues. Cad Saude Publica2004 Mar-Apr; 20: 580-8.
2. Landesman S. Quality of life and personal life satisfaction: definition and measurement issues. Ment Retard1986 Jun; 24: 141-3.
3. Ciconelli R, Ferraz M, Santos W, Meinão I, Quaresma M. 1. Introdução 1.1. Contexto. Rev bras reumatol 1999; 39: 143-50.
4. Drossman DA, Patrick DL, Whitehead WE, Toner BB, Diamant NE, Hu Y, Jia H, Bangdiwala SI. Further validation of the IBS-QOL: a disease-specific quality-of-life questionnaire. Am J Gastroenterol2000 Apr; 95: 999-1007.
5. Walters SJ, Campbell MJ, Lall R. Design and analysis of trials with quality of life as an outcome: a practical guide. J Biopharm Stat 2001; 11: 155-76.
6. Ananth CV, Kleinbaum DG. Regression models for ordinal responses: a review of methods and applications. Int J Epidemiol1997 Dec; 26: 1323-33.
7. Lall R, Campbell M, Walters S, Morgan K, Co-operative MRCC. A review of ordinal regression models applied on health-related quality of life assessments. Statistical Methods in Medical Research 2002; 11: 49-67.
8. Anderson JA. Regression and ordered categorical variables. Journal of the royal statistical society Series B (Methodological) 1984: 1-30.
9. Bender R, Grouven U. Ordinal logistic regression in medical research. J R Coll Physicians Lond1997 Sep-Oct; 31: 546-51.

[Downloaded from irje.tums.ac.ir on 2024-08-01]

10. Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression. Biometrics1990: 1171-8.
11. Greenland S. Alternative models for ordinal logistic regression. Statistics in Medicine1994; 13: 1665-77.
12. Hosmer D, Lemeshow S. Interpretation of the fitted logistic regression model. Applied Logistic Regression 2nd ed New York: John Wiley & Sons, INC2000: 47-90.
13. McCullagh P. Regression models for ordinal data. Journal of the royal statistical society Series B (Methodological) 1980: 109-42.
14. McCullagh P, Nelder JA. Generalized linear models: Chapman & Hall/CRC; 1989.
15. Abreu MN, Siqueira AL, Cardoso CS, Caiaffa WT. Ordinal logistic regression models: application in quality of life studies. Cad Saude Publica2008;24 Suppl 4:s581-91.
16. Jafari H, Lahsaeizadeh S, Jafari P, Karimi M. Quality of life in thalassemia major: reliability and validity of the Persian version of the SF-36 questionnaire. J Postgrad Med 2008 Oct-Dec; 54: 273-5.
17. Jafari P, Forouzandeh E, Bagheri Z, Karamizadeh Z, Shalileh K. Health related quality of life of Iranian children with type 1 diabetes: reliability and validity of the Persian version of the PedsQL Generic Core Scales and Diabetes Module. Health Qual Life Outcomes2011; 9: 104.
18. Jafari P, Ghanizadeh A, Akhondzadeh S, Mohammadi MR. Health-related quality of life of Iranian children with attention deficit/hyperactivity disorder. Qual Life Res2011 Feb; 20: 31-6.
19. Khani H, Majdi MR, Marzabadi EA, Montazeri A, Ghorbani A, Ramezani M. Quality of life of iranian beta-thalassaemia major patients living on the southern coast of the Caspian Sea. East Mediterr Health J2012 May; 18: 539-45.
20. Abedzadeh Kalarhoudi M, Taebi M, Sadat Z, Saberi F. Assessment of quality of life in menopausal periods: a population study in kashan, iran. Iran Red Crescent Med J2011 Nov; 13: 811-7.
21. Jafari P, Bagheri Z, Ayatollahi SM, Soltani Z. Using Rasch rating scale model to reassess the psychometric properties of the Persian version of the PedsQL 4.0 Generic Core Scales in school children. Health Qual Life Outcomes2012; 10: 27.
22. Agresti A. Categorical data analysis. Parra et al Análisis exploratorio de factores incidentes en la evaluación docente por parte de los estudiantes Lectiva2000; 4: 279-98.
23. Fienberg SE. The analysis of cross-classified categorical data. 1980.
24. Kemmler G, Holzner B, Kopp M, Dünser M, Greil R, Hahn E, Sperner-Unterweger B. Multidimensional scaling as a tool for analysing quality of life data. Quality of Life Research2002; 11: 223-33.
25. Peterson B, Harrell Jr FE. Partial proportional odds models for ordinal response variables. Applied Statistics1990: 205-17.
26. Kemmler G, Holzner B, Kopp M, Dunser M, Greil R, Hahn E, Sperner-Unterweger B. Multidimensional scaling as a tool for analysing quality of life data. Qual Life Res2002 May; 11: 223-33.

[Downloaded from irje.tums.ac.ir on 2024-08-01]

Iranian Journal of Epidemiology 2014; 10(1): 15-24.

Original Article

Application of the Ordinal Logistic Regression Models in Quality of life Studies

Asadollahi Z[1], Jafari P[2], Rezaeian M[3]

1- Department of Social Medicine, School of Medicine, Rafsanjan University of Medical Sciences, Rafsanjan, Iran
2- Department of Biostatistics, Shiraz University of Medical Sciences, Shiraz, Iran
3- Environmental Health Center, Rafsanjan University of Medical Sciences, Rafsanjan, Iran
Corresponding author: Asadollahi Z., asadollahi.zahra@gmail.com

Background & Objectives: Due to the increasing tendency to measure the quality of life in recent years and the extensive quality of life questionnaires, it is important to determine the appropriate method of analyzing data derived from these studies. The aim of the present study was to introduce ordinal logistic regression models as an appropriate method for analyzing the data of quality of life.

Methods: The data was derived from a cross-sectional study on quality of life survey of 938 students. For data analysis, two binary logistic regression models and ordinal logistic regression models were used and the results of these models were compared.

Results: The results of goodness of fit showed that all three models were fitted well. Based on the ordinal logistic regression models, the three variables out of the explanatory variables were statistically associated with the response while based on the binary logistic regression model, after combining two categories of response variable, only two variables were significant. Therefore, combining the categories of the response variable should be avoided as much as possible because it may lead to data loss due to ignoring some of the response categories.

Conclusion: It is concluded that to analyze quality of life data, due to the nature of the response variable, ordinal logistic regression models are recommended considering the fewer parameter estimates and easier interpretation of the results.

Keywords: Quality of life, Ordinal logistic regression models, Binary Logistic Regression Model

[Downloaded from irje.tums.ac.ir on 2024-08-01]