

عوامل تأثیرگذار بر استفاده از خریدهای اینترنتی در خانوارهای ایرانی

اشکان شباک^۱، مریم میرزایی^۲، حامد لرونند^۳

تاریخ تایید: ۰۰/۱/۴

تاریخ دریافت: ۹۹/۹/۲

DOI: 10.22034/JCSC.2021.137389.2270

چکیده

با توجه به اهمیت گسترش استفاده از فناوری اطلاعات در زمینه‌های اجتماعی و اقتصادی جامعه و بهره‌گیری از آن در راستای پیشرفت و توسعه، داشتن تحلیل درست و دقیق از چگونگی استفاده افراد جامعه از فناوری اطلاعات و ارتباطات و متغیرهای تأثیرگذار بر آن ضروری است. در همین راستا در این مقاله تلاش می‌شود تا با استفاده از نتایج آمارگیری استفاده‌ی خانوارها و افراد از فناوری اطلاعات و ارتباطات مرکز آمار ایران (سال ۱۳۹۶) و روش‌های مدل‌بندی رگرسیونی، به بررسی متغیرهای تأثیرگذار در استفاده از خریدهای اینترنتی، پرداخته شود. در این مطالعه خریدهای اینترنتی به صورت خرید y کالا از m کالا بیان می‌شود. خریدهای اینترنتی (تعداد کالاهای خریداری شده) شامل خرید در حیطه ملزومات زندگی، سرگرمی و خدمات آموزشی و فرهنگی می‌شود که این به عنوان متغیر پاسخ در نظر گرفته می‌شود و تأثیرگذاری عواملی مانند شهری یا روستایی بودن، جنسیت، سن، وضع سواد و مدرک تحصیلی و وضع فعالیت بر آن بررسی می‌شود. البته شایان گفتن است که با توجه به ساختار سلسله‌مراتبی داده‌های مورد استفاده و نبود فرض استقلال مشاهدات از مدل رگرسیون چندسطحی با پاسخ دوجمله‌ای استفاده شده است. در این مقاله با اجرای شبیه‌سازی مناسب بودن مدل رگرسیون دوسطحی با پاسخ دوجمله‌ای برای داده‌های سلسله‌مراتبی تأیید شد و با اجرای مدل رگرسیون دوسطحی با پاسخ دوجمله‌ای روی داده‌ها و مقایسه با رگرسیون معمولی با پاسخ دوجمله‌ای با توجه به مقدار AIC مدل رگرسیون دوسطحی با پاسخ دو جمله‌ای بهتر عمل کرد. **واژه‌های کلیدی:** رگرسیون چندسطحی، ساختار سلسله‌مراتبی، فناوری اطلاعات و ارتباطات.

۱ استادیار، هیات علمی پژوهشکده آمار، تهران (نویسنده مسئول)؛ a.shabbak@gmail.com

۲ کارشناس پژوهشی پژوهشکده آمار، تهران. mmaryammirzaie73@gmail.com

۳ استادیار دانشکده علوم ریاضی، دانشگاه صنعتی اصفهان؛ lorvandhamed@iut.ac.ir

مقدمه

در سال‌های اخیر فناوری اطلاعات و ارتباطات که یکی از شاخص‌های مهم توسعه‌یافتگی کشورها به شمار می‌رود، تحولات گسترده‌ای را در همه‌ی زمینه‌ها و امور بشر به همراه داشته است. فرآیند شکل‌گیری و انتشار اینترنت، منجر به تولد ساختار جدید رسانه‌های شبکه‌ای و شکل‌دهی به فرهنگ کاربران این شبکه شد. جهان‌شمولی زبان دیجیتال و منطق کاملاً شبکه‌ای سیستم ارتباطی اینترنت، زمینه‌های تکنولوژیک ارتباط افقی و جهانی را فراهم کرد و به دوره‌ی مخاطبان توده، پایان داد و با ایجاد شبکه‌های تعاملی ارتباطات، تحولی عظیم در پارادایم ارتباطات از «مونولوگ» به «دیالوگ» یا از «جریان یک سویه»^۱ ارتباطات به «جریان دو سویه»^۲ و یا «مخاطب منفعل»^۳ به «مخاطب فعال»^۴ پدید آورد. «فرانسوا سابا»^۵ در سال ۱۹۸۵ چنین پیش‌بینی کرد که:

«رسانه‌های دیجیتالی جدید [و پیشاپیش همه آنها، اینترنت]، مخاطبان پراکنده و متمایزی را تعیین می‌کنند که با وجود این که از نظر تعداد پُرشمارند، از نظر هم‌زمانی و یک‌دستی پیام‌هایی که دریافت می‌کنند، دیگر به عنوان مخاطبان انبوه به شمار نمی‌آیند. رسانه‌های نوین، دیگر به مفهوم سنتی کلمه، رسانه‌های همگانی نیستند که پیام‌های محدودی را برای انبوهی از مخاطبان متجانس پخش کنند. به دلیل تعدد پیام‌ها و منابع، مخاطبان، قدرت انتخاب بیشتری پیدا کرده‌اند. مخاطبان مورد نظر دوست دارند پیام‌های خود را انتخاب کنند و این امر، دسته‌بندی آنان را عمیق‌تر می‌سازد و رابطه فردی میان فرستنده و گیرنده را افزایش می‌دهد» (کاستلز، ۱۳۸۵، ج ۱: ۳۹۵).

به طور کلی، صاحب‌نظران حوزه ارتباطات جمعی، ویژگی‌های منحصر به فردی را در مقایسه با رسانه‌های قدیمی‌تر برای اینترنت مطرح کرده‌اند که از جمله آنها می‌توان به مواردی چون: تعاملی بودن یا دوسویه‌گی، بی‌زمان و بی‌مکان شدگی، ظرفیت نامحدود، انعطاف‌پذیری زیاد، امکان شخصی‌سازی اشاره نمود.

بر اساس آخرین آمار منتشر شده^۶، در حال حاضر (۳۰ سپتامبر ۲۰۲۰) کاربران اینترنت، آماری نزدیک به ۵ میلیارد نفر از جمعیت ۷/۷ میلیارد نفری جهان را به خود اختصاص داده‌اند

^۱ One-Way Flow

^۲ Tow-Way Flow

^۳ Passive Audience

^۴ Active Audience

^۵ F. Saba

^۶ <http://www.internetworldstats.com/stats.htm>

که از این آمار، ۵۶/۱ درصد در آسیا و اقیانوسیه، ۱۴/۸ درصد در اروپا، ۱۶/۳ درصد در قاره آمریکا و ۱۲/۸ درصد در آفریقا زندگی می‌کنند.

بنابر جدیدترین گزارش‌های رسمی^۱، در حال حاضر ایران با دارا بودن ۶۷ میلیون نفر کاربر اینترنت، رتبه نخست را در استفاده از اینترنت در منطقه خاورمیانه به خود اختصاص داده است که این رقم معادل ۳۶/۹ درصد کاربران منطقه خاورمیانه و ۱/۳ درصد کل کاربران اینترنت در سطح دنیا می‌باشد. این در حالی است که تعداد کاربران اینترنت در ایران در سال ۲۰۰۰ تنها ۲۵۰ هزار نفر بوده است. رشد چشمگیری که در استفاده از اینترنت دیده می‌شود، ظرفیت و توانایی آن را به عنوان یک عنصر اساسی در امر ارتباطات و مبادله اطلاعات در عصر حاضر نشان می‌دهد.

با توجه به اهمیت تأثیر اینترنت بر امور اجتماعی و فرهنگی و همچنین بر رفتار افراد و خانوارها، در این مقاله تلاش شده است تا بر اساس اطلاعات طرح آمارگیری «استفاده‌ی خانوارها و افراد از فناوری اطلاعات و ارتباطات مرکز آمار ایران در سال ۱۳۹۶ برای ۳۱ استان کشور» به بررسی عوامل تأثیرگذار بر استفاده از اینترنت در زمینه‌های اجتماعی و فرهنگی مانند خرید لوازم سرگرمی، خرید ملزومات زندگی و خرید خدمات آموزشی فرهنگی پرداخته شود. از آنجایی که داده‌های به کارگرفته شده از طرح آمارگیری «استفاده‌ی خانوارها و افراد از فناوری اطلاعات و ارتباطات مرکز آمار ایران در سال ۱۳۹۶ برای ۳۱ استان کشور» دارای ساختار سلسله‌مراتبی هستند، به کار بردن رگرسیون معمولی به دلیل در نظر گرفتن دو فرض استقلال مشاهدات و ناهمبستگی خطای مدل (Wong and Mason, 1985: 391) و نادیده گرفتن تغییراتی که به واسطه‌ی سطوح مختلف ایجاد می‌شود، مناسب نیست (Maas and Hox, 2004: 2). معمولاً در بررسی نحوه استفاده افراد از فناوری اطلاعات، مطالعات مختلف نشان می‌دهند که نحوه استفاده از این ابزار در گروه‌های مختلف جمعیتی یا نواحی مختلف یک کشور یکسان نیست. در طرح آمارگیری مرکز آمار نیز میزان استفاده خانوارها از فناوری اطلاعات و ارتباطات در استان‌های مختلف کشور متفاوت است. بنابراین با در نظر گرفتن این تفاوت‌ها است که داده‌ها را سلسله‌مراتبی در نظر گرفته و از مدل رگرسیون چندسطحی برای برازش داده‌ها استفاده می‌شود. در این مقاله سطح اول افراد و سطح دوم استان می‌باشد. بنابراین ابتدا مقدمات لازم و رگرسیون دوسطحی با پاسخ دوجمله‌ای مطرح می‌شود و سپس مدل برازنده روی داده‌های شبیه‌سازی شده و داده‌های واقعی برازش داده می‌شود.

^۱ www.internetworldstats.com

به طور کلی هدف این مقاله بررسی میزان تاثیرگذاری متغیرهای «شهری یا روستایی بودن، جنسیت، سن، وضع سواد و مدرک تحصیلی و وضع فعالیت افراد» بر تعداد خرید اینترنتی در حیطه‌های مختلف اشاره شده در فوق است. به این دلیل که متغیر پاسخ بیانگر خرید y کالا از بین m کالا است از رگرسیون دو جمله‌ای استفاده شده است. همچنین در این مقاله چگونگی گروه سنی و تحصیلات، شغل و جنسیت، افراد در خرید اینترنتی در حیطه‌های اشاره شده در فوق نیز، بررسی می‌شود. در این مقاله از نرم‌افزار R نسخه ۳,۵,۰ برای انجام تحلیل‌ها استفاده شده است.

پیشینه تحقیق

یکی از جنبه‌های استفاده از اینترنت به‌کارگیری آن جهت انجام خریدهای غیر حضوری در حیطه‌های مختلف است. از آنجا که خرید اینترنتی در ایران موضوعی نوپاست و حجم زیادی از خرید و فروش در سراسر دنیا، به صورت الکترونیکی انجام می‌پذیرد، کشور ما نیز در آینده‌ای نزدیک ناگزیر از پذیرش و بکارگیری آن خواهد بود. بنابراین شناسایی عوامل مؤثر بر این پدیده می‌تواند به پذیرش بیشتر آن توسط مشتریان کمک نموده و بستر مناسبی برای حرکت به سوی آن فراهم آورد (الفت و همکاران، ۱۳۹۰: ۷). رندی و همکاران (۱۳۹۳: ۴۳) عوامل مؤثر بر رفتار خرید مشتریان اینترنتی در ایران را با توجه به مدل پذیرش فناوری بررسی کردند، آنها ۱۲ عامل تاثیرگذار بر رفتار خرید مشتریان اینترنتی را شناسایی کردند. از همین روی، مرکز آمار ایران در سال‌های ۱۳۸۷ و ۱۳۸۹ و سپس در سال‌های ۱۳۹۴ و ۱۳۹۶ به صورت تخصصی‌تر و جامع‌تر به جمع‌آوری اطلاعات در زمینه‌های مختلف فناوری اطلاعات و ارتباطات در قالب آمارگیری «استفاده‌ی خانوارها و افراد از فناوری اطلاعات و ارتباطات» پرداخته است. در این آمارگیری، داده‌های لازم در چارچوب آمارگیری نمونه‌ای گردآوری شده‌اند و تنها به تحلیل نمودار و جداول توصیفی بسنده شده است و تحلیل خاصی روی داده‌ها و ارتباط میان آن‌ها صورت نگرفته است. قدمت نفوذ اینترنت در زندگی بشر به کمتر از سه دهه برمی‌گردد. (Salman and Hasim 2011: 2) و Rogers (1995)، مدل نفوذ فناوری در صنایع و سازمان‌های مختلف را بررسی کردند. در همین راستا McCoy and Mbarika (2005: 2)، با پیروی از مطالعه راجرز الگوی نفوذ اینترنت را در کشورهای توسعه یافته و در حال توسعه مورد تجزیه و تحلیل قرار دادند. همچنین Lawson et al. (2002: 2) تأثیر نفوذ اینترنت را بر سود اقتصادی با تأکید بر تجارت الکترونیک در بلند مدت

بررسی کردند. بابایی فرد و حیدریان (۱۳۹۷: ۵۶) به بررسی تأثیر اینترنت بر سرمایه‌ی فرهنگی و شکل‌گیری سرمایه‌ی فرهنگی آنلاین در فضای مجازی پرداختند. آنها با استفاده از تحلیل‌های رگرسیونی نشان دادند که متغیرهای «مهارت در بهره‌گیری از اینترنت»، «استفاده از سایت‌های علمی اینترنت»، «خرید اینترنتی کالاهای فرهنگی» و «استفاده از سایت‌های هنری اینترنت» بر «سرمایه‌ی فرهنگی» را تأثیر مستقیم دارند. نتایج آنها نشان داد که بین ابعاد گوناگون بهره‌گیری از اینترنت، یعنی «مدت و میزان استفاده از اینترنت»، «مهارت در بهره‌گیری از اینترنت»، «استفاده از سایت‌های علمی اینترنت»، «استفاده از سایت‌های هنری اینترنت»، «خرید اینترنتی کالاهای فرهنگی» و «استفاده از شبکه‌های اجتماعی مجازی» با «سرمایه‌ی فرهنگی» رابطه‌ی معنادار و مستقیم وجود دارد. مهدی‌زاده و عنبرین (۱۳۸۸: ۱۷) به بررسی رابطه‌ی میزان و نحوه استفاده از اینترنت و هویت فرهنگی جوانان پرداختند. مطالعه آنها نشان داد که اولاً بین هویت فرهنگی کاربران و غیر کاربران اینترنت تفاوت وجود دارد، ثانیاً این میزان و نحوه استفاده از اینترنت با هویت فرهنگی کاربران در رابطه است. یعنی به هر میزان که استفاده از سایتها و مطالب غیر بومی بیشتر می‌شود، هویت فرهنگی تضعیف می‌گردد. همچنین خانیکی و عابدی (۱۳۹۳: ۳۶) تأثیر اینترنت را بر نگرش فرهنگی روحانیون بررسی کردند. آنها با استفاده از روشهای آماری نشان دادند بین استفاده از اینترنت (در دو بُعد میزان استفاده و نوع محتوای مورد استفاده) با نگرش فرهنگی روحانیون ارتباط معناداری وجود دارد؛ بدین ترتیب که نگرش فرهنگی روحانیونی که به میزان زیادتری از اینترنت استفاده می‌کنند، بیشتر به سمت نوگرایی تمایل دارد و بر عکس هر چه میزان استفاده از اینترنت کم باشد، نگرش فرهنگی آنها بیشتر به سمت سنت‌گرایی میل می‌کند.

روش‌شناسی

• مدل‌سازی عوامل تأثیرگذار بر استفاده از اینترنت در تعداد خرید لوازم در

حیطه‌های مختلف سرگرمی، ملزومات زندگی و خدمات آموزشی فرهنگی

بررسی عوامل تأثیرگذار بر استفاده از اینترنت جهت انجام خرید در حیطه‌های مختلف می‌تواند یک معیار سنجش برای بررسی پیشرفت یک کشور در زمینه فناوری اطلاعات محسوب شود. در همین راستا به بررسی عوامل تأثیرگذار بر خرید اینترنتی افراد در حیطه‌های مختلف پرداخته می‌شود. در این بخش ابتدا مدل خطی تعمیم‌یافته از آن جهت که متغیر پاسخ مورد

مطالعه به صورت دوجمله‌ای و از خانواده نمایی است مطرح می‌شود. همچنین مدل رگرسیونی با پاسخ دوجمله‌ای برای داده‌های مورد مطالعه بیان می‌شود، سپس مدل رگرسیون دو سطحی با پاسخ دوجمله‌ای (برای مثال تعداد خرید لوازم سرگرمی) از آن جهت که داده‌های مورد نظر دارای ساختار سلسله مراتبی هستند مطرح می‌شود.

الف - مدل خطی تعمیم‌یافته: مدل خطی تعمیم‌یافته^۱ تعمیم رگرسیون خطی برای داده‌هایی که فرض توزیعی آن در خانواده نمایی قرار می‌گیرد، به کار می‌رود. مدل‌های خطی تعمیم‌یافته به وسیله‌ی سه مؤلفه‌ی زیر مشخص می‌شوند:

۱- مؤلفه‌ی تصادفی: مؤلفه تصادفی، توزیع احتمالی متغیر پاسخ را مشخص می‌کند. این مؤلفه شامل مشاهدات مستقل $y_1, \dots, y_n$ از متغیر تصادفی Y با توزیعی از یک خانواده نمایی^۲ است. هر یک از این مشاهدات، دارای تابع چگالی از خانواده زیر است:

$$f_Y(y|\theta) = b(y_i) a(\theta_i) \exp(Q(\theta_i) T(y_i))$$

که $Q(\theta_i)$ پارامتر طبیعی، $a(\theta_i)$ تابعی از پارامتر θ_i و $b(y_i)$ تابعی از y_i است.

به عنوان مثال، توزیع‌های برنولی، پواسون، چندجمله‌ای و نرمال که بسیار مورد استفاده قرار می‌گیرند، مثال‌هایی از اعضای این خانواده هستند. (Holland and Leinhardt, 1981: 373).

۲- مؤلفه‌ی سیستماتیک: مؤلفه‌ی سیستماتیک شامل یک بردار η است که ترکیب خطی از مجموعه‌ی متغیرهای پیشگو است، یعنی

$$\eta = X'\beta = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$$

که در آن $X = (X_1, \dots, X_p)$ یک برداری از p متغیرهای کمکی و β بردار پارامترهای مدل است. η نیز یک برآوردگر خطی است که پیشگوی خطی نامیده می‌شود.

۳- تابع ربط: ارتباط بین مؤلفه‌ی سیستماتیک ($\eta = X'\beta$) و مقادیر مورد انتظار مؤلفه‌های تصادفی را توصیف می‌کند. تابع ربط، تابعی است مشتق‌پذیر و یکنوا که میانگین پاسخ را به مؤلفه‌ی سیستماتیک پیوند می‌دهد.

فرض کنید $\mu_i = E(y_i)$ آنگاه رابطه $\eta_i = g(\mu_i)$ برقرار است که در آن g یک تابع مشتق‌پذیر یکنوا است. در نتیجه داریم:

^۱ Generalized Linear Model

^۲ Exponential Family

$$g(\mu_i) = X_i' \beta = \sum_{j=1}^p x_{ij} \beta_j. \quad i = 1, \dots, n$$

به عنوان مثال، اگر $g(\mu_i) = \mu_i$ آن‌گاه رابطه $\mu_i = \sum_{j=1}^p x_{ij} \beta_j$ برقرار است که به آن تابع ربط همانی می‌گویند. در حالت کلی اگر $g(\mu_i) = Q(\theta_i)$ ، آن‌گاه حالت زیر رخ می‌دهد:

$$Q(\theta_i) = \sum_{j=1}^p x_{ij} \beta_j$$

به این تابع ربط، تابع ربط کانونی می‌گویند.

ب - مدل رگرسیون با پاسخ دوجمله‌ای: مدل‌های رگرسیونی با پاسخ دوجمله‌ای، حالت خاص مدل‌های خطی تعمیم‌یافته هستند که شامل رگرسیون لوژیستیک، پروبیت، لگ-لگ و لگ-لگ مکمل است (Wacholder, 1986: 3). در مسائلی که متغیر پاسخ تعداد موفقیت‌ها در دنباله‌ای شامل m آزمایش مستقل دودویی با احتمال موفقیت π باشد رگرسیون با پاسخ دوجمله‌ای مورد استفاده قرار می‌گیرد. برای بیان ساختار مدل رگرسیون با پاسخ دوجمله‌ای لازم است تابع چگالی احتمال y_i در قالب خانواده نمایی به صورت زیر مطرح شود:

$$f_{Y_i}(y_i|\pi_i) = \binom{m}{y_i} \exp \left(y_i \log \left(\frac{\pi_i}{1 - \pi_i} \right) + m \log (1 - \pi_i) \right), \quad (1)$$

$$i = 1, \dots, n$$

$f_{Y_i}(y_i|\pi_i)$ در رابطه (۱) عضو خانواده‌ی نمایی طبیعی است، n تعداد مشاهدات مورد مطالعه و π_i احتمال موفقیت فرد i ام است. در این مدل تابع ربط، همان پارامتر طبیعی توزیع یعنی لوجیت π_i است که برای فرد i ام به صورت زیر تعریف می‌شود:

$$\text{logit}(\pi_i) = \log \left(\frac{\pi_i}{1 - \pi_i} \right) = \beta_0 + \sum_{k=1}^p x_{ik} \beta_k \quad (2)$$

یا به عبارتی

$$\pi_i = \frac{\exp \left(\beta_0 + \sum_{k=1}^p \beta_k X_{ik} \right)}{1 + \exp \left(\beta_0 + \sum_{k=1}^p \beta_k X_{ik} \right)} \quad (3)$$

در رابطه (۲) و (۳)، β_k ضریب رگرسیونی و p تعداد ضرایب رگرسیونی متغیرهای کمکی است و با در نظر گرفتن عرض از مبدأ در مجموع $p + 1$ ضریب رگرسیونی داریم. چون در این

مدل ضرایب به صورت غیر خطی با متغیر پاسخ در ارتباط هستند، بنابراین نمی‌توان برای برآورد آن از تخمین‌زننده‌های خطی مانند تخمین‌زننده کمترین توان‌های دوم مانده‌ها^۱ استفاده کرد. به همین دلیل از روش ماکسیمم درست‌نمایی^۲ برای برآورد پارامترهای این مدل استفاده می‌شود. مزیت اصلی مدل لجیت این است که مقدار برآوردشده برای احتمال وقوع متغیر پاسخ در این مدل الزاماً در بازه [۰, ۱] خواهد بود.

از آنجایی که داده‌های تحت بررسی دارای متغیر پاسخ Y (تعداد خرید در لوازم سرگرمی، تعداد خرید در لوازم ملزومات زندگی و یا تعداد خرید در خدمات آموزشی-فرهنگی) از توزیع دوجمله‌ای^۳ (برای اطلاعات بیشتر در رابطه با توزیع دوجمله‌ای به (Wackerly et al., 2014: 3) مراجعه شود) با m آزمایش مستقل برنولی^۴ (دودویی) با احتمال موفقیت π است و متغیرهای کمکی شهری یا روستایی بودن (Living) (دارای دو رده شهری بودن (Urban) و روستایی بودن (Rural) است که روستایی بودن به عنوان مبنا در نظر گرفته شده است)، جنسیت (Gender) (دارای دو رده مرد (Male) و زن (Female) است که زن به عنوان مبنا در نظر گرفته شده است)، سن (Age)، وضع سواد و مدرک تحصیلی (Education) (دارای ۳ رده دیپلم (Diploma)، زیر دیپلم (Under_Diploma) و تحصیلات دانشگاهی (Bachelor_Higher) است که دیپلم به عنوان مبنا در نظر گرفته می‌شود) و وضع فعالیت (Job) (دارای ۶ رده شاغل (Employed)، بیکار (Unemployed)، محصل (Student)، دارای درآمد بدون کار (Pensionary)، خانه‌دار (Housewife) و سایر (Inactive) است که سایر به عنوان مبنا در نظر گرفته می‌شود)، با این تفاسیر با در نظر گرفتن عرض از مبدأ ۱۱ ضریب رگرسیونی داریم که باید برآورد شود. مدل خطی تعمیم‌یافته یا مدل رگرسیون با پاسخ دوجمله‌ای برای داده‌های مورد مطالعه برای فرد i ام به صورت زیر تعریف می‌شود:

$$\text{logit}(\pi_i) = \beta_0 + \beta_1 \text{Living}_{i1} + \beta_2 \text{Gender}_{i1} + \beta_3 \text{Age}_i + \beta_4 \text{Education}_{i1} + \beta_5 \text{Education}_{i2} + \beta_6 \text{Job}_{i1} + \beta_7 \text{Job}_{i2} + \beta_8 \text{Job}_{i3} + \beta_9 \text{Job}_{i4} + \beta_{10} \text{Job}_{i5} \quad (4)$$

در رابطه (۴)، Living_{i1} بیانگر روستایی بودن و Gender_{i1} بیانگر زن بودن و Age_i بیانگر سن فرد i ام است. Education_{i1} بیانگر دارا بودن مدرک زیر دیپلم و Education_{i2} بیانگر دارا

¹ Least Residual Squares

² Maximum Likelihood Method

³ Binomial Distribution

⁴ Bernoulli

بودن مدرک تحصیلی تحصیلات دانشگاهی، فرد λ_m است. Job_{i1} بیانگر شاغل بودن، Job_{i2} بیانگر بیکار بودن، Job_{i3} بیانگر محصل بودن، Job_{i4} بیانگر دارای درآمد بدون کار بودن، Job_{i5} بیانگر خانه‌دار بودن فرد λ_m است.

حال برای برآورد ضرایب در مدل رگرسیونی زمانی که پاسخ به صورت دوجمله‌ای با تابع ربط لوجیت باشد از طریق بیشینه‌سازی لگاریتم درستنمایی یا کمینه‌سازی منفی لگاریتم درستنمایی حاصل می‌شود. تابع درستنمایی در این حالت به صورت زیر است:

$$L(\beta, Y) = \prod_{i=1}^n f(y_i) = \prod_{i=1}^n \binom{m}{y_i} \pi_i^{y_i} (1 - \pi_i)^{m-y_i}$$

حاصل لگاریتم درستنمایی به صورت زیر است:

$$l(\beta, Y) = \sum_{i=1}^n \left(y_i \log \pi_i + (m - y_i) \log (1 - \pi_i) + \log \binom{m}{y_i} \right) \quad (5)$$

حال با جایگذاری رابطه (۳) در رابطه (۵) بردار β به صورت زیر برآورد می‌شود:

$$\hat{\beta} = \underset{\beta \in R^{p+1}}{\text{Arg min}} \left\{ -\sum_{i=1}^n \left(y_i \log \pi_i + (m - y_i) \log (1 - \pi_i) + \log \binom{m}{y_i} \right) \right\} \quad (6)$$

ج- مدل رگرسیون دو سطحی با پاسخ دوجمله‌ای: در یک مدل رگرسیون دو سطحی^۱ فرض می‌شود سطح یک آن شامل افراد و سطح دو آن شامل z گروه است. این مدل دوسطحی با متغیر پاسخ Y (تعداد خرید اینترنتی در حیطه‌های مختلف سرگرمی، ملزومات زندگی و یا خدمات آموزشی فرهنگی) که دارای توزیع دوجمله‌ای با m آزمایش مستقل دودویی با احتمال موفقیت π است و متغیرهای کمکی همانطور که در بخش مدل رگرسیون با پاسخ دوجمله‌ای به طور کامل مطرح شدند؛ شهری یا روستایی بودن (Living)، جنسیت (Gender)، سن (Age)، وضع سواد و مدرک تحصیلی (Education) و وضع فعالیت (Job). مدل لوجیت برای فرد λ_m درون گروه λ_m به صورت زیر تعریف می‌شود:

$$\begin{aligned} \text{logit}(\pi_{ij}) = & \beta_{.j} + \beta_{1j} \text{Living}_{ij1} + \beta_{2j} \text{Gender}_{ij1} + \beta_{3j} \text{Age}_{ij} + \beta_{4j} \text{Education}_{ij1} \\ & + \beta_{5j} \text{Education}_{ij2} + \beta_{6j} \text{Job}_{ij1} + \beta_{7j} \text{Job}_{ij2} + \beta_{8j} \text{Job}_{ij3} + \beta_{9j} \text{Job}_{ij4} \\ & + \beta_{10j} \text{Job}_{ij5} \end{aligned}$$

¹ Two-Level Regression

$$i = 1, 2, \dots, I; j = 1, 2, \dots, J \quad (7)$$

$\beta_{.j}$ عرض از مبدأ $\beta_{1j}, \beta_{2j}, \dots, \beta_{10j}$ شیب‌های خط رگرسیونی است که به کمک نمونه مورد نظر برآورد می‌شوند. پارامترهای $\beta_{.j}, \beta_{1j}, \dots, \beta_{10j}$ برای هر گروه (سطح) به صورت زیر تعریف می‌شوند:

$$\begin{aligned} \beta_{.j} &= \gamma_{..} + u_{.j} \\ \beta_{1j} &= \gamma_{1.} + u_{1j} \\ \beta_{2j} &= \gamma_{2.} + u_{2j} \\ &\vdots \\ \beta_{10j} &= \gamma_{10.} + u_{10j} \end{aligned} \quad (8)$$

در عبارت‌های تصادفی $\beta_{.j}, \beta_{1j}, \dots, \beta_{10j}$ عبارت‌های $u_{.j}, u_{1j}, \dots, u_{10j}$ خطاهای تصادفی حاصل از همبستگی بین افراد هر گروه (استان) است. با جایگذاری رابطه‌ی (۸) در (۷)، معادله‌ی کلی برای مدل دو سطحی با پاسخ دوجمله‌ای حاصل می‌شود. در این بخش مدل‌های رگرسیون با پاسخ دوجمله‌ای و رگرسیون دو سطحی با پاسخ دوجمله‌ای مطرح شد. برای بررسی مناسب بودن مدل پیشنهادی برای داده‌های سلسله‌مراتبی در بخش بعدی شبیه‌سازی مطرح می‌شود.

• شبیه‌سازی

برای بررسی مناسبیت مدل رگرسیون دوسطحی با پاسخ دوجمله‌ای یک مجموعه داده با ساختار سلسله‌مراتبی با چهار متغیر کمکی که هر چهار متغیر کمکی به صورت رده‌بندی شده (طبقه‌بندی شده) هستند. متغیر کمکی X_1 دارای دو رده (طبقه)، X_2 سه رده، X_3 چهار رده و X_4 دو رده است. زمانی که متغیرهای کمکی به صورت اسمی یا رده‌بندی باشند، یک رده به عنوان مبنا در نظر گرفته می‌شود و ضرایب رگرسیونی برای سایر رده‌ها بر اساس مبنا حاصل می‌شود. برای مثال X_1 دارای دو رده است که یک رده آن به عنوان مبنا در نظر گرفته می‌شود و تنها ضریب رگرسیونی مربوط به یک رده آن محاسبه می‌شود. با این تفاسیر با در نظر گرفتن مبنا برای هر متغیر، در کل با در نظر گرفتن عرض از مبدأ ۸ ضریب رگرسیونی اولیه به اندازه نمونه $n = 30 \times 100$ از طریق نمونه‌گیری تصادفی با احتمال‌های مختلف تولید می‌شود و برای ایجاد ساختار سلسله‌مراتبی داده‌ها به ۳۰ گروه با تعداد نمونه ۱۰۰ در هر گروه

تقسیم‌بندی شده است. ضرایب رگرسیونی متغیرهای فوق با تعیین مبنا از رده‌های متغیرهای کمکی با در نظر گرفتن عرض از مبدأ به صورت

$$\beta = \{-0.15, -0.3, 0.5, -0.15, 0.5, -0.15, 0.8, 0.8\}$$

که عرض از مبدأ -0.15 در نظر گرفته شده است.

با در دست داشتن متغیرهای کمکی و ضرایب رگرسیونی تعیین‌شده تابع لجیت به صورت زیر حاصل می‌شود:

$$\pi_{ij} = \frac{\exp\left(\beta_{.j} + \sum_{k=1}^p \beta_{kj} X_{ijk}\right)}{1 + \exp\left(\beta_{.j} + \sum_{k=1}^p \beta_{kj} X_{ijk}\right)} \quad (9)$$

$$j = 1, \dots, 30; \quad i = 1, \dots, 30 \times 100; \quad k = 1, \dots, p$$

که β_{kj} به صورت زیر تولید می‌شود:

$$\beta_{kj} = \gamma_{k.} + u_{kj}$$

X_{ijk} متغیر کمکی k ام برای فرد i ام درون گروه j ام است. u_{ij} مؤلفه خطای تصادفی است که به میزان انحراف ضریب رگرسیونی یک گروه با ضریب رگرسیون کلی اشاره دارد. $\gamma_{k.}$ ضریب رگرسیونی کلی بین متغیر پاسخ و متغیر کمکی سطح یک است. u_j از توزیع نرمال چندمتغیره با میانگین μ و ماتریس همبستگی Σ تولید می‌شود. در اینجا از پکیج mvtnorm و lme4 وجود در نرم‌افزار R استفاده شده است.

که μ و Σ به صورت زیر تعریف می‌شوند:

$$\Sigma = \begin{pmatrix} \sigma_{.1}^2 & \sigma_{.1} & \dots & \sigma_{.p} \\ \sigma_{.1} & \sigma_{.1}^2 & \dots & \sigma_{.1p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{.p} & \sigma_{.1p} & \dots & \sigma_{.p}^2 \end{pmatrix}, \quad \mu = \begin{pmatrix} . \\ . \\ \vdots \\ . \end{pmatrix}, \quad u_j = \begin{pmatrix} u_{.j} \\ u_{1j} \\ \vdots \\ u_{pj} \end{pmatrix}$$

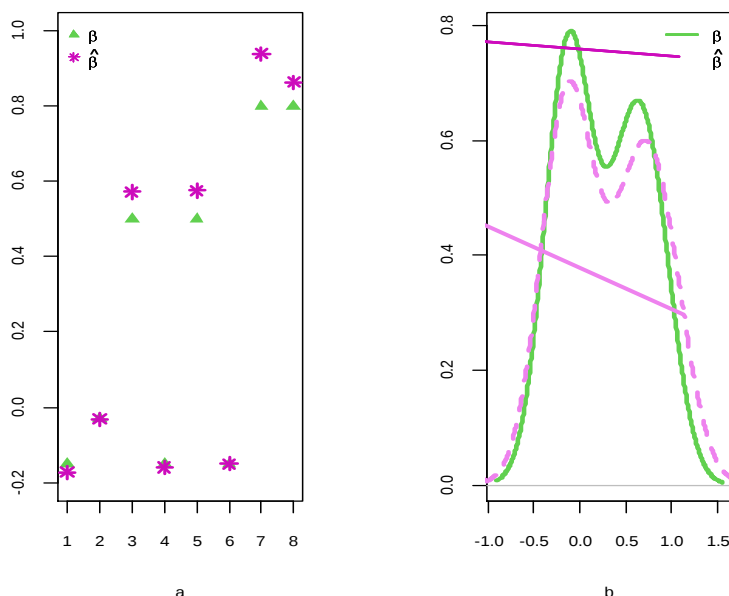
متغیر پاسخ مورد نظر به صورت تصادفی از توزیع دوجمله‌ای با اندازه تکرار آزمایش $m = 4$ و احتمال موفقیت π_{ij} برای $j = 1, \dots, 30$ شبیه‌سازی می‌شود. تابع درست‌نمایی را می‌توان به صورت زیر نوشت:

$$\begin{aligned}
L(\boldsymbol{\beta}, \boldsymbol{\Sigma} | \mathbf{y}, \mathbf{x}) &= \prod_{j=1}^{r_0} \left\{ \Pr(Y_{1j} = y_{1j}, \dots, Y_{1..j} = y_{1..j} | x, \boldsymbol{\beta}) \right\} \\
&= \\
&\prod_{j=1}^{r_0} \left\{ \int \dots \int \prod_{i=1}^{100} \left\{ \Pr(Y_{ij} = y_{ij} | u_{.j}, \dots, u_{pj}) f_{u_j}(u_{.j}, \dots, u_{pj}) du_{.j}, \dots, du_{pj} \right\} \right\} \\
&= \\
&\prod_{j=1}^{r_0} \left\{ \int \dots \int \prod_{i=1}^{100} \left\{ \binom{m}{y_{ij}} \pi_{ij}^{y_{ij}} (1 - \right. \right. \\
&\left. \left. \pi_{ij} \right)^{m-y_{ij}} \frac{\exp\left(-\frac{1}{2} \mathbf{u}_j^T \boldsymbol{\Sigma}^{-1} \mathbf{u}_j\right)}{\sqrt{2 \times \pi}^{p+1}} du_{.j}, \dots, du_{pj} \right\} \right\}
\end{aligned}$$

ضرایب رگرسیونی را با بیشینه‌سازی لگاریتم درستنمایی یا کمینه‌سازی منفی لگاریتم درستنمایی و با جایگذاری π_{ij} حاصل از رابطه (۱۰) با ۱۰۰۰ بار تکرار به روش مونت کارلو صورت زیر حاصل می‌شود:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\theta} \in \mathbb{R}^{p+1}}{\text{Arg min}} \left\{ -\log(L(\boldsymbol{\beta}, \boldsymbol{\Sigma} | \mathbf{y}, \mathbf{x})) \right\} \quad (10)$$

در رابطه (۱۰)، p تعداد ضرایب رگرسیونی متغیرهای کمکی است و با در نظر گرفتن عرض از مبدأ در مجموع $p + 1$ ضریب رگرسیونی داریم. به علت پیچیده بودن حل انتگرال مسئله به صورت صریح روش‌های جایگزین مانند روش تکراری عددی حل می‌شود.



نمودار ۱- نمودارهای ضرایب واقعی و ضرایب برآوردشده متغیرهای کمکی

نمودار ۱ برای بررسی کارکرد مدل پیشنهادی مفید است. منحنی (شکل ۳-۱(b)) و مثلث (شکل ۳-۱(a)) مربوط به ضرایب رگرسیونی واقعی هستند. خط تیره (شکل ۳-۱(b)) و ستاره (شکل ۳-۱(a)) مربوط به ضرایب رگرسیونی برآوردشده از روش رگرسیون دوسطحی با پاسخ دوجمله‌ای است. با توجه به شکل‌های فوق ضرایب رگرسیونی برآوردشده با استفاده از روش رگرسیون دوسطحی با پاسخ دوجمله‌ای چندان تفاوتی با ضرایب رگرسیونی واقعی ندارد. در نتیجه می‌توان گفت که روش رگرسیون دوسطحی با پاسخ دوجمله‌ای برای داده‌های سلسله مراتبی یا دوسطحی مناسب است.

یافته‌های تحقیق

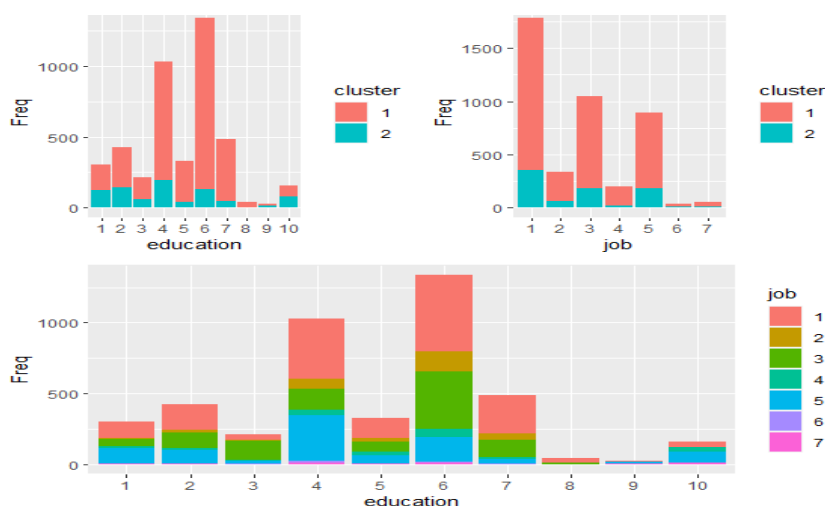
فناوری اطلاعات و ارتباطات تحولات گسترده‌ای را در زمینه‌های مختلف اجتماعی، فرهنگی، اقتصادی و پزشکی و غیره به همراه داشته است و به عنوان یکی از شاخص‌های مهم توسعه‌یافتگی کشورها به شمار می‌رود. در همین راستا به تحلیل داده‌های اخذ شده از اطلاعات «آمارگیری استفاده‌ی خانوارها و افراد از فناوری اطلاعات و ارتباطات مرکز آمار ایران در سال ۱۳۹۶ برای ۳۱ استان کشور» که دارای ساختار دوسطحی یا سلسله‌مراتبی هستند، پرداخته

می‌شود. چنان‌که افراد در سطح یک و استان‌ها در سطح دو قرار دارند. در واقع در این مطالعه افراد به صورت تصادفی از استان‌ها انتخاب شده‌اند. برای چنین داده‌هایی به کار بردن روش‌های سنتی به دلیل نبود فرض استقلال مشاهدات در استان یکسان مناسب نیست و جهت دستیابی به نتایج دقیق‌تر بهتر است رگرسیون دو سطحی مورد استفاده قرار گیرد و با توجه به این‌که متغیر Y تعداد کالاهایی است که یک فرد در حیطه‌های مختلف (سرگرمی، ملزومات زندگی و خدمات آموزشی فرهنگی) موفق به خرید آن‌ها به صورت اینترنتی می‌شود که با در نظر گرفتن فرض یکسان بودن آزمایش‌ها و یکسان بودن احتمال موفقیت هر آزمایش به صورت دوجمله‌ای است، از مدل رگرسیون دوسطحی با پاسخ دوجمله‌ای استفاده می‌شود. در این مطالعه تأثیرگذاری عوامل مختلف مانند شهری یا روستایی بودن (Living)، جنسیت (Gender)، سن (Age)، وضع سواد و مدرک تحصیلی (Education) و وضع فعالیت (Job) بر متغیر پاسخ مورد بررسی قرار می‌گیرد.

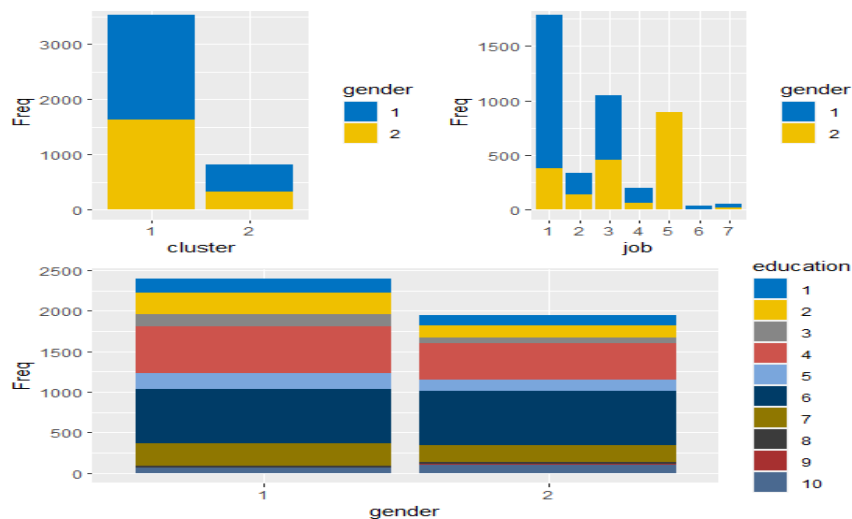
نمودارهای ۲، ۳ و ۴ برای توصیف کامل داده‌ها شامل فراوانی و روابط بین متغیرهای مختلف رسم شده‌اند. داده‌ها به صورت کلی شامل شهری یا روستایی بودن (Cluster) (شهری بودن (کد ۱) و روستایی بودن (کد ۲))، جنسیت (Gender) (مرد (کد ۱) و زن (کد ۲))، سن (Age)، وضع سواد و مدرک تحصیلی (Education) (ابتدایی (کد ۱)، راهنمایی یا مقطع اول متوسطه (کد ۲)، مقطع دوم متوسطه (کد ۳)، دیپلم و پیش‌دانشگاهی (کد ۴)، کاردانی (کد ۵)، کارشناسی (کد ۶)، کارشناسی ارشد و دکترای حرفه‌ای (کد ۷)، دکترای تخصصی (کد ۸)، سوادآموزی و غیررسمی (کد ۹)، بی‌سواد (کد ۱۰)) و وضع فعالیت (Job) (شاغل (کد ۱)، بیکار (کد ۲)، محصل (کد ۳)، دارای درآمد بدون کار (کد ۴)، خانه‌دار (کد ۵)، سرباز وظیفه (کد ۶)، سایر (کد ۷)) هستند. با توجه به نمودار ۲ جمعیت بسیاری از افرادی که در شهر زندگی می‌کنند دارای تحصیلات بالاتری نسبت به افرادی که در روستا زندگی می‌کنند، هستند، همچنین اکثر افرادی که در شهر زندگی می‌کنند دارای تحصیلات کارشناسی و دیپلم و پیش‌دانشگاهی هستند. از این رو می‌توان گفت شهری یا روستایی بودن روی میزان تحصیلات تأثیرگذار است. اکثر افرادی که دارای مدرک کارشناسی ارشد و دکترای حرفه‌ای هستند، دارای شغل هستند. همچنین با افزایش میزان تحصیلات اشتغال به کار نیز افزایش می‌یابد. نسبت افراد شاغل و محصل که در شهر زندگی می‌کنند از افرادی که در روستا زندگی می‌کنند بیش‌تر است. با توجه به نمودار ۳ شهری یا روستایی بودن روی جنسیت تأثیرگذار نیست. تعداد مردان و زنان در شهر و روستا تقریباً یکسان

است. نسبت اشتغال به کار در مردان از زنان بیش‌تر است. فراوانی محصلان زن و مرد تقریباً یکسان است و نسبت افراد دارای درآمد بدون کار در مردان بیش‌تر است. فراوانی مردان و زنان با تحصیلات کارشناسی تقریباً یکسان است. به صورت کلی می‌توان گفت جنسیت بر شغل تأثیرگذار است و تحصیلات نیز چندان تأثیری روی جنسیت ندارد.

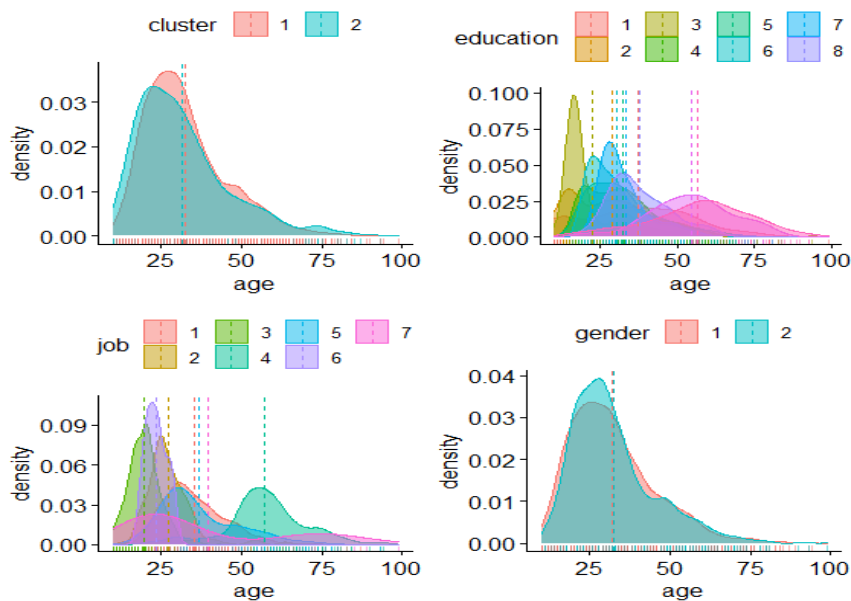
با توجه به نمودار ۴ محدوده سنی افراد شهری یا روستایی چندان تفاوتی ندارد تنها محدوده سنی افراد شهری اندکی جوان‌تر است. معمولاً افرادی که سن بالاتری دارند تحصیلات پایین‌تری دارند. و افرادی که در سنین ۲۵-۴۰ هستند دارای تحصیلات بالاتری هستند که این بستگی به پیشرفت دانش دارد. افرادی که در سنین بالاتری هستند در دوران تحصیلات خود از امکانات کافی برخوردار نبوده‌اند. با پیشرفت دانش و جامعه سطح تحصیلات افراد نیز افزایش یافته است. افرادی که در سنین جوان‌تری هستند اکثراً، محصل هستند. افراد در سنین تقریباً ۴۵-۸۰ دارای درآمد بدون کار هستند و در سن ۶۰ به بیش‌ترین مقدار خود می‌رسد. و این می‌تواند به دلیل بازنشستگی یا پس‌انداز افراد باشد. جنسیت نیز بستگی به سن ندارد یعنی زن و مرد در همه محدوده‌های سنی تقریباً به میزان یکسانی در این مطالعه حضور دارند، تنها محدوده سنی زنان اندکی جوان‌تر است.



نمودار ۲- ارتباط متغیرهای تحصیلات، شغل و شهری یا روستایی بودن و فراوانی آن‌ها



نمودار ۳- ارتباط بین متغیرهای تحصیلات، شغل، جنسیت و شهری یا روستایی بودن و فراوانی آن‌ها



نمودار ۴- ارتباط بین متغیر سن با شهری یا روستایی بودن، تحصیلات، شغل، جنسیت و فراوانی آن‌ها

تحلیل داده‌ها

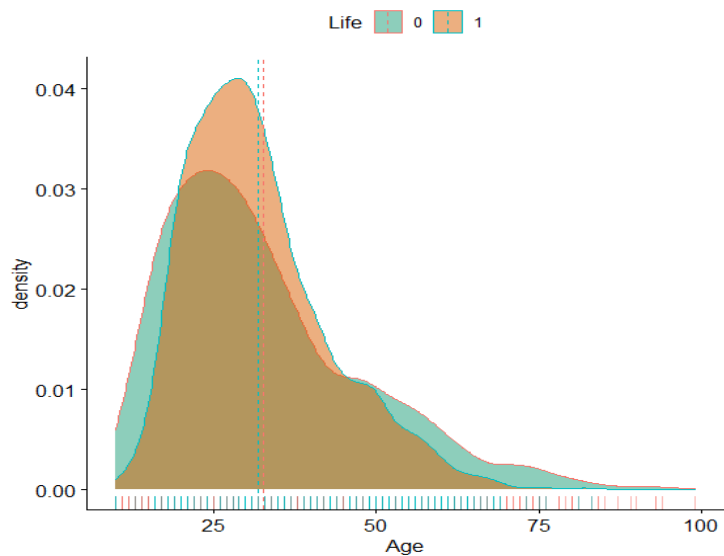
• بررسی عوامل مؤثر بر تعداد خرید اینترنتی در حیطه ملزومات زندگی

در این مطالعه متغیر پاسخ تعداد خریده‌ها یا کالاهایی است که یک فرد به صورت اینترنتی در حیطه ملزومات زندگی انجام می‌دهد در واقع خرید Y کالا از 6 کالا است که Y می‌تواند مقادیر $0, 1, \dots, 6$ را اختیار کند، این 6 کالا که در حیطه ملزومات زندگی قرار می‌گیرد عبارت‌اند از:

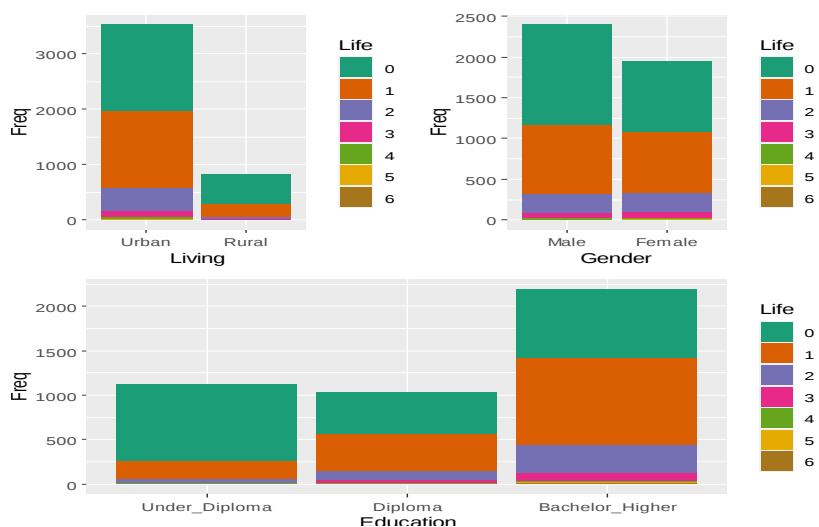
۱. پوشاک و کفش
۲. لوازم خانگی
۳. لوازم آرایشی، بهداشتی، عطر و زیور آلات
۴. لوازم ورزشی و سرگرمی
۵. غذا، خواروبار و نوشیدنی
۶. دارو و لوازم پزشکی

در این بخش بررسی می‌شود که چه متغیرهایی بر تعداد خرید اینترنتی در حیطه ملزومات زندگی تأثیرگذار هستند. متغیرهای کمکی مورد نظر در این مطالعه شهری یا روستایی بودن، جنسیت، سن، وضع سواد و مدرک تحصیلی و وضع فعالیت هستند. نمودار ۵ برای بررسی ارتباط بین متغیر کمکی سن با خرید مردم در حیطه ملزومات زندگی رسم شده است که در این نمودار شکل زرد رنگ بیانگر انجام حداقل یک خرید به صورت اینترنتی در حیطه ملزومات زندگی است، شکل آبی رنگ بیانگر این است که هیچ خرید اینترنتی صورت نگرفته است. و از آنجایی که دو منحنی روی هم منطبق نیستند می‌توان گفت سن بر تعداد خرید افراد به صورت اینترنتی در حیطه ملزومات زندگی تأثیرگذار است و معمولاً افرادی که در این حیطه خرید خود را انجام می‌دهند جوان‌تر از افرادی هستند که هیچ خرید اینترنتی انجام نداده‌اند. که با آزمون به صورت دقیق‌تر مورد بررسی قرار می‌گیرد. نمودار ۶ و ۷ برای بررسی ارتباط بین متغیرهای کمکی مختلف با خرید مردم در حیطه ملزومات زندگی رسم شده است که با توجه به آن وضع فعالیت بر خرید به صورت اینترنتی تأثیرگذار است. افراد شاغل و محصل و خانه‌دار و بیکار اکثراً خرید اینترنتی داشته‌اند. خرید اینترنتی در زنان بیش‌تر انجام شده است. پس می‌توان گفت جنسیت نیز روی تعداد خرید اینترنتی افراد در زمینه ملزومات زندگی تأثیرگذار است. افرادی که در شهر زندگی می‌کنند نسبت به افرادی که در روستا زندگی می‌کنند بیش‌تر،

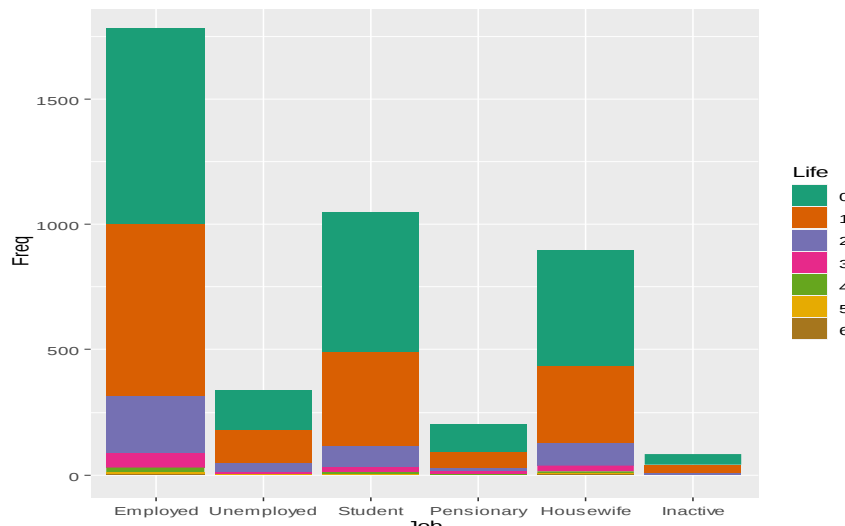
خرید خود را به صورت اینترنتی انجام می‌دهند. هر چه تحصیلات افراد بالاتر باشد، خرید افراد به صورت اینترنتی در این حیطه افزایش می‌یابد.



نمودار ۵- ارتباط متغیر پاسخ (تعداد خرید در حیطه ملزومات زندگی) و متغیر کمکی سن



نمودار ۶- ارتباط متغیر پاسخ (تعداد خرید در حیطه ملزومات زندگی) و متغیرهای کمکی جنسیت، شهری یا روستایی بودن و تحصیلات



نمودار ۷- ارتباط متغیر پاسخ (تعداد خرید در حیطه ملزومات زندگی) و متغیرهای کمکی شغل

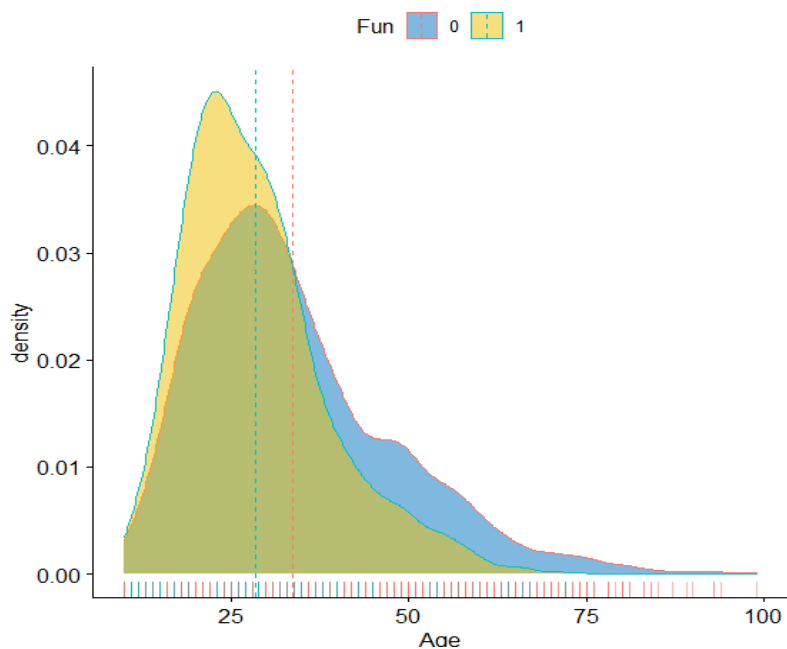
• بررسی عوامل مؤثر بر تعداد خرید اینترنتی در حیطه سرگرمی

در این مطالعه متغیر پاسخ تعداد خریدهای اینترنتی است که یک فرد انجام می‌دهد. در واقع خرید y کالا از 4 کالا است که y می‌تواند مقادیر $0, 1, \dots, 4$ را اختیار کند، این 4 کالا که در حیطه سرگرمی قرار می‌گیرد عبارت‌اند از:

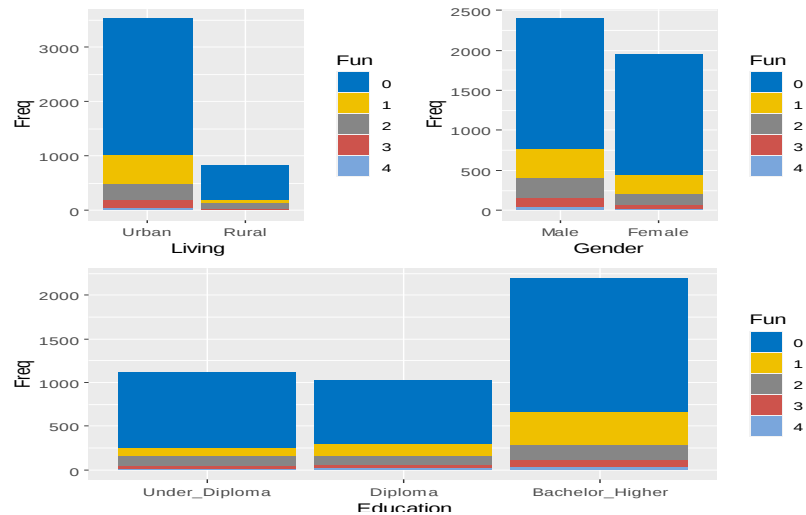
۱. موسیقی
۲. فیلم
۳. بازی‌های رایانه‌ای یا ویدئویی
۴. نرم‌افزار رایانه‌ای

در این بخش بررسی می‌شود که چه متغیرهایی بر تعداد خرید اینترنتی در حیطه سرگرمی تأثیرگذار هستند. متغیرهای کمکی مورد نظر در این مطالعه شهری یا روستایی بودن، جنسیت، سن، وضع سواد و مدرک تحصیلی و وضع فعالیت هستند. نمودار ۸ برای بررسی ارتباط بین متغیر کمکی سن با تعداد خرید مردم در حیطه سرگرمی رسم شده است که در این نمودار شکل آبی رنگ بیانگر حداقل یک خرید به صورت اینترنتی در حیطه سرگرمی است، شکل صورتی رنگ بیانگر این است که هیچ خرید اینترنتی صورت نگرفته

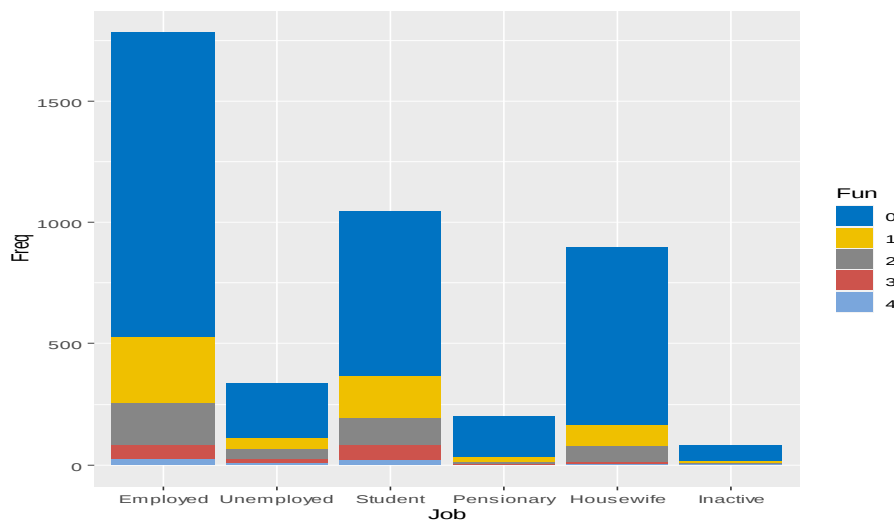
است. و از آنجایی که دو منحنی روی هم منطبق نیستند، می‌توان گفت سن بر تعداد خرید افراد به صورت اینترنتی در حیطه سرگرمی تأثیرگذار است که با آزمون به صورت دقیق‌تر مورد بررسی قرار می‌گیرد. نمودار ۹ و ۱۰ برای بررسی ارتباط بین متغیرهای کمکی مختلف با تعداد خرید مردم در حیطه سرگرمی رسم شده است که با توجه به این نمودار جنسیت بر تعداد خرید افراد در حیطه سرگرمی به صورت اینترنتی تأثیرگذار است. نسبت خرید مردان به زنان بیش‌تر است. اکثراً افراد شاغل و محصل خرید خود را در این حیطه انجام می‌دهند. افراد شهری نیز نسبت به روستاییان خرید بیش‌تری داشته‌اند. هر چه تحصیلات نیز بیش‌تر شود خرید اینترنتی افزایش می‌یابد.



نمودار ۸- ارتباط متغیر پاسخ (تعداد خرید در حیطه سرگرمی) و متغیر کمکی سن



نمودار ۹- ارتباط متغیر پاسخ (تعداد خرید در حیطه سرگرمی) و متغیرهای کمکی جنسیت، شهری یا روستایی بودن و تحصیلات



نمودار ۱۰- ارتباط متغیر پاسخ (تعداد خرید در حیطه سرگرمی) و متغیرهای کمکی شغل

• بررسی عوامل مؤثر بر تعداد خرید اینترنتی در خدمات آموزشی فرهنگی

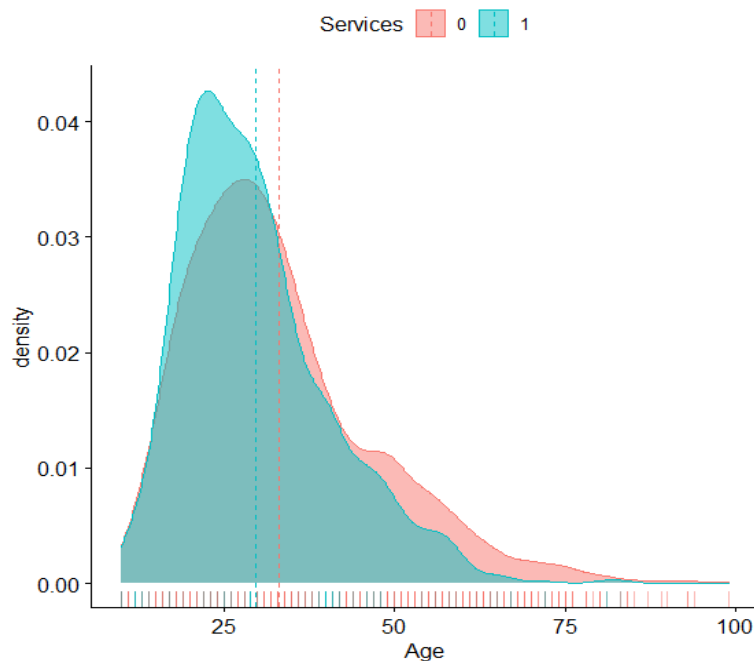
در این مطالعه متغیر پاسخ تعداد خریده‌ها یا کالاهایی است که یک فرد در حیطه خدمات آموزشی فرهنگی موفق به خرید آن‌ها می‌شود در واقع خرید y کالا از ۴ کالا است که y

می‌تواند مقادیر 0,1,...,4 را اختیار کند، این 4 کالا که در حیطه خدمات آموزشی فرهنگی قرار می‌گیرد عبارت‌اند از:

۱. خدمات آموزشی
۲. خدمات فرهنگی
۳. خدمات چاپ و ترجمه
۴. کتاب، مجله، روزنامه، مقاله

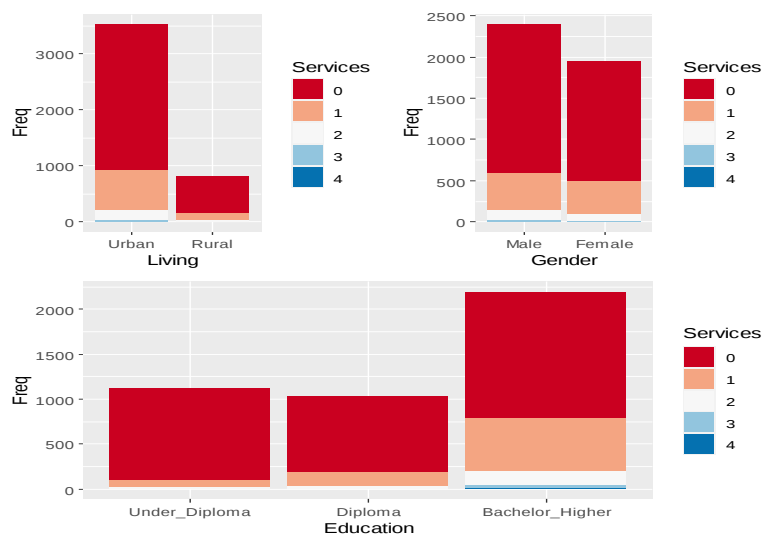
در این بخش بررسی می‌شود که چه متغیرهایی بر تعداد خرید اینترنتی در حیطه خدمات آموزشی فرهنگی تأثیرگذار هستند. متغیرهای کمکی مورد نظر در این مطالعه شهری یا روستایی بودن، جنسیت، سن، وضع سواد و مدرک تحصیلی و وضع فعالیت هستند. نمودار ۱۱ برای بررسی ارتباط بین تعداد خرید اینترنتی در حیطه خدمات آموزشی فرهنگی و سن افراد رسم شده است که در این نمودار به ترتیب رنگ‌هایی که با شماره مشخص شده‌اند بیانگر تعداد خرید افراد در حیطه خدمات آموزشی و فرهنگی است. عدد صفر بیانگر این است که هیچ خریدی در این حیطه صورت نگرفته است. و از آنجایی که منحنی‌ها بسیار نزدیک به هم هستند و میانگین‌های آن‌ها نزدیک به هم است، می‌توان گفت سن بر تعداد خرید افراد به صورت اینترنتی در حیطه خدمات آموزشی فرهنگی تأثیرگذار نیست که با آزمون به صورت دقیق‌تر مورد بررسی قرار می‌گیرد.

نمودار ۱۲ و ۱۳ نیز برای بررسی روابط بین تعداد خرید در حیطه خدمات آموزشی فرهنگی و سایر متغیرهای کمکی رسم شده است که با توجه به نمودار ۱۲ جنسیت بر تعداد خرید افراد در حیطه خدمات آموزشی فرهنگی به صورت اینترنتی تأثیرگذار است. اکثراً افراد شاغل، محصل و خانه‌دار خرید خود را در این حیطه انجام می‌دهند. افراد شهری نیز نسبت به روستاییان خرید بیش‌تری داشته‌اند. هر چه تحصیلات نیز بیش‌تر شود خرید اینترنتی افزایش می‌یابد.



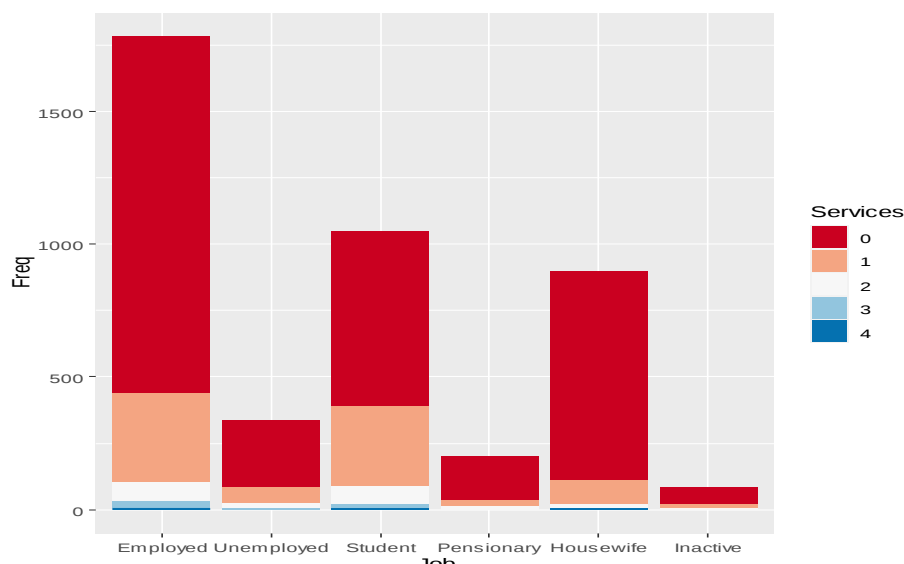
نمودار ۱۱- ارتباط متغیر پاسخ (تعداد خرید در حیطه خدمات آموزشی فرهنگی) و متغیر

کمکی سن



نمودار ۱۲- ارتباط متغیر پاسخ (تعداد خرید در حیطه خدمات آموزشی فرهنگی) و

متغیرهای کمکی جنسیت، شهری یا روستایی بودن و تحصیلات



نمودار ۱۳- ارتباط متغیر پاسخ (تعداد خرید در حیطه خدمات آموزشی فرهنگی) و متغیر کمکی شغل

• نتایج حاصل از برازش مدل تعداد خرید لوازم در حیطه‌های مختلف سرگرمی، ملزومات زندگی و خدمات آموزشی فرهنگی

نتایج حاصل از اجرای مدل رگرسیون دوسطحی با پاسخ دوجمله‌ای (رابطه ۷) در حیطه سرگرمی، ملزومات زندگی و خدمات آموزشی و فرهنگی به ترتیب در جدول‌های ۱، ۲ و ۳ آمده است. در این جداول مقدار p-value در سطح ۰۰۰۵ برای هریک از متغیرهای کمکی محاسبه شده است. همانطور که نمودار ۷ گویای این بود که سن بر تعداد خرید اینترنتی در حیطه سرگرمی تأثیرگذار است نتایج به دست آمده در جدول ۱ نیز این موضوع را تأیید می‌کند به این دلیل که مقدار p-value برای سن در حیطه سرگرمی کمتر از ۰۰۰۵ شده است، بنابراین سن بر تعداد خرید اینترنتی در حیطه سرگرمی تأثیرگذار است به عبارتی معنی‌دار است، همچنین با توجه به نمودارهای ۵ و ۹ که به آن اشاره شد، سن بر تعداد خرید در حیطه ملزومات زندگی تأثیرگذار و بر تعداد خرید در حیطه خدمات آموزشی فرهنگی تأثیرگذار نیست و نتایج جدول ۲ و جدول ۳ این را تأیید می‌کنند زیرا مقدار p-value برای سن در حیطه ملزومات زندگی کمتر از ۰/۰۵ و در حیطه خدمات آموزشی فرهنگی بزرگتر از ۰۰۰۵ است. با توجه به مقدار p-value کمتر از ۰۰۰۵ برای جنسیت و شهری یا روستایی

بودن (نوع اقامت) در حیطه‌های سرگرمی و ملزومات زندگی و خدمات آموزشی و فرهنگی جنسیت و نوع اقامت بر تعداد خرید اینترنتی افراد در حیطه سرگرمی (جدول ۱) و ملزومات زندگی (جدول ۲) و خدمات آموزشی و فرهنگی (جدول ۳) تأثیرگذار است. این موضوع را می‌توان به صورت نموداری در نمودار ۱۰ ملاحظه نمود. به همین ترتیب با توجه به نتایج هر سه جدول ۱، ۲، ۳ می‌توان گفت سطح تحصيلات بر تعداد خرید اینترنتی افراد در همه حیطه‌های سرگرمی، ملزومات زندگی و خدمات آموزشی و فرهنگی تأثیرگذار است به دلیل اینکه مقدار p-value برای اکثر متغیرها معنی دار است و هرچه سطح تحصيلات افراد بالاتر باشد خرید اینترنتی بیش‌تری در این حیطه‌ها صورت می‌گیرد.

برای بررسی تأثیرگذاری وضع فعالیت بر تعداد خرید اینترنتی در حیطه سرگرمی با توجه به نتایج جدول ۱ به دلیل اینکه p-value‌های مربوط به متغیرها کمتر از ۰۰۰۵ هستند، وضع فعالیت بر تعداد خرید اینترنتی در حیطه سرگرمی تأثیرگذار هستند. همچنین با بررسی تأثیرگذاری وضع فعالیت بر تعداد خرید اینترنتی در حیطه ملزومات زندگی نتایج جدول ۲ به دلیل اینکه p-value‌های مربوط به متغیرها به غیر از شاغل کوچکتر از ۰۰۰۵ هستند، بیانگر این است که وضع فعالیت تقریباً بر تعداد خرید اینترنتی در حیطه ملزومات زندگی تأثیرگذار هستند و این از روی نمودار ۶ قابل بیان است. برای بررسی تأثیرگذاری وضع فعالیت بر تعداد خرید اینترنتی در حیطه خدمات آموزشی و فرهنگی با مراجعه به نتایج جدول ۳ می‌توان این نتیجه را گرفت که به دلیل اینکه p-value‌های مربوط به محصل و خانه‌دار کوچکتر از ۰۰۰۵ هستند، محصل بودن و خانه‌دار بودن بر تعداد خرید اینترنتی در حیطه خدمات آموزشی و فرهنگی تأثیرگذار هستند و این از روی نمودار ۱۰ نیز مشخص است.

در جدول ۴ نیز با توجه به معیار آکائیکه^۱ (AIC) رگرسیون دوسطحی با پاسخ دوجمله‌ای (رابطه ۷) در همه حیطه‌های ملزومات زندگی، سرگرمی و خدمات آموزشی فرهنگی به دلیل ساختار سلسله مراتبی داده‌ها از رگرسیون معمولی با پاسخ دوجمله‌ای (رابطه ۴) بهتر عمل می‌کند (هر چه مقدار AIC کمتر باشد مدل بهتر است، برای اطلاعات بیش‌تر در رابطه با این معیار به Sakamoto et al. (1986: 81) مراجعه فرمایید.

¹ Akaike Information Criterion

جدول ۱- برآورد ضرایب رگرسیونی در حیطه سرگرمی به روش رگرسیون دوسطحی با پاسخ دوجمله‌ای

متغیرهای تحت مطالعه	برآورد ضرایب	انحراف استاندارد	p-value
عرض از مبدأ	-2/453	0/279	< ۱۶e-۲*
سن	-۰/۰۳۷	0/003	< ۱۶e-۲*
شهری	0/247	0/069	0/0003*
پایه (روستایی)	-	-	-
جنسیت	-	-	-
مرد	0/385	0/061	3/96e-۱۰ *
پایه (زن)	-	-	-
وضع سواد و مدرک تحصیلی	-	-	-
پایه (دیپلم)	-0/842	0/141	2/87e-9 *
زیر دیپلم	-0/186	0/105	۰/۰۷
تحصیلات دانشگاهی	-	-	-
وضع فعالیت	-	-	-
شاغل	0/687	0/216	0/001*
بیکار	0/848	0/227	0/0001*
محصل	0/572	0/217	0/008*
دارای درآمد بدون کار	0/501	0/279	۰/۰۷
خانه‌دار	0/468	0/229	0/041*
پایه (سایر)	-	-	-

جدول ۲- برآورد ضرایب رگرسیونی در حیطه ملزومات زندگی به روش رگرسیون دوسطحی با پاسخ دوجمله‌ای

متغیرهای تحت مطالعه	برآورد ضرایب	انحراف استاندارد	p-value
عرض از مبدأ	-2/147	0/181	< ۱۶e-۲*
سن	-0/007	0/002	0/0001 *
شهری	0/518	0/06	< ۱۶e-۲*
پایه (روستایی)	-	-	-
جنسیت	-	-	-

متغیرهای تحت مطالعه	برآورد ضرایب	انحراف استاندارد	p-value
مرد	-0/281	0/046	1/63e-9 *
پایه (زن)	-		-
وضع سواد و مدرک تحصیلی	-	-	-
پایه (دیپلم)	-0/894	0/114	6e-15 *
زیر دیپلم	-0/459	0/113	5/24e-5 *
تحصیلات دانشگاهی			
وضع فعالیت			
شاغل	0/359	0/158	0/023*
بیکار	0/05	0/17	0/765
محصل	-0/098	0/161	0/544
دارای درآمد بدون کار	0/211	0/189	0/264
خانه‌دار	0/01	0/164	0/949
پایه (سایر)	-	-	-

جدول ۳- برآورد ضرایب رگرسیونی در حیطه خدمات آموزشی فرهنگی به روش

رگرسیون دوسطحی با پاسخ دوجمله‌ای

متغیرهای تحت مطالعه	برآورد ضرایب	انحراف استاندارد	p-value
عرض از مبدأ	-2/797	0/275	< ۱۶e-۲*
سن	0/0005	0/003	0/866
شهری	0/212	0/081	0/008*
پایه (روستایی)	-		-
جنسیت			
مرد	-0/213	0/064	0/0009*
پایه (زن)	-		-
وضع سواد و مدرک تحصیلی			
پایه (دیپلم)	-	-	-
زیر دیپلم	-1/081	0/189	1/11e-08 *
تحصیلات دانشگاهی	-0/762	0/157	1/31e-06*

متغیرهای تحت مطالعه	برآورد ضرایب	انحراف استاندارد	p-value
وضع فعالیت			
شاغل	0/281	0/242	0/244
بیکار	0/255	0/258	0/322
محصل	0/727	0/245	0/003*
دارای درآمد بدون کار	-0/177	0/297	0/55
خانه‌دار	-0/638	0/258	0/013*
پایه (سایر)	-	-	-

جدول ۴- مقایسه رگرسیون دوسطحی با پاسخ دوجمله‌ای و رگرسیون معمولی بر اساس معیار آکائیکه (AIC)

AIC			روش مورد استفاده
خدمات آموزشی فرهنگی	ملزومات زندگی	سرگرمی	
6176/2	9503/4	7599/2	رگرسیون دو سطحی برای پاسخ دوجمله‌ای
6232/5	9575/9	8671/7	رگرسیون معمولی برای پاسخ دوجمله‌ای

بحث و نتیجه‌گیری

امروزه اهمیت تأثیر استفاده از اینترنت در امور مختلف زندگی بشر بر کسی پوشیده نیست. با توجه به میزان این اهمیت در این مقاله سعی شد از اطلاعات طرح آمارگیری «استفاده‌ی خانوارها و افراد از فناوری اطلاعات و ارتباطات مرکز آمار ایران در سال ۱۳۹۶ برای ۳۱ استان کشور» برای بررسی عوامل تأثیرگذار در استفاده از اینترنت در زمینه‌های سرگرمی، خرید ملزومات زندگی و خدمات آموزشی-فرهنگی استفاده شود. مدل رگرسیون چند سطحی رویکردی است که امکان همبستگی پاسخ‌ها در گروه یکسان و مدل‌سازی داده‌هایی با ساختار سلسله‌مراتبی را ایجاد می‌کند. در این مطالعه یک سری داده با ساختار سلسله‌مراتبی تولید شد و مدل رگرسیون دو سطحی با پاسخ دوجمله‌ای روی آن اجرا شد که نتایج حاصل از شبیه‌سازی بیانگر مناسب بودن این مدل برای تحلیل داده‌هایی با ساختار سلسله‌مراتبی است. متغیر پاسخ به صورت تعداد کالاهایی که یک فرد در حیطه‌های مختلف (سرگرمی، ملزومات

زندگی و خدمات آموزشی فرهنگی) موفق به خرید آن‌ها به صورت اینترنتی می‌شود، تعریف می‌شود و متغیرهای کمکی به کاربرده شده در این مطالعه شامل سن، سطح تحصیلات (بیسواد، ابتدایی و راهنمایی، متوسطه و پیش دانشگاهی، تحصیلات غیر تکمیلی دانشگاهی و تحصیلات تکمیلی دانشگاهی)، جنس، وضع فعالیت (شاغل، بیکار، خانه دار و غیرفعال مانند سرباز و محصل و...) و شهری یا روستایی بودن، هستند. در واقع در این مطالعه تأثیرگذاری متغیرهای کمکی بر متغیر پاسخ با استفاده از رگرسیون دو سطحی با پاسخ دوجمله‌ای مورد بررسی قرار گرفت که نتایج بیانگر تأثیرگذاری متغیر سن بر خرید اینترنتی افراد در حیطه سرگرمی است ولی این متغیر بر تعداد خرید افراد در حیطه ملزومات زندگی و خدمات آموزشی فرهنگی تأثیرگذار نیست. جنسیت و نوع اقامت نیز بر تعداد خرید اینترنتی افراد در حیطه سرگرمی و ملزومات زندگی تأثیرگذار است ولی بر تعداد خرید افراد در حیطه خدمات آموزشی فرهنگی تأثیرگذار نیست. سطح تحصیلات بر تعداد خرید اینترنتی افراد در حیطه‌های مختلف سرگرمی، ملزومات زندگی و خدمات آموزشی و فرهنگی تأثیرگذار است و هرچه سطح تحصیلات افراد بالاتر باشد تعداد خرید اینترنتی بیش‌تری در این حیطه‌ها صورت می‌گیرد. وضع فعالیت نیز تأثیر چندانی بر تعداد خرید اینترنتی افراد در حیطه‌های مختلف سرگرمی، ملزومات زندگی و خدمات آموزشی و فرهنگی ندارد. اجرای مدل رگرسیونی دوسطحی با پاسخ دو جمله‌ای و مدل رگرسیون معمولی با پاسخ دوجمله‌ای بر داده‌ها و مقایسه آن‌ها با معیار AIC بهتر بودن مدل رگرسیون دوسطحی با پاسخ دو جمله‌ای برای داده‌ها را تأیید کرد.

منابع

- الفت، لعیا. خسروانی، فرزانه. جلالی، رضا. (۱۳۹۰). «شناسایی عوامل موثر بر خرید اینترنتی و اولویت‌بندی آنها با استفاده از ANP از فازی». **دانشکده مدیریت دانشگاه تهران، مدیریت بازرگانی**. شماره ۷. صص ۳۶-۱۹.
- بابائی فرد، اسداله. حیدریان، امین. (۱۳۹۷). «مطالعه تأثیر اینترنت بر سرمایه فرهنگی دانشجویان». **فصلنامه انجمن ایرانی مطالعات فرهنگی و ارتباطات**، شماره ۵۶. صص ۲۲۵-۲۵۰.
- بهار، مهری. حاجی‌محمدی، علی. (۱۳۸۶). «دانشجویان دانشگاه تهران و استفاده از اینترنت: مطالعه بهره‌وری و خرسندی». **فصلنامه انجمن ایرانی مطالعات فرهنگی و ارتباطات**، ۱۰.
- پاک‌سرشت، سلیمان. نوری‌نیا، حسین. (۱۳۸۶). «بررسی پیامدهای کاربرد فراغتی اینترنت بر رفتارهای فراغتی جوانان تهرانی». **فصلنامه انجمن ایرانی مطالعات فرهنگی و ارتباطات**، ۱۰.
- خانیکی، هادی. عابدی، حمید. (۱۳۹۳). «تأثیر اینترنت بر نگرش فرهنگی روحانیون (مطالعه موردی طلاب حوزه‌های علمیه قم)». **فصلنامه انجمن ایرانی مطالعات فرهنگی و ارتباطات**، شماره ۳۶. صص ۴۳-۱۱.
- دادگران، محمد. (۱۳۸۵). **مبانی ارتباطات جمعی**. تهران: فیروزه.
- رندی، وحید. محسن خون، سیاوش. معصومی، بهروز. (۱۳۹۳). «عوامل موثر بر رفتار خرید مشتریان اینترنتی در ایران با توجه به مدل پذیرش فناوری (TAM)». **مجله مدیریت توسعه و تحول (ویژه نامه)**، شماره ۴۳. صص ۱۱۸-۱۰۹.
- ساروخانی، باقر. (۱۳۸۳). **اندیشه‌های بنیادین علم ارتباطات**. تهران: خجسته.
- سورین، ورنر جوزف. تانکارد، جیمز دبلیو. (۱۳۸۴). **نظریه‌های ارتباطات**. ترجمه: علیرضا دهقان، تهران: دانشگاه تهران، مؤسسه انتشارات و چاپ.
- عبداللهمیان، حمید. (۱۳۸۴). «نوع‌شناسی و بازتعریف آسیب‌های اینترنتی و تغییرات هویتی در ایران». **فصلنامه انجمن ایرانی مطالعات فرهنگی و ارتباطات**، ۳ و ۲.
- فاضلی، نعمت‌الله. (۱۳۸۵). «دیجتالی شدن به سبک ایرانی و ایرانی شدن به سبک دیجیتالی». **فصلنامه رسانه**، ۶۷.
- کاستلز، مانوئل. (۱۳۸۵). **عصر اطلاعات: اقتصاد، جامعه و فرهنگ (پایان هزاره)**. ترجمه: احد علیقلیان و افشین خاکباز، تهران: طرح نو.
- معتمدنژاد، کاظم. (۱۳۸۹). **جامعه اطلاعاتی: اندیشه‌های بنیادی، دیدگاه‌های انتقادی و چشم‌اندازهای جهانی**. تهران: میراث قلم.

- مهدی‌زاده، شراره. عنبرین، زینب. (۱۳۸۸). «بررسی رابطه میزان و نحوه استفاده از اینترنت و هویت فرهنگی جوانان». فصلنامه انجمن ایرانی مطالعات فرهنگی و ارتباطات، شماره ۱۷، صص ۱-۲۸.
- وبستر، فرانک. (۱۳۸۳). *نظریه‌های جامعه‌اطلاعاتی*. ترجمه: اسماعیل قدیمی، تهران: قصیده‌سرا.
- Holland, P. W. and Leinhardt, S. (1981). "An exponential family of probability distributions for directed graphs", *Journal of The American Statistical association*, No.373. PP: 33-50.
- Lawson, R. Alcock, C. and Cooper, J. (2002). "Diffusion of electronic commerce in small and medium enterprises", *Australasian Journal of Information Systems*, No.2. PP:123-134.
- Maas, C. J. and Hox, J. J. (2004). "Robustness issues in multilevel regression analysis", *Statistica Neerlandica*, No.2. PP: 127-137.
- McCoy, S. and Mbarika, V. (2005). "Teledensity growth strategies for Latin America: the case of Colombia and Ecuador", *Communications of the Association for Information Systems (CAIS)*, No.2. PP: 26-56.
- Nelder, J. A. and Wedderburn, R. W. 1972. "Generalized linear models", *Journal of the Royal Statistical Society: Series A (General)*, No.3. PP: 370-384.
- Rogers, E. M. 1995. "Diffusion of innovations, in diffusion of innovation", The Free Press, New York.
- Sakamoto, Y. Ishiguro, M. and Kitagawa, G. 1986. "Akaike information criterion statistics", *Dordrecht, The Netherlands: D. Reidel*, No. 81.
- Salman, A. and Hasim, M.S. 2011. "Internet usage in a Malaysian sub-urban community: a study of diffusion of ICT innovation", *The Innovation Journal: The Public Sector Innovation Journal*, No.2. PP: 1-15.
- Wacholder, S. (1986). "Binomial regression in GLIM: estimating risk ratios and risk differences", *American journal of epidemiology*, No.3. PP: 174-184.
- Wackerly, D. Mendenhall, W. and Scheaffer, R. L. 2014. "Mathematical statistics with applications", Cengage Learning.
- Wong, G. Y. and Mason, W. M. (1985). "The hierarchical logistic regression model for multilevel analysis", *Journal of the American Statistical Association*, No.391. PP:513-524.