© Statistical Research and Training Center
Archive of SID
مجلّه‌ی پژوهش‌های آماری ایران
دوره‌ی ۵، شماره‌ی ۱، بهار و تابستان ۱۳۸۷، صص ۷۵–۹۴
J. Statist. Res. Iran 5 (2008): 75–94

مدل انتقالی برای تحلیل داده‌های طولی با پاسخ‌های دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی

زهرا رضایی قهرودی،[†] مجتبی گنجعلی[†,‡,*] و فاطمه هرندی[†]

[†] پژوهشکده‌ی آمار
[‡] دانشگاه شهید بهشتی

چکیده. در بسیاری از مطالعات طولی، پاسخ‌های دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی اندازه‌گیری می‌شوند. در این مطالعات هدف، بررسی اثر متغیرهای کمکی بر روی پاسخ‌های وابسته به زمان است. یک تحلیل رگرسیونی برای این نوع داده‌ها، باید همبستگی بین پاسخ‌ها در طول زمان را درنظر گیرد. برای تحلیل این پاسخ‌های اسمی–ترتیبی، با استفاده از یک روش وزن‌دهی مناسب، یک مدل انتقالی دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی پیشنهاد شده است و سپس از روش ماکسیمم درست‌نمایی برای یافتن برآورد پارامترها استفاده شده است. تابع درست‌نمایی در این روش به‌گونه‌ای تفکیک شده است که استفاده از نرم‌افزارهای موجود ممکن باشد. این روش در طرح آمارگیری نیروی کار مرکز آمار ایران به کار برده شده است که در آن پاسخ ترتیبی در زمان اول، طول مدت بیکاری افراد بیکار و پاسخ اسمی در زمان دوم، وضع فعالیت اقتصادی افراد است. آن‌چه در این مطالعه مد نظر است پیدا کردن عوامل مؤثر بر بیکار باقی‌ماندن و یا تغییر وضع فعالیت از بیکار به شاغل یا غیر فعال است.

واژگان کلیدی. طول مدت بیکاری؛ وضع فعالیت اقتصادی؛ داده‌های طولی؛ مدل انتقالی؛ شبه R^2ی ناگل‌کرک.

۱ مقدمه

اغلب در آمارگیری‌های طولی، پاسخ‌های وابسته که به‌عنوان متغیرهای آمیخته‌ی ترتیبی و اسمی ثبت شده باشند، در حضور اطلاعات کمکی مشاهده می‌شوند. چنین اندازه‌گیری‌های مکرر به‌عنوان ابزار مهمی در

[*] نویسنده‌ی عهده‌دار مکاتبات

www.SID.ir

Archive OF SID

تحقیقات اجتماعی-اقتصادی به حساب می‌آیند. چون این پاسخ‌های ترتیبی-اسمی برای افراد یکسان در طول زمان جمع‌آوری می‌شوند، بنا بر این، این پاسخ‌ها با هم همبستگی دارند. یک روش ساده برای تحلیل چنین داده‌هایی در نظر نگرفتن این همبستگی و استفاده از تحلیل‌های استاندارد تک‌متغیره مانند استفاده از مدل لوجیت یا پروبیت تجمعی برای پاسخ‌های ترتیبی و مدل لوجیت تعمیم‌یافته برای پاسخ‌های اسمی است. لیانگ و زیگر (۱۹۸۶، قضیه‌ی ۱) به این مطلب اشاره کرده‌اند که اگر چه برآورد پارامترها از طریق تحلیل‌های تک متغیره بدون در نظر گرفتن همبستگی سازگار است ولی زمانی که میزان همبستگی بالا باشد، برآوردهایی با کارایی پایین به دست می‌آیند و از دست دادن کارایی ممکن است منجر به بیش‌برآورد دقت اثرات متغیرهای کمکی شود. بنا بر این موضوعات مورد علاقه در این مطالعات احتساب همبستگی بین پاسخ‌ها و ارزیابی اثرات متغیرهای کمکی روی پاسخ‌ها به طور همزمان می‌باشد.

با توجه به پیامدهای سیاسی و اجتماعی-اقتصادی ناشی از بیکاری در جامعه و ضرورت آشنایی با مشخصه‌های نیروی کار، به خصوص مشخص کردن ساختار و چگونگی تغییر وضع فعالیت اقتصادی افراد جامعه با هدف برنامه‌ریزی صحیح برای دست‌یابی به اهداف توسعه‌ی فرهنگی و اجتماعی-اقتصادی، اطلاعات افراد بیکار در یک مطالعه‌ی طولی برای بررسی انتخاب شده است. به این ترتیب که در این مقاله با پاسخ‌های دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی سر و کار داریم که طول مدت بیکاری افراد بیکار (که به‌صورت ترتیبی ثبت شده است، بخش بعد را ببینید) به‌عنوان پاسخ ترتیبی دور اول آمارگیری و وضع فعالیت اقتصادی این افراد به‌عنوان پاسخ اسمی دور دوم آمارگیری اندازه‌گیری شده‌اند.

رهیافت‌های مختلفی برای در نظر گرفتن همبستگی بین پاسخ‌ها در مطالعات طولی وجود دارد. یکی از این رهیافت‌ها مدل‌های حاشیه‌ای است که استنباط‌هایی برای پارامترهای مدل بر اساس میانگین‌گیری روی کل جامعه انجام می‌دهد (لیانگ و دیگران، ۱۹۹۲، استیراتلی و دیگران ۱۹۸۴، تن هاو و دیگران، ۱۹۹۶، ورمونت و دیگران، ۲۰۰۱). روش دیگر برای تحلیل این‌گونه داده‌ها مدل‌بندی اثرات تصادفی است که ناهمگنی بین پاسخگوها توسط این مدل بررسی می‌شود. هارویل و می (۱۹۸۴)، بریج و دوس‌سانتوس (۱۹۹۶)، وربک و لسافره (۱۹۹۶)، توتز (۲۰۰۵) و لئون و دیگران (۲۰۰۷) از جمله افرادی هستند که از مدل‌های اثرات تصادفی برای تحلیل این‌گونه داده‌ها استفاده کرده‌اند. این مدل عموماً برای داده‌های پانلی بلندمدت مناسب است که در این مقاله به‌دلیل بررسی یک مطالعه‌ی پانلی کوتاه‌مدت، استفاده از این روش پیشنهاد نمی‌شود. سومین روش، استفاده از مدل‌های مارکوف یا انتقالی است که به بررسی تغییرات میان‌دوره‌های آمارگیری بر اساس پاسخ‌های دوره‌ی قبل می‌پردازد. این مدل توسط بارتلت (۱۹۵۱)، اندرسون و گودمن (۱۹۵۷)، بیلینگزلی (۱۹۶۱)، گاربر (۱۹۸۹)، چانگ و دیگران (۲۰۰۵) و سنگال و دیگران (۲۰۰۷) پیشنهاد شده است. پاسخ به این سؤال که کدام مدل باید انتخاب شود عمدتاً به سئوال

www.SID.ir

Archive OF SID

یا سئوالات تحقیقاتی بستگی دارد که باید به آن‌ها جواب داده شود. در مدل‌های انتقالی، امید ریاضی پاسخ در یک موقعیت یا زمان مشخص بر حسب پاسخ‌ها در موقعیت‌های قبلی مدل‌بندی می‌شوند. در مدل‌های حاشیه‌ای، امید ریاضی پاسخ‌ها وابسته به متغیرهای تبیینی هستند و مدل‌های اثرهای تصادفی به‌دلیل دخالت پارامترهایی که برای آزمودنی‌ها یا واحدها خاص هستند، متفاوت از دو مدل قبلی است. چون هدف ما در این مقاله بررسی دلیل بیکار باقی ماندن افراد بیکار در دور اول آمارگیری یا تغییر وضع فعالیت آن‌ها به شاغل یا غیر فعال شدن در دور دوم آمارگیری است و سؤال اصلی و مورد علاقه بررسی چگونگی انتقال بین دو دوره‌ی آمارگیری است، برای پاسخ به این سؤال علمی یک روش مناسب استفاده از مدل‌های انتقالی است. برای مطالعه‌ی بیش‌تر به اگرستی (۲۰۰۲) و دیگل و دیگران (۲۰۰۲) مراجعه شود.

همه‌ی مطالعاتی که تاکنون درباره‌ی مدل‌های انتقالی معرفی شده‌اند، فقط اندازه‌گیری‌های مکرری را بررسی کرده‌اند که مقیاس اندازه‌گیری یکسانی داشته باشند، مانند پاسخ‌های چندمتغیره‌ی اسمی و یا پاسخ‌های چندمتغیره‌ی ترتیبی. در این مقاله، کاربردی از مدل انتقالی بیان شده است که مقیاس اندازه‌گیری پاسخ‌ها با زمان تغییر می‌کند و با تلفیقی از پاسخ‌های ترتیبی و اسمی مواجه هستیم. به‌همین دلیل است که آن را پاسخ‌های دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی نامیده‌ایم. مدل انتقالی که در این مقاله معرفی شده است، روی داده‌های نیروی کار مرکز آمار ایران در سال ۱۳۸۵ به کار برده شده است. روش برآورد پارامترها استفاده از روش ماکسیمم درست‌نمایی است. توزیع پاسخ‌های اسمی به‌عنوان تابعی از متغیرهای کمکی و پاسخ‌های دوره‌ی قبل (پاسخ ترتیبی) در نظر گرفته شده است. از آن جا که به‌منظور دست‌یابی به استنباط‌های آماری معتبر برای جامعه، طرح نمونه‌گیری در تحلیل داده‌ها باید درنظر گرفته شود، برای آشنایی با آمارگیری نیروی کار مرکز آمار ایران به طور مختصر در بخش ۲ توضیحاتی داده می‌شود. همچنین در این بخش نتایج تحلیل توصیفی اولیه از داده‌های بیکاران بیان شده است. برای برآورد پارامترهای مدل، با در نظر گرفتن طرح نمونه‌گیری، بخش ۳، جزییات تابع درست‌نمایی را برای مدل انتقالی دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی بیان می‌کند و سپس معیارهای نیکویی برازش این مدل ارائه می‌شود. در بخش ۴ نتایج استفاده از این مدل برای داده‌های نیروی کار ارائه می‌شود. در این بخش نتایج آزمون نیکویی برازش برای داده‌های مورد بررسی نیز بیان و در انتها نتیجه‌گیری ارائه خواهد شد.

۲ توصیف آمارگیری نیروی کار و داده‌های بیکاران

در این بخش کلیاتی در مورد آمارگیری نیروی کار مرکز آمار ایران و برخی از نتایج تحلیل توصیفی افراد بیکار در طرح مزبور بیان می‌شود.

www.SID.ir

Archive OF SID

۲/۱ معرفی آمارگیری نیروی کار

آمارگیری نیروی کار، یک آمارگیری نمونه‌ای چرخشی فصلی و یکی از مهم‌ترین آمارگیری‌های نمونه‌ای خانواری در ایران است. طراحی این آمارگیری بر اساس آخرین توصیه‌های سازمان بین‌المللی کار انجام شده است و برای اولین بار در سال ۱۳۸۴ با ۴۴۵۶۸ خانوار نمونه در هر فصل اجرا شد تا با استفاده از آن بتوان آمارهای جاری نیروی کار کشور و همچنین تغییرات مرتبط را در دوره‌های زمانی کوتاه بدست آورد. این آمارگیری به‌منظور تأمین نیازهای آماری سیاست‌گذاران، تصمیم‌گیران و محققان در بخش‌های اقتصادی و اجتماعی برای ارزیابی سیاست‌های قبلی دولت و برنامه‌ریزی برای تصمیمات آینده اجرا می‌شود.

جامعه‌ی هدف طرح، شامل همه‌ی اعضای خانوارهای معمولی ساکن در نقاط شهری و روستایی کشور است. مرکز آمار ایران برای آمارگیری‌های خانواری، یک نمونه‌ی پایه از واحدهای جغرافیایی دارد که برای آمارگیری نیروی کار نیز از آن استفاده می‌کند. چارچوب نمونه‌ی پایه‌ی طرح، فهرست بلوک‌ها و آبادی‌های کشور است که از اطلاعات فایل فرم فهرست‌برداری سرشماری عمومی کارگاهی ۱۳۸۱ به دست آمده است. این فهرست علاوه بر اطلاعات جغرافیایی، تعداد خانوار و جمعیت بلوک‌ها و آبادی‌ها را نیز شامل می‌شود. از این چارچوب برای ساخت خوشه‌ها استفاده می‌شود. خوشه‌ها واحدهایی هستند که شامل حد اقل ۱۰۰ و حد اکثر ۲۰۰ خانوار می‌باشند (برای جزییات بیش‌تر به طرح آمارگیری نیروی کار مرکز آمار ایران سال ۱۳۸۵ مراجعه شود). به‌منظور برآورد تغییرات بین دوره‌ها علاوه بر برآوردهای سطوح جاری، از نمونه‌گیری چرخشی استفاده شده است (کوکران، ۱۹۷۷، ص ۳۴۴-۳۴۷). الگوی چرخش انتخابی برای این طرح، یک الگوی ۲-۲-۲ است. یعنی از هر خانوار نمونه، چهار بار آمارگیری به عمل می‌آید، به این ترتیب که خانوار دو فصل متوالی در نمونه است، سپس به طور موقت برای دو فصل متوالی از نمونه خارج می‌شود و مجدداً برای دو فصل متوالی به نمونه باز می‌گردد و پس از آن برای همیشه از نمونه خارج می‌شود.

روش نمونه‌گیری، خوشه‌ای دو مرحله‌ای با طبقه‌بندی است که واحد نمونه‌گیری مرحله‌ی اول خوشه و واحد نمونه‌گیری مرحله‌ی دوم، زیرگروهی شامل سه خانوار مجاور است. در هر خوشه‌ی نمونه، با استفاده از فهرست خانوارهای ساکن در خوشه (چارچوب ثانویه)، زیر گروه‌های سه خانواری از خانوارهای مجاور ساخته می‌شود. هر یک از این زیرگروه‌ها یک گروه چرخش نامیده می‌شوند و در هر فصل آمارگیری، از هر خوشه‌ی نمونه، ۴ گروه چرخش (شامل ۱۲ خانوار) آمارگیری می‌شود. برای هر خانوار نمونه، یک فرم خانواری پر می‌شود که مشخصات همه‌ی اعضای خانوار در آن ثبت می‌شود و یک فرم فردی نیز برای افراد ۱۰ ساله و بیش‌تر عضو خانوار پر می‌شود. در این طرح واحد اطلاع‌گیری، خانوار ساکن در محل سکونت خانوار نمونه است.

برای برآورد پارامترهای جامعه، ضروری است که داده‌هایی که از این آمارگیری برای هر عضو خانوارهای

www.SID.ir

Archive OF SID

نمونه جمع‌آوری شده است به روش مناسبی وزن‌دهی شود. این وزن‌ها پس از اعمال سه مرحله‌ی وزن‌دهی «محاسبه‌ی وزن پایه»، «تعدیل وزن برای واحدهای بی‌پاسخ» و «تعدیل وزن بر اساس پیش‌بینی‌های جمعیتی» (دویل و سارندال، ۱۹۹۲) به دست آمده‌اند که در بخش‌های بعدی از این وزن‌ها در تابع درست‌نمایی برای برآورد پارامترها استفاده می‌شود.

۲٫۲ توصیف داده‌ها

در این مقاله همان‌طور که قبلاً اشاره شد علاقه‌مند به بررسی الگوی تغییر وضع فعالیت اقتصادی افراد بیکار پس از گذشت ۳ ماه با توجه به طول مدت بیکاری آن‌ها هستیم. به‌منظور مطالعه‌ی اثر مشخصه‌های فردی و خانواری روی الگوی نام‌برده، در این مقاله از داده‌های آمارگیری نیروی کار مرکز آمار ایران در بهار و تابستان ۱۳۸۵ استفاده شده است.

جامعه‌ی مورد مطالعه افراد بیکار کشور در بهار ۱۳۸۵ است. به‌منظور بررسی الگوی تغییر وضع فعالیت اقتصادی این افراد پس از ۳ ماه و ارتباط آن با مدت بیکاری آن‌ها، افراد بیکار در نمونه‌ی فصل بهار ۱۳۸۵ آمارگیری نیروی کار را که در فصل بعد آمارگیری نیز در نمونه باقی مانده‌اند (۲۵۲۵ نمونه) را مورد مطالعه قرار داده‌ایم. در داده‌های اصلی، طول مدت بیکاری این افراد در بهار ۱۳۸۵ (پاسخ اولیه) بر حسب ماه جمع‌آوری شده است که با توجه به‌دلایل زیر این پاسخ‌ها به‌صورت رده‌های ترتیبی طول مدت بیکاری بر حسب ماه در فاصله‌ی (۱: [۳ – ۱]، ۲: [۶ – ۴]، ۳: [۱۲ – ۷]، ۴: [۲۴ – ۱۳] و ۵: ۲۵ ⩽) طبقه‌بندی شده‌اند. یکی از دلایل ترتیبی کردن طول مدت بیکاری این است که پاسخ‌های افراد معمولاً به‌طور تقریبی پاسخ‌های درستی هستند و ممکن است که این پاسخ‌ها با خطا همراه باشند. از طرفی برای سیاست‌گذاران مسایل بازار کار بررسی توزیع زمان بیکاری بر اساس رده‌های طول مدت بیکاری معرفی شده، برای ارائه‌ی تحلیل‌هایی در مورد بیکاری کوتاه‌مدت، میان‌مدت و بلندمدت از اهمیت بیش‌تری برخوردار است. وضع فعالیت اقتصادی بیکاران بهار ۱۳۸۵ در تابستان ۱۳۸۵ (پاسخ پی‌گیر) نیز در سه رده‌ی اسمی (۱: بیکار[۱]، ۲: شاغل[۲]، ۳: غیر فعال[۳]) طبقه‌بندی شده است.

جدول ۱ توزیع نمونه‌ای دو پاسخ مورد نظر که طول مدت بیکاری بیکاران در بهار ۱۳۸۵ و وضع فعالیت اقتصادی این افراد (واحدهای نمونه) در تابستان ۱۳۸۵ است را بیان می‌کند. احتمال نمونه‌ای در جدول ۱ نشان می‌دهد که با افزایش طول مدت بیکاری، درصد افرادی که پس از ۳ ماه شاغل می‌شوند، کاهش می‌یابد. یعنی احتمال شاغل شدن بیکاران با طول مدت بیکاری آن‌ها رابطه‌ی عکس دارد. همچنین این جدول نشان می‌دهد که برای طول مدت بیکاری بین [۳ – ۱] و [۶ – ۴] ماه، درصد شاغل شدن خیلی

www.SID.ir

Archive OF SID

جدول ۱. توزیع نمونه‌ای طول مدت بیکاری و وضع فعالیت اقتصادی به‌ترتیب به‌عنوان پاسخ‌های دوره‌ی اول و دوم آمارگیری (تعداد مشاهدات، درصد سطری و درصد ستونی)

طول مدت بیکاری افراد نمونه‌ی بیکار در بهار ۱۳۸۵	وضع فعالیت اقتصادی (تابستان ۱۳۸۵)			مجموع
	بیکار	شاغل	غیر فعال	
[۱ − ۳]	۱۵۳	۲۹۱	۸۶	۵۳۰
	۲۸٫۸۷%	۵۴٫۹۱%	۱۶٫۲۳%	۱۰۰٫۰%
	۱۵٫۹۴%	۳۲٫۹۹%	۱۲٫۵۹%	۶۱٫۵۲%
[۴ − ۶]	۱۶۳	۱۸۲	۷۳	۴۱۸
	۳۹٫۰۰%	۴۳٫۵۴%	۱۷٫۴۶%	۱۰۰٫۰%
	۱۶٫۹۸%	۲۰٫۶۳%	۱۰٫۶۹%	۴۸٫۳%
[۷ − ۱۲]	۲۲۲	۱۵۴	۱۹۹	۵۷۵
	۳۸٫۶۱%	۲۶٫۷۸%	۳۴٫۶۱%	۱۰۰٫۰%
	۲۳٫۱۳%	۱۷٫۴۶%	۲۹٫۱۴%	۵۴٫۴۱%
[۱۳ − ۲۴]	۲۱۲	۱۲۸	۱۴۹	۴۹۰
	۴۳٫۳۵%	۲۶٫۱۸%	۳۰٫۴۷%	۱۰۰٫۰%
	۲۲٫۰۸%	۱۴٫۵۱%	۲۱٫۸۲%	۵۸٫۴۱%
⩾ ۲۵	۲۱۰	۱۲۷	۱۷۶	۵۱۷
	۴۰٫۹۴%	۲۴٫۷۶%	۳۴٫۳۱%	۱۰۰٫۰%
	۲۱٫۸۸%	۱۴٫۴۰%	۲۵٫۷۷%	۶۲٫۰۵%
مجموع	۹۶۰	۸۸۲	۶۸۳	۲۵۲۵
	۳۸٫۰۲%	۳۴٫۹۳%	۲۷٫۰۵%	۱۰۰٫۰%
	۱۰۰٫۰%	۱۰۰٫۰%	۱۰۰٫۰%	۱۰۰٫۰%

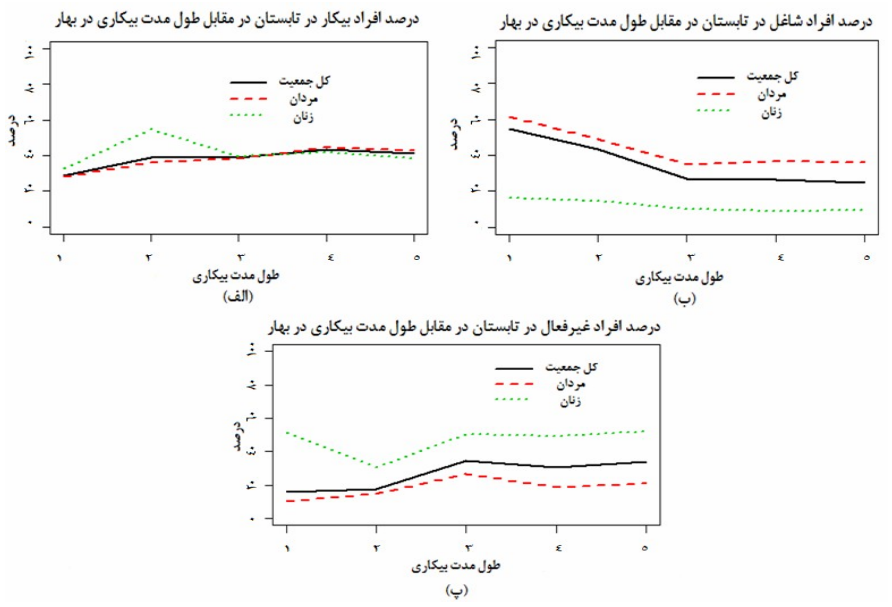
بیش‌تر از بیکار ماندن یا غیر فعال شدن است. برای رده‌های دیگر طول مدت بیکاری، بالاترین درصد مربوط به بیکار ماندن است. بنا بر این، به نظر می‌رسد که بین این دو پاسخ همبستگی وجود دارد و برای بدست آوردن استنباط صحیح، این همبستگی باید در نظر گرفته شود.

به‌دلیل این‌که ممکن است توزیع توأم طول مدت بیکاری و وضع فعالیت اقتصادی برای مردان با زنان متفاوت باشد، الگوی تغییرات وضع فعالیت اقتصادی بیکاران در بهار، در تابستان به طور مجزا برای مردان، زنان و کل افراد نمونه مورد بررسی قرار گرفته شده است. شکل ۱ (الف) نشان می‌دهد که درصد زنان بیکار در فصل بهار با طول مدت بیکاری حد اکثر ۱۲ ماه (حد اکثر ۳ در مقیاس ترتیبی) که در فصل تابستان نیز بیکار باقی‌مانده‌اند بیش‌تر از مردان با این ویژگی است. برای افراد با بیش از ۱۲ ماه

www.SID.ir

Archive OF SID

بیکاری در بهار، درصد بیکار باقی‌ماندن پس از ۳ ماه برای مردان بیش‌تر از زنان است. با توجه به شکل ۱ (الف) و (پ)، به نظر می‌رسد که زنان، فقط مدت کوتاهی بیکار ماندن را تحمل می‌کنند و چنان‌چه نتوانستند کاری پیدا کنند، از نیروی کار خارج شده و غیر فعال می‌شوند. شکل ۱ (ب) نیز نشان می‌دهد که درصد مردان بیکاری که در تابستان شاغل شده‌اند برای هر سطح طول مدت بیکاری در بهار، بیش‌تر از زنان است. همچنین این شکل نشان می‌دهد که با افزایش طول مدت بیکاری تا یک سال، احتمال شاغل شدن دارای روندی کاهشی است ولی پس از یک سال، این احتمال، تقریباً هم برای مردان و هم برای زنان ثابت می‌ماند. برای افرادی که در تابستان از نیروی کار خارج شده‌اند شکل ۱ (پ) یک جهت معکوس در مقایسه با شکل (ب) نشان می‌دهد، به این صورت که درصد زنان بیکاری که پس از ۳ ماه غیر فعال شده‌اند، برای همه‌ی سطوح طول مدت بیکاری در بهار، بیش‌تر از مردان است. شکل ۱ (پ) همچنین نشان می‌دهد که شکل منحنی برای افراد بیکار با بیش از ۶ ماه بیکاری در بهار، برای زنان، مردان و کل جامعه شبیه است اما برای طول مدت بیکاری ۶ ماه و کم‌تر در فصل بهار، این الگو برای زنان متفاوت از مردان و کل جامعه است.



شکل ۱. درصد افراد بیکار (الف)، افراد شاغل (ب) و افراد غیر فعال (پ) در تابستان سال ۱۳۸۵ در مقابل طول مدت بیکاری در بهار سال ۱۳۸۵ به‌تفکیک کل جمعیت، مردان و زنان

www.SID.ir

Archive OF SID

۳ مدل انتقالی دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی و نیکویی برازش آن

در این بخش، مدل انتقالی دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی معرفی می‌شود و تحلیل آن با استفاده از روش ماکسیمم درست‌نمایی بیان می‌شود. برای پاسخ اولیه (پاسخ در دور اول آمارگیری)، مدل لوجیت تجمعی حاشیه‌ای، اثرات متغیرهای کمکی را روی لگاریتم بخت انتخاب رده‌ی کوچک‌تر نسبت به بزرگ‌تر به‌صورت زیر برآورد می‌کند.

$$\log\left\{\frac{\Pr(Y_{i1}\leqslant l_1;\alpha_1,\beta_1)}{\Pr(Y_{i1}>l_1;\alpha_1,\beta_1)}\right\}=\alpha_{1,l_1}+\sum_{k=1}^{K_1}\beta_{1,k}X_{i1k}\qquad l_1=1,...,L_1, \tag{۱}$$

که در آن Y_{i1} پاسخ اولیه‌ی iامین فرد، $X_{i1}=(X_{i11},...,X_{i1K_1})'$ بردار متغیرهای کمکی با مؤلفه‌ی X_{i1k} متناظر با kامین متغیر کمکی iامین فرد در اولین زمان، L_1+1 تعداد رده‌های ترتیبی متغیرهای پاسخ، $l_1=1,...,L_1$ نقاط برش در زمان اول است که نشان‌دهنده‌ی لگاریتم بخت انتخاب رده‌های کوچک‌تر نسبت به رده‌های بزرگ‌تر زمانی که متغیرهای کمکی صفر در نظر گرفته شده باشند، است. $\alpha_1=(\alpha_{1,1},...,\alpha_{1,L_1})'$ بردار پارامترهای برش است که $\alpha_{1,1}\leqslant\alpha_{1,2}\leqslant...\leqslant\alpha_{1,L_1}$ و $\beta_1=(\beta_{1,1},...,\beta_{1,K_1})'$ بردار ضرایب رگرسیونی برای متغیرهای کمکی است. احتمال این‌که پاسخ فرد iام در دور اول آمارگیری در اولین سطح پاسخ ($l_1=1$) باشد، به‌صورت زیر است:

$$\pi_{i1,1}=\Pr(Y_{i1}=1|\alpha_1,\beta_1)=\frac{e^{\alpha_{1,1}+X_{i1}'\beta_1}}{1+e^{\alpha_{1,1}+X_{i1}'\beta_1}} \tag{۲}$$

و برای $l_1=2,...,L_1$ داریم

$$\begin{aligned}\pi_{i1,l_1}&=\Pr(Y_{i1}=l_1|\alpha_1,\beta_1)\\&=\Pr(Y_{i1}\leqslant l_1|\alpha_1,\beta_1)-Pr(Y_{i1}\leqslant l_1-1|\alpha_1,\beta_1)\\&=\frac{e^{\alpha_{1,l_1}+X_{i1}'\beta_1}}{1+e^{\alpha_{1,l_1}+X_{i1}'\beta_1}}-\frac{e^{\alpha_{1,(l_1-1)}+X_{i1}'\beta_1}}{1+e^{\alpha_{1,(l_1-1)}+X_{i1}'\beta_1}}.\end{aligned} \tag{۳}$$

برای پاسخ اسمی به‌عنوان پاسخ در زمان دوم با L_2+1 سطح، مدل لوجیت انتقالی برای پاسخ‌های دودویی ($L_2=1$) به مدل لوجیت انتقالی تعمیم‌یافته (مک‌فادن، ۱۹۷۴) برای $l_1=1,...,L_1+1$

www.SID.ir

Archive OF SID

به‌صورت زیر تعمیم می‌یابد:

$$\eta_{i2,l_2|l_1} = \log\left\{\frac{\Pr(Y_{i2}=l_2|Y_{i1}=l_1;\gamma,\alpha_2,\beta_2)}{\Pr(Y_{i2}=L_2+1|Y_{i1}=l_1;\gamma,\alpha_2,\beta_2)}\right\}$$

$$= \alpha_{2,l_2} + \sum_{k=1}^{K_2}\beta_{2,l_2,k}X_{i2k} + \gamma_{l_2}Y_{i1} \qquad l_2 = 1,...,L_2. \tag{۴}$$

در این معادله $\alpha_2 = (\alpha_{2,1},...,\alpha_{2,L_2})'$، بردار پارامترهای ثابت با L_2 عضو، $\beta_2 = (\beta_{2,1,1},...,\beta_{2,L_2,K_2})'$ بردار ضرایب رگرسیونی برای متغیرهای کمکی با L_2K_2 مؤلفه و $X_{i2} = (X_{i21},...,X_{i2K_2})'$ بردار متغیرهای کمکی با مؤلفه‌ی X_{i2k} متناظر با kامین متغیر کمکی برای iامین فرد در زمان دوم است. $\gamma = (\gamma_1,...\gamma_{L_2})'$ بردار با L_2 پارامتر است که نشان‌دهنده‌ی اثر پاسخ ترتیبی زمان قبل، روی پاسخ اسمی زمان حال است. این مدل می‌تواند برحسب احتمال‌های واقعی $\pi_{i2,l_2|l_1}$ به جای لگاریتم بخت بیان شود. با شروع از معادله‌ی (۴) و پذیرش این قید که $\eta_{i2,(L_2+1)|l_1} = 0$ باشد، احتمال‌های پاسخ برای دوره‌ی زمان دوم با استفاده از مدل لوجیت انتقالی تعمیم‌یافته به‌صورت زیر بیان می‌شود:

$$\pi_{i2,l_2|l_1} = \Pr(Y_{i2}=l_2|Y_{i1}=l_1;\gamma,\alpha_2,\beta_2)$$

$$= \frac{\exp\{\eta_{i2,l_2|l_1}\}}{1+\sum_{r=1}^{L_2}\exp\{\eta_{i2,r|l_1}\}}, \qquad l_2 = 1,..,L_2+1,\ l_1 = 1,..,L_1+1 \tag{۵}$$

با استفاده از مدل انتقالی، تابع درست‌نمایی برای مدل دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی به‌صورت

$$L = \prod_{i=1}^{n}\left[\,\pi_{i1,y_{i1}}\right]^{W_{i1}}\left[\pi_{i2,y_{i2}|y_{i1}}\right]^{W_{i2}} \tag{۶}$$

است که n تعداد افراد در مطالعه، W_{i1} وزن نمونه‌گیری برای iامین فرد در اولین زمان و W_{i2} وزن نمونه‌گیری برای iامین فرد در زمان دوم است. لگاریتم تابع درست‌نمایی به‌صورت زیر است:

$$l(\theta) = \sum_{i=1}^{n}\left\{W_{i1}\log(\pi_{i1,y_{i1}}) + W_{i2}\log(\pi_{i2,y_{i2}|y_{i1}})\right\} \tag{۷}$$

www.SID.ir

Archive OF SID

که در آن $\pi_{i۱,y_{i۱}}$ و $\pi_{i۲,y_{i۲}|\,y_{i۱}}$ احتمال‌های پاسخ در اولین و دومین زمان است و

$$\pi_{i۱,(L_۱+۱)} = ۱ - \sum_{l_۱=۱}^{L_۱} \pi_{i۱,l_۱}, \qquad \pi_{i۲,(L_۲+۱)|\,l_۱} = ۱ - \sum_{l_۲=۱}^{L_۲} \pi_{i۲,l_۲|,l_۱}.$$

در نظر گرفتن نیکویی برازش مدل نیز یک مسئله‌ی مهم است که باید به آن پرداخته شود. برای پاسخ‌های ترتیبی، آزمون نیکویی برازش نسبتی (آزمون نسبت درست‌نمایی) ممکن است برای ارزیابی برازش این مدل مناسب نباشد. در بعضی شبیه‌سازی‌ها نشان داده شده است که آماره‌ی کیبش Deviance زمانی که داده‌ها خیلی پراکنده باشند، دارای توزیع خی‌دو نیست. به‌عنوان مثال زمانی که تعداد مشاهدات کمی در خانه‌هایی که از طریق طبقه‌بندی متقاطع همه‌ی سطوح متغیرهای کمکی بدست آمده باشد، وجود داشته باشند، آماره‌ی کیبش مناسب نیست (که‌لر، ۱۹۸۶، که‌لر و لارنتز، ۱۹۸۰، لارنتز، ۱۹۷۸). در داده‌هایی که ما در این مقاله استفاده کرده‌ایم، به‌دلیل این که متغیرهای کمکی زیادی در مدل وجود دارند و اغلب آن‌ها گسسته هستند، استفاده از آزمون نسبت درست‌نمایی پیشنهاد نمی‌شود. بنا بر این، باید به دنبال معیار دیگری باشیم. برای پاسخ‌های ترتیبی، دو نسخه‌ی تغییریافته‌ی ضریب همبستگی چندگانه‌ی $R^۲$ وجود دارد. یک نسخه‌ی آن توسط کاکس و اسنل (۱۹۸۹) بسط داده شده است و دیگری توسط ناگل‌کرک (۱۹۹۱) به‌گونه‌ای تغییر یافته است که بتواند به مقدار یک برسد. شبه $R^۲$ی ناگل‌کرک که ما برای نیکویی برازش مدل لوجیت تجمعی به کار برده‌ایم به‌صورت زیر تعریف می‌شود که در آن به جای اندازه‌ی نمونه از جمع وزن‌ها که معرف اندازه‌ی جامعه است، استفاده شده است.

$$R^۲ = \frac{۱ - \exp\left\{-۲(\sum_{i=۱}^{n} W_{i۱})^{-۱}(\ln L_{\text{Full}} - \ln L_{\text{Null}})\right\}}{۱ - \exp\left\{۲(\sum_{i=۱}^{n} W_{i۱})^{-۱} \ln L_{\text{Null}}\right\}} \qquad (۸)$$

که در آن L_{Full} و L_{Null} به ترتیب تابع درست‌نمایی مدل برازش‌شده و مدل صفر (تنها با در نظر گرفتن مقادیر ثابت) است. شبه $R^۲$ می‌تواند به‌عنوان نسبتی از واریانس که توسط مدل تشریح می‌شود، تفسیر شود. برای پاسخ‌های اسمی، ضریب پیشایندی $R_c^۲$ (آلدریچ و نلسون، ۱۹۸۴) به‌عنوان یک آماره‌ی نیکویی برازش استفاده می‌شود که به‌صورت زیر تعریف می‌شود:

$$R_c^۲ = \frac{G_M}{G_M + \sum_{i=۱}^{n} W_{i۲}} \qquad (۹)$$

که در آن $G_M = -۲[\ln L_{\text{Null}} - \ln L_{\text{Full}}]$ است. این آماره در بازه‌ی $(۰, ۱)$ مقدار اختیار می‌کند.

www.SID.ir

Archive OF SID

۴ مدل و نتایج برای داده‌های نیروی کار

در این بخش، مدل انتقالی دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی برای داده‌های نیروی کار معرفی می‌شود. متغیرهای کمکی پیوسته و گسسته‌ای که در این مدل استفاده می‌شوند و کدهای مربوط به آن‌ها در جدول ۲ بیان شده‌اند. مدل‌های لوجیت دومتغیره‌ای که در بخش قبل آن‌ها را معرفی کردیم برای داده‌های مورد استفاده در این مقاله با در نظر گرفتن متغیرهای کمکی که در جدول ۲ معرفی شده‌اند، برازش داده می‌شوند.

در مدل‌های مورد استفاده برای پاسخ‌های ترتیبی و اسمی، متغیرهای کمکی که بیش از ۲ سطح داشته باشند مانند سطح تحصیلات، وضع زناشویی و . . .، به‌عنوان فاکتور در نظر گرفته و با تعریف معمول به متغیرهای دودویی تبدیل می‌شوند. اکنون با استفاده از مدل انتقالی دومتغیره‌ی آمیخته‌ی ترتیبی و اسمی (مدل ۱ و ۴) تحلیل داده‌های نیروی کار سال ۱۳۸۵ می‌پردازیم. با توجه به این که در تابع درست‌نمایی رابطه‌ی (۶) وزن هر یک از افراد نمونه در نظر گرفته شده است، بنا بر این، نتایج مدل‌بندی قابل تعمیم به کل جامعه می‌باشد. بنا بر این، جدول ۳ نتایج مدل‌بندی حاشیه‌ای طول مدت بیکاری به‌عنوان پاسخ ترتیبی زمان اول برای کل جامعه و به طور مجزا به تفکیک جنس را بیان می‌کند. نتایج این جدول نشان می‌دهد که هرچه سن بیش‌تر باشد، طول مدت بیکاری با احتمال بیش‌تری بالا می‌رود. همچنین احتمال این که سرپرست خانوار طول مدت بیکاری کم‌تری را نسبت به بقیه‌ی اعضای خانوار داشته باشد، بیش‌تر است. چنین تعبیری برای افرادی که مهاجر هستند و کسانی که در حال تحصیل هستند نیز برقرار است. همچنین مردها با احتمال بیش‌تری طول مدت بیکاری کم‌تر دارند. اثر سطح تحصیلات نشان می‌دهد که افراد در سطح دیپلم یا پیش‌دانشگاهی طول مدت بیکاری بیش‌تری دارند. نتایج همچنین نشان می‌دهد که مردانی که ازدواج کرده‌اند، طول مدت بیکاری کم‌تری دارند.

جدول ۳ همچنین نشان می‌دهد که افراد یا مردانی که تجربه‌ی کار قبلی دارند، طول مدت بیکاری کم‌تری دارند و احتمال داشتن طول مدت بیکاری کم‌تر برای افرادی که قبل از بیکاری‌شان شاغل بوده‌اند نسبت به بقیه‌ی افراد (خانه‌دار، بازنشسته و...) بالاتر است. این احتمال برای افرادی که قبل از بیکاری تحصیل می‌کرده‌اند پایین‌تر است. این اثر برای زنان و کل افراد نیز معنی‌دار است.

با توجه به تعریف تعدیل‌شده‌ی بیکاری (ILO، ۱۹۹۲) نتایج ما نشان می‌دهد، زنانی که به‌دلایل موجه در جستجوی کار نبوده‌اند (به‌دلیل این که یا کاری یافته‌اند و قرار است آن را دیرتر شروع کنند و یا این که در انتظار بازگشت به شغل قبلی خود هستند) طول مدت بیکاری کم‌تری نسبت به زنانی که در مؤسسات کاریابی ثبت‌نام کرده‌اند و یا با کارفرما تماس داشته‌اند و یا روش‌های دیگری را برای جستجوی کار انتخاب کرده‌اند، دارند. مردانی که به دنبال منابع مالی و امکاناتی برای شروع فعالیت خود اشتغالی هستند و یا تقاضای جواز کسب یا پروانه‌ی کار دارند، طول مدت بیکاری بیش‌تری نسبت به بقیه‌ی افراد دارند که دلیل

www.SID.ir

Archive OF SID

جدول ۲. متغیرهای کمکی و کدهای مربوط به آن‌ها

متغیرهای کمکی	کدهای مربوط
سن	متغیر پیوسته
بستگی با سرپرست خانوار	۱: سرپرست خانوار ۲: بقیه‌ی اعضای خانوار
وضع مهاجرت	۱: مهاجر ۲: غیر مهاجر
وضع تحصیلات	۱: در حال تحصیل بودن ۲: در حال تحصیل نبودن
محل زندگی	۱: شهری ۲: روستایی
جنسیت	۱: مرد ۲: زن
سطح تحصیلات	۱: زیر دیپلم ۲: دیپلم یا پیش‌دانشگاهی ۳: تحصیلات دانشگاهی
وضع زناشویی	۱: ازدواج کرده ۲: طلاق گرفته یا بیوه ۳: هرگز ازدواج نکرده
تجربه‌ی کار قبلی	۱: تجربه داشتن ۲: تجربه نداشتن
وضع فعالیت قبلی	۱: شاغل ۲: در حال تحصیل ۳: بقیه‌ی موارد
روش جستجوی کار	۱: جستجو نکردن ۲: پرس و جو از دوستان و آشنایان ۳: جستجوی منابع مالی برای شروع فعالیت خود اشتغالی و یا تقاضای جواز کسب ۴: درج آگهی در روزنامه یا مطالعه‌ی آگهی‌های استخدام ۵: بقیه‌ی موارد
تعداد افراد شاغل در خانوار	۱: ۰ ۲: ۱ ۳: $\geqslant$ ۲
تعداد افراد در خانوار	۱: ۲ — ۱ ۲: ۴ — ۳ ۳: $\geqslant$ ۵

www.SID.ir

Archive OF SID

جدول ۳. برآورد پارامترها و انحراف معیار در مدل لوجیت تجمعی برای طول مدت بیکاری به‌عنوان پاسخ ترتیبی برای داده‌های نیروی کار (مقادیری که با * مشخص شده‌اند، در سطح ۵ درصد معنی‌دارند)

پارامترها	کل جمعیت		مردان		زنان	
	برآورد	انحراف معیار	برآورد	انحراف معیار	برآورد	انحراف معیار
$\alpha_{1,1}$	−۰/۱۹۴۵	۰/۳۳۳	۰/۲۹۲	۰/۴۲۹	۰/۶۱۰	۰/۶۷۳
$\alpha_{1,2}$	۰/۸۲۰*	۰/۳۲۹	۱/۳۸۳*	۰/۴۳۱	۱/۳۵۷*	۰/۶۹۰
$\alpha_{1,3}$	۱/۸۶۹*	۰/۳۳۷	۲/۳۱۸*	۰/۴۳۷	۲/۸۱۴*	۰/۶۹۹
$\alpha_{1,4}$	۲/۹۶۳*	۰/۳۳۵	۳/۳۷۷*	۰/۴۳۳	۴/۰۱۵*	۰/۷۱۳
سن	−۰/۰۶۳*	۰/۰۰۷	−۰/۰۵۶*	۰/۰۰۸	−۰/۱۱۴*	۰/۰۲۰
بستگی با سرپرست خانوار						
مبنا: بقیه‌ی اعضای خانوار	–	–	–	–	–	–
سرپرست خانوار	۰/۷۹۱*	۰/۱۹۹	۰/۵۰۰*	۰/۲۳۱	۰/۵۷۲	۰/۶۶۵
وضع مهاجرت						
مبنا: غیر مهاجر	–	–	–	–	–	–
مهاجر	۰/۶۰۶*	۰/۱۹۱	۰/۵۴۸*	۰/۲۲۶	۰/۹۴۸*	۰/۳۵۲
وضع تحصیلات						
مبنا: در حال تحصیل نبودن	–	–	–	–	–	–
در حال تحصیل بودن	۰/۷۷۳*	۰/۲۷۱	۰/۸۵۳*	۰/۳۵۹	۰/۶۴۳	۰/۴۱۱
جنسیت						
مبنا: زن	–	–	–	–	–	–
مرد	۰/۲۹۷*	۰/۱۲۳	–	–	–	–
سطح تحصیلات						
مبنا: تحصیلات دانشگاهی	–	–	–	–	–	–
زیر دیپلم	−۰/۰۵۷	۰/۱۴۷	−۰/۱۷۴	۰/۱۹۸	−۰/۱۶۶	۰/۳۰۱
دیپلم یا پیش دانشگاهی	−۰/۳۴۸*	۰/۱۳۶	−۰/۵۱۰*	۰/۱۹۷	−۰/۳۹۵	۰/۲۰۲
وضع زناشویی						
مبنا: هرگز ازدواج نکرده	–	–	–	–	–	–
ازدواج کرده	۰/۲۲۴	۰/۱۴۰	۰/۴۴۷*	۰/۲۰۰	۰/۰۴۹	۰/۲۰۸
طلاق‌گرفته یا بیوه‌شده	۰/۳۶۱	۰/۳۸۲	۰/۹۶۶	۰/۵۷۹	۰/۳۵۸	۰/۵۹۵
تجربه‌ی کار قبلی						
مبنا: تجربه نداشتن	–	–	–	–	–	–
تجربه داشتن	۰/۵۴۱*	۰/۱۸۴	۰/۶۸۸*	۰/۲۳۳	۰/۴۵۶	۰/۳۰۴
وضع فعالیت قبلی						
مبنا: بقیه‌ی موارد	–	–	–	–	–	–
شاغل	۰/۴۰۴*	۰/۱۹۵	۰/۴۱۵	۰/۲۳۸	۰/۲۶۹	۰/۳۷۵
در حال تحصیل	−۰/۴۰۲*	۰/۱۷۹	−۰/۳۲۰	۰/۲۴۴	−۰/۸۵۹*	۰/۳۰۷

www.SID.ir

Archive OF SID

ادامه‌ی جدول ۳. برآورد پارامترها و انحراف معیار در مدل لوجیت تجمعی برای طول مدت بیکاری به‌عنوان پاسخ ترتیبی برای داده‌های نیروی کار (مقادیری که با * مشخص شده‌اند، در سطح ۵ درصد معنی‌دارند)

پارامترها	کل جمعیت		مردان		زنان	
	برآورد	انحراف معیار	برآورد	انحراف معیار	برآورد	انحراف معیار
روش جستجوی کار						
مبنا: بقیه‌ی موارد	–	–	–	–	–	–
جستجو نکردن	۰/۴۶۰	۰/۲۸۵	۰/۰۱۴	۰/۳۳۴	۱/۳۳۱*	۰/۶۲۶
پرس و جو از دوستان و آشنایان	−۰/۳۰۱	۰/۲۰۰	−۰/۶۹۴*	۰/۲۵۶	۰/۴۰۰	۰/۳۴۹
شروع خود اشتغالی و یا تقاضای جواز کسب	−۰/۷۷۰*	۰/۲۳۹	−۱/۱۱۶*	۰/۲۹۸	−۰/۱۰۵	۰/۴۳۳
درج یا مطالعه‌ی آگهی در روزنامه	−۰/۸۳۴*	۰/۲۰۴	−۱/۳۳۸*	۰/۲۷۰	−۰/۰۷۴	۰/۳۳۸
$-2\log L_{\text{Null}}$	۳۳۸۵۱۲۶/۰		۲۴۰۹۸۹۵/۴		۸۹۷۳۸۶/۸	
$-2\log L_{\text{Full}}$	۳۱۴۷۱۴۹/۶		۲۲۷۷۳۵۹/۴		۸۳۵۰۳۲/۱	

آن به خاطر فرایند طولانی‌تر به انجام رسیدن اقدامات برای این نوع کارها است. این نتایج برای مردان و کل جمعیت مورد مطالعه که از طریق درج آگهی در روزنامه‌ها یا مطالعه‌ی آگهی‌های استخدام به دنبال کار می‌گردند نسبت به بقیه‌ی افراد مشابه است.

نتایج مدل‌بندی انتقالی وضع فعالیت اقتصادی به‌عنوان پاسخ اسمی در زمان دوم به‌شرط طول مدت بیکاری به‌عنوان پاسخ زمان اول در جدول ۴ ارائه شده است. نتایج این جدول که بر اساس مدل لوجیت تعمیم‌یافته (وضع غیر فعال به‌عنوان مبنا) بدست آمده است، اثرات متغیرهای کمکی مختلف و طول مدت بیکاری را به‌عنوان پاسخ زمان قبل، روی وضع فعالیت اقتصادی نشان می‌دهد.

به‌منظور بررسی وجود همبستگی بین دو پاسخ، برای کل جامعه‌ی مورد بررسی دو مدل را به‌صورت زیر تعریف می‌کنیم. مدلی شامل همه‌ی اثرات متغیرهای کمکی و پاسخ زمان قبل را مدل II و مدلی را با حذف اثرات پاسخ زمان قبل مدل I می‌نامیم. با مقایسه‌ی این دو مدل با استفاده از آماره‌ی

$$G^2 = -2(\log L_{\text{Full}_I} - \log L_{\text{Full}_{II}}) = ۱۷۰۴۵/۴$$

که دارای توزیع خی‌دو (تحت فرض این‌که مدل I مدل درست است) با ۲ درجه‌ی آزادی است، می‌بینیم که اثر معنی‌دار بسیار قوی از پاسخ ترتیبی زمان قبل روی پاسخ اسمی زمان حال وجود دارد.

اکنون به بررسی اثر بقیه‌ی متغیرهای کمکی روی بخت بیکار ماندن، نسبت به غیر فعال شدن می‌پردازیم. این نتایج همچنین برای مردان و زنان به طور مجزا بیان شده‌اند. برای مردان، اثر سن

www.SID.ir

Archive OF SID

جدول ۴. برآورد پارامترها و انحراف معیار مدل لوجیت تعمیم‌یافته برای وضع فعالیت اقتصادی (بیکار) در نیروی کار (مدل I: مدل (۴) بدون در نظر گرفتن پاسخ زمان قبل، مدل II: مدل (۴))، مقادیری که با * مشخص شده‌اند، در سطح ۵ درصد معنی‌دارند.

پارامترها	کل جمعیت				مردان		زنان	
	مدل I		مدل II					
	برآورد	انحراف معیار	برآورد	انحراف معیار	برآورد	انحراف معیار	برآورد	انحراف معیار
$\alpha_{۲,۱}$	۰/۱۷۳	۰/۳۱۴	۰/۳۳۳	۰/۳۲۶	۱/۵۶۰*	۰/۳۷۴	−۰/۰۱۳	۰/۴۹۳
سن	−۰/۰۱۸	۰/۰۱۱	−۰/۰۱۷	۰/۰۱۱	−۰/۰۳۵*	۰/۰۱۴	۰/۰۱۲	۰/۰۱۸
بستگی با سرپرست خانوار								
مبنا: بقیه اعضای خانوار	–	–	–	–	–	–	–	–
سرپرست خانوار	۰/۴۹۰	۰/۲۶۳	۰/۴۶۵	۰/۲۶۸	−۰/۲۶۹	۰/۳۴۸	۰/۷۰۴	۰/۷۰۴
وضع تحصیلات								
مبنا: در حال تحصیل نبودن	–	–	–	–	–	–	–	–
در حال تحصیل بودن	−۲/۰۴۳*	۰/۳۹۶	−۲/۰۷۴*	۰/۳۹۰	−۲/۹۸۴*	۰/۴۲۹	−۰/۷۳۷	۰/۴۶۰
محل زندگی								
مبنا: روستا	–	–	–	–	–	–	–	–
شهر	۰/۴۶۷*	۰/۱۶۹	۰/۴۷۷*	۰/۱۷۱	۰/۷۴۵*	۰/۲۱۳	۰/۱۰۸	۰/۲۵۹
جنسیت								
مبنا: زن	–	–	–	–	–	–	–	–
مرد	۱/۰۲۷*	۰/۱۶۰	۰/۹۹۷*	۰/۱۶۶	–	–	–	–
وضع زناشویی								
مبنا: هرگز ازدواج نکرده	–	–	–	–	–	–	–	–
ازدواج کرده	−۰/۶۳۷*	۰/۱۷۹	−۰/۶۵۲*	۰/۱۸۱	۰/۵۱۳	۰/۳۲۴	−۱/۰۴۹*	۰/۲۵۰
طلاق‌گرفته یا بیوه	−۰/۵۹۲	۰/۵۶۵	−۰/۶۰۴	۰/۵۶۸	−۰/۵۵۶	۰/۸۱۸	−۱/۱۰۰	۰/۸۵۳
$\gamma_۱(Y_۱)$	–	–	−۰/۰۷۲	۰/۰۵۱	−۰/۰۷۲	۰/۰۶۱	−۰/۰۹۰	۰/۰۸۹

معنی‌دار است اما این اثر برای کل جامعه و زنان معنی‌دار نیست. این اثر نشان می‌دهد که هر چه سن مردان افزایش می‌یابد، بخت بیکار ماندن آن‌ها نسبت به غیر فعال شدنشان کاهش می‌یابد. اثر سرپرست خانوار بودن نیز معنی‌دار نیست. نتایج همچنین نشان می‌دهند که اثر وضع تحصیل در تابستان برای همه‌ی افراد و مردان معنی‌دار است و این اثر نشان می‌دهد که بخت برآورد شده‌ی بیکار ماندن نسبت به غیر فعال شدن برای افرادی که در حال تحصیل هستند نسبت به افرادی که تحصیل نمی‌کنند، کم‌تر است. همچنین این بخت در مناطق شهری نسبت به مناطق روستایی برای مردان و کل جمعیت به نسبت بالاتر است. نتایج همچنین نشان می‌دهد که بخت برآورد شده‌ی بیکار ماندن نسبت به غیر فعال شدن برای مردان بیش‌تر است. این به این معناست که احتمال باقی ماندن در نیروی کار (به‌عنوان بیکار) در طول مدت مطالعه برای مردان بیکار نسبت به زنان بیکار بیش‌تر است. اثر وضع زناشویی نیز نشان می‌دهد که بخت برآورد شده‌ی بیکار ماندن نسبت به غیر فعال شدن برای زنانی که ازدواج کرده‌اند نسبت به زنانی که ازدواج نکرده‌اند کم‌تر است. این به این معناست که زنانی که ازدواج کرده‌اند با احتمال بالایی غیر فعال می‌شوند. اثر وضع

www.SID.ir

Archive OF SID

ادامه‌ی جدول ۴. برآورد پارامترها و انحراف معیار مدل لوجیت تعمیم‌یافته برای وضع فعالیت اقتصادی (شاغل) در نیروی کار (مدل I: مدل (۴) بدون در نظر گرفتن پاسخ زمان قبل، مدل II: مدل (۴))، مقادیری که با * مشخص شده‌اند، در سطح ۵ درصد معنی‌دارند.

پارامترها	کل جمعیت				مردان		زنان	
	مدل I		مدل II					
	برآورد	انحراف معیار	برآورد	انحراف معیار	برآورد	انحراف معیار	برآورد	انحراف معیار
$\alpha_{۲,۲}$	−۰/۰۸۲	۰/۳۳۷	۰/۳۰۷	۰/۳۴۳	۲/۷۸۴*	۰/۳۴۷	−۱/۱۵۴	۰/۷۵۴
سن	−۰/۰۴۸*	۰/۰۱۰	−۰/۰۴۲*	۰/۰۱۱	−۰/۰۶۱*	۰/۰۱۳	۰/۰۲۸	۰/۰۲۳
بستگی با سرپرست خانوار								
مبنا: بقیه اعضای خانوار	–	–	–	–	–	–	–	–
سرپرست خانوار	۰/۸۹۳*	۰/۲۴۸	۰/۷۸۷*	۰/۲۵۷	۰/۱۳۶	۰/۳۳۲	−۰/۸۸۱	۱/۰۷۹
وضع تحصیلات								
مبنا: در حال تحصیل نبودن	–	–	–	–	–	–	–	–
در حال تحصیل بودن	−۲/۲۰۹*	۰/۳۷۸	−۲/۲۳۴*	۰/۳۹۲	−۲/۵۱۰*	۰/۴۰۴	−۲/۱۱۴*	۰/۶۸۶
محل زندگی								
مبنا: روستا	–	–	–	–	–	–	–	–
شهر	−۰/۲۰۱	۰/۱۶۷	−۰/۱۶۷	۰/۱۶۹	۰/۰۳۳	۰/۱۹۸	−۰/۲۲۷	۰/۳۸۰
جنسیت								
مبنا: زن	–	–	–	–	–	–	–	–
مرد	۲/۳۱۰*	۰/۲۰۴	۲/۱۸۵*	۰/۲۰۶	–	–	–	–
وضع زناشویی								
مبنا: هرگز ازدواج نکرده	–	–	–	–	–	–	–	–
ازدواج کرده	۰/۲۹۶	۰/۱۹۴	۰/۲۴۰	۰/۱۹۸	۱/۳۸۵*	۰/۳۲۴	−۰/۹۱۶*	۰/۳۷۸
طلاق‌گرفته یا بیوه	۰/۹۰۲	۰/۵۵۰	۰/۸۰۵	۰/۵۶۸	۰/۸۷۲	۰/۷۲۵	۰/۸۲۴	۰/۶۹۳
$\gamma_{۲}(Y_{۱})$	–	–	−۰/۲۲۳*	۰/۰۵۴	−۰/۲۳۵*	۰/۰۶۰	−۰/۱۷۳	۰/۱۳۰
$-۲logL_{\text{Null}}$	۲۳۵۹۳۳۳/۴		۲۳۵۷۰۳۲/۶		۱۵۹۷۹۳۷/۵		۵۹۸۸۳۴/۰۴	
$-۲\log L_{\text{Full}}$	۲۰۸۱۶۴۳/۱		۲۰۶۴۵۹۷/۷		۱۴۵۶۴۹۷/۵		۵۷۴۱۹۲/۱۶	

زناشویی برای مردان معنی‌دار نشده است.

به‌منظور مقایسه‌ی اثر متغیرهای کمکی روی بخت برآورد شده‌ی شاغل شدن نسبت به غیر فعال شدن، نتایج جدول ۴ به‌صورت زیر بیان شده است. اثر سن نشان می‌دهد که بخت برآورد شده‌ی شاغل شدن نسبت به غیر فعال شدن برای افراد مسن کم‌تر است. نتیجه‌ی مشابه برای مردان وجود دارد. اثر سرپرست خانوار بودن نشان می‌دهد که بخت برآورد شده‌ی شاغل شدن نسبت به غیر فعال شدن برای سرپرست خانوار نسبت به بقیه‌ی اعضای خانوار بیش‌تر است. نتایج همچنین نشان می‌دهد که بخت برآورد شده‌ی شاغل شدن نسبت به غیر فعال شدن برای افرادی که درحال تحصیل هستند نسبت به افرادی که تحصیل نمی‌کنند، کم‌تر است. برای اثر جنسیت، نتایج نشان می‌دهد که این بخت برای مردان نسبت به زنان بیش‌تر است. اثر وضع زناشویی نشان می‌دهد که این بخت برای مردانی که ازدواج کرده‌اند نسبت به مردانی که ازدواج نکرده‌اند، بیش‌تر است. همچنین این بخت برای زنانی که ازدواج کرده‌اند نسبت به زنانی که هرگز ازدواج نکرده‌اند، کم‌تر است. اثر طول مدت بیکاری در بهار نیز روی احتمال شاغل شدن نسبت به غیر

www.SID.ir

Archive OF SID

فعال شدن معنی‌دار است در حالی که این اثر روی احتمال بیکار ماندن نسبت به غیر فعال شدن معنی‌دار نبوده است.

در مورد نیکویی برازش مدل برای داده‌هایی که در این مقاله استفاده شده است، شبه R^2ی ناگل‌کرک برای مدل لوجیت تجمعی با حضور ۱۰ متغیر کمکی در مدل ۰٫۲۱ برای کل جمعیت، ۰٫۱۷ برای مردان و ۰٫۲۰ برای زنان است که به طور نسبی برازش خوبی را نشان داده است. کاهش این مقدار برای مردان و زنان در مقایسه با کل جمعیت به‌دلیل کاهش اندازه‌ی نمونه و نگاه کردن صرف به مردان یا زنان بوده است.

ضریب تعیین برای مدل لوجیت انتقالی تعمیم‌یافته مقدار ۰٫۲۱ برای کل جمعیت، ۰٫۱۵ برای مردان و ۰٫۰۷ برای زنان بوده است که به طور نسبی برازش خوبی را برای کل جمعیت و مردان نشان می‌دهد. این کمّیت برای زنان به‌دلیل این که اکثر متغیرهای کمکی در این مدل معنی‌دار نشده‌اند، مناسب نیست.

۵ نتیجه‌گیری

در این مقاله، مدل انتقالی برای پاسخ‌های طولی آمیخته‌ی ترتیبی و اسمی معرفی شد. همچنین از نرم‌افزارهای موجود (برای مثال SAS، رویه‌ی surveylogistic) برای برآورد پارامترهای مدل در داده‌های بیکاری نیروی کار (بهار و تابستان ۱۳۸۵) مرکز آمار ایران استفاده شد. در این داده‌ها، پاسخ ترتیبی در زمان اول، طول مدت بیکاری افراد بیکار در بهار ۱۳۸۵ و پاسخ اسمی در زمان دوم، وضع فعالیت اقتصادی این افراد در تابستان ۱۳۸۵ است. هر عضو خانوار نمونه با وزن نهایی محاسبه‌شده در سه مرحله‌ی ۱) محاسبه‌ی وزن پایه ۲) تعدیل وزن بر اساس بی‌پاسخی ۳) تعدیل وزن بر اساس پیش‌بینی جمعیتی، وزن‌دهی شدند. برای این داده‌ها، به دست آمد که طول مدت بیکاری به‌عنوان پاسخ زمان اول تأثیر معنی‌دار بسیاری روی وضع فعالیت اقتصادی به‌عنوان پاسخ زمان دوم دارد. در این مقاله به این نتیجه دست یافتیم که هر چه طول مدت بیکاری افراد تا بهار ۱۳۸۵ بیش‌تر باشد، احتمال آن که فرد در تابستان ۱۳۸۵ نیز بیکار باقی بماند، بیش‌تر است. لازم به ذکر است که این نتیجه‌گیری از آمار توصیفی به دست آمده است. این نتایج و نتایج مشابه را می‌توان از طریق مدل‌بندی به دست آورد با این تفاوت که مدل‌بندی آماری می‌تواند شدت تأثیرها را نیز بیان کند. به‌عنوان مثال بخت شاغل شدن به غیر فعال شدن یک فرد در تابستان ۱۳۸۵ با طول مدت بیکاری ۱۳-۲۴ ماه در بهار ۱۳۸۵، ۰٫۵۱۲ برابر بخت شاغل شدن به غیر فعال شدن یک فرد در تابستان ۱۳۸۵ با طول مدت بیکاری ۱-۳ ماه در بهار ۱۳۸۵ است (اگر تأثیر بقیه‌ی متغیرهای کمکی را ثابت در نظر بگیریم).

مدل انتقالی معرفی‌شده در این مقاله برای هر تعداد از متغیرهای کمکی قابل تعمیم است. برای داده‌های طولی با بیش از ۲ دوره‌ی زمانی نیز این مدل انتقالی به آسانی می‌تواند بسط داده شود. برای چنین

www.SID.ir

Archive OF SID

داده‌هایی، مدل‌های اثرات تصادفی برای در نظر گرفتن باقی‌مانده‌ی ناهمگنی بین پاسخگوها (آلبرت و فولمن، ۲۰۰۳، هکمن ۱۹۸۱) می‌تواند در نظر گرفته شود. در این صورت می‌توان برای مثال در پیشگوی خطی مدل، اثر تصادفی را به‌صورت خطی در نظر گرفت و به‌شرط این اثر، پاسخ جاری به پاسخ قبلی وابسته خواهد بود. این مدل به مدل اثرات تصادفی انتقالی موسوم است. هر چه برآورد واریانس متغیر اثر تصادفی بیش‌تر باشد ناهمگنی بیش‌تری بین پاسخگویان وجود دارد. در صورت استفاده از مدل اثرات تصادفی انتقالی، دیگر استفاده از نرم‌افزارهای موجود برای برآورد پارامترها امکان‌پذیر نیست. در این حالت می‌توان از الگوریتم‌های بهینه‌سازی مانند الگوریتم نیوتن-رافسون و یا الگوریتم شبه نیوتن-رافسون برای ماکسیمم کردن تابع درست‌نمایی استفاده کرد.

سپاس‌گزاری

از پژوهشکده‌ی آمار برای حمایت مالی این مقاله تشکر می‌نماییم. همچنین از داوران مجله برای توصیه‌های ارزشمند آن‌ها سپاس‌گزاریم.

توضیحات

۱. بیکار به تمام افراد ۱۰ ساله و بیش‌تر اطلاق می‌شود که در هفته‌ی مرجع فاقد کار باشند، در هفته‌ی مرجع و یا هفته‌ی بعد از آن آماده برای کار باشند و در هفته‌ی مرجع و سه هفته قبل از آن جویای کار باشند.

۲. شاغل به تمام افراد ۱۰ ساله و بیش‌تر اطلاق می‌شود که در طول هفته‌ی مرجع، طبق تعریف کار، حد اقل یک ساعت کار کرده و یا بنا به دلایلی به طور موقت کار را ترک کرده باشند.

۳. غیر فعال به تمام افراد ۱۰ ساله و بیش‌تر اطلاق می‌شود که در طول هفته‌ی مرجع طبق تعریف در هیچ یک از دو گروه شاغلان و بیکاران قرار نمی‌گیرد. برای اطلاعات بیش‌تر به نشریه‌ی نتایج آمارگیری از نیروی کار ۱۳۸۵، مرکز آمار ایران مراجعه شود.

مرجع‌ها

Agresti, A. (2002). *Categorical Data Analysis*. Wiley, New York.

Albert, P.S. and Follmann, D.A. (2003). A random effects transition model for longitudinal binary data with informative missingness. *Statist. Neerlandica* **57**, 100-111.

Aldrich, J.H. and Nelson, F.D. (1984). *Linear Probability, Logit and Probit Models*. Sage, California.

Anderson, T.W. and Goodman, L.A. (1957). Statistical inference about Markov Chains. *Ann. Math. Statist.* **28**, 89-110.

www.SID.ir

Archive OF SID

Bartlett, M.S. (1951). The frequency goodness of fit test for probability Chains. Proceedings of the *Cambridge Philosophical Society* **47**, 86-95.

Berridge, D.M. and Dos Santos, D.M. (1996). Fitting a random effects model to ordinal recurrent events using existing software. *J. Stat. Comput. Simul.* **55**, 73-86.

Billingsley, P. (1961). Statistical methods in Markov Chains. *Ann. Math. Statist.* **32**, 12-40.

Chung, H., Park, Y., and Lanza, S.T. (2005). Latent transition analysis with covariates: pubertal timing and substance use behaviours in adolescent females. *Stat. Med.* **24**, 2895-2910.

Cochran, W.G. (1977). *Sampling Techniques*. Wiley, New York.

Cox, D.R. and Snell, E.J. (1989).*Analysis of Binary Data*. Chapman and Hall, London.

Deville, J.C. and Sarndal, C.E. (1992). Calibration estimators in survey sampling. *J. Amer. Statist. Assoc.* **87**, 376-382.

Diggle, P.J., Heagerty, P., Liang, K.Y., and Zeger, S.L. (2002). *Analysis of Longitudinal Data*. Oxford University Press, Oxford.

Garber, A.M. (1989). A discrete-time model of the acquisition of antibiotic-resistant infections in hospitalized patients. *Biometrics* **45**, 797-816.

Harville, D.A. and Mee, R.W. (1984). A mixed model procedure for analyzing ordered categorical data. *Biometrics* **40**, 393-408.

Heckman, J.J. (1981). *Statistical Models for Discrete Panel Data*. In Structural Analysis of Discrete Data with Econometric Application, eds. C.F. Manski and D. McFadden, MIT Press, Cambridge.

Koehler, K.J. (1986). Goodness-of-fit tests for log-linearmodels in sparse contingency tables. *J. Amer. Statist. Assoc.* **81**, 483-493.

Koehler, K. and Larntz, K. (1980) . An empirical investigation of goodness-of-fit statistics for sparse multinomials. *J. Amer. Statist. Assoc.* **75**, 336-344.

Larntz, K. (1978). Small-sample comparisons of exact levels for chi-squared goodness-of-fit statistics. *J. Amer. Statist. Assoc.* **73**, 253-263.

Leon, A.C., Hedeker, D., and Teres, J.J. (2007). Bias reduction in effectiveness analyses of longitudinal ordinal doses with a mixed-effects propensity adjustment. *Stat. Med.* **26**, 110-123.

Liang, K.Y. and Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73**, 13-22.

Liang, K.Y., Zeger, S.L., and Qaqish, B.F. (1992). Multivariate regression analyses for categorical data (with Discussion). *J. Roy. Statist. Soc.* **54**, 3-40.

© ۱۳۸۷ پژوهشگاه‌ی آمار www.SID.ir

Archive OF SID

McFadden, D. (1974). The measurement of urban travel demand. *J. Public Econ.* **3**, 303-328.

Nagelkerke, N.J.D. (1991). A note on a general definition of the coefficient of determination. *Biometrika* **78**, 691-692.

Sengul, T.K., Stoffer, D.S., and Day, N.L. (2007). A residuals-based transition model for longitudinal analysis with estimation in the presence of missing data. *Stat. Med.* **26**: 3330-3341.

Stiratelli, R., Laird, N., and Ware, J.H. (1984). Random-effects models for serial observations with binary response. *Biometrics* **40**, 961-971.

Ten Have, T.R., Landis, J.R., and Hartzel, J. (1996). Population-averaged and cluster specific models for clustered ordinal response data. *Stat. Med.* **15**, 2573-2588.

Tutz, G. (2005). Modelling of repeated ordered measurements by isotonic sequential regression. *Statist. Mod.* **5**, 269-287.

Verbeke, G. and Lesaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population. *J. Amer. Statist. Assoc.* **91**, 217-221.

Vermunt, J.K., Rodrigo, M.F., and Ato-Garcia, M. (2001). Modeling joint and marginal distribution in the analysis of categorical panel data. *Sociol. Methods Res.* **30**, 170-196.

زهرا رضایی قهرودی
پژوهشکده‌ی آمار،
شماره‌ی ۱۴۵، خیابان سرتیپ فکوری،
خیابان باباطاهر، خیابان دکتر فاطمی،
تهران، ایران.
رایانشانی: *z_rezaei@srtc.ac.ir*

فاطمه هرندی
پژوهشکده‌ی آمار،
شماره‌ی ۱۴۵، خیابان سرتیپ فکوری،
خیابان باباطاهر، خیابان دکتر فاطمی،
تهران، ایران.
رایانشانی: *harandi@srtc.ac.ir*

مجتبی گنجعلی
گروه آمار، دانشکده‌ی علوم ریاضی،
دانشگاه شهید بهشتی،
تهران، ایران.
و
پژوهشکده‌ی آمار،
شماره‌ی ۱۴۵، خیابان سرتیپ فکوری،
خیابان باباطاهر، خیابان دکتر فاطمی،
تهران، ایران.
رایانشانی: *m-ganjali@sbu.ac.ir*

www.SID.ir