

بررسی وضع اقتصادی اعضای هیأت علمی دانشگاهها در مقایسه با دارندگان تخصصهای مشابه در سایر مراکز دولتی و غیر دولتی

محمد رضا مشکانی
دانشگاه شهید بهشتی

زهرا سادات مشکانی
دانشگاه علوم پزشکی تهران

چکیده

در این بررسی، بر اساس نمونه‌ای تصادفی، به مقایسه وضع اقتصادی و رفاهی اعضای هیأت علمی دانشگاهها و مؤسسات آموزش عالی کشور و همالان آنها در سایر مراکز دولتی و غیر دولتی پرداخته‌ایم. مبنای مقایسه، هشت قلم هزینه عمده خانوار است در سال تحصیلی (۷۲-۷۳). نظر به تغییرات زیاد این هزینه‌ها در بین افراد نمونه، به کمک روشهای آماری چند متغیری از قبیل رگرسیون چندگانه و رگرسیون لوجستیک چندگانه، عوامل دخیل در هزینه‌ها را مشخص کرده‌ایم. پس از تعیین عوامل مؤثر و تثبیت اثر آنها، اقلام هزینه را بر حسب شغل اصلی، یعنی عضویت در هیأت علمی دانشگاهها یا خدمت در سایر مراکز، مقایسه کرده‌ایم. نتایج به دست آمده حاکی از آن است که هزینه افراد نمونه در مقایسه با هزینه خانوارهای شهری، که مرکز آمار ایران منتشر کرده، کمتر است و این خود نشانه نامناسب بودن وضع اقتصادی و رفاهی تحصیلکردگان است.

واژگان کلیدی: وضع اقتصادی، اعضای هیأت علمی، هزینه خانوار، رگرسیون لوجستیک، رگرسیون چندگانه

۱- کلیات

۱-۱. مقدمه

اعضای هیأت علمی دانشگاههای کشور، تربیت متخصصان و مدیران آینده مملکت در بخشهای فنی، علمی، فرهنگی، اقتصادی و اجتماعی رابه عهده دارند. وظیفه مهم دیگری که در کنار آموزش به عهده آنان گذاشته شده انجام پژوهشهای علمی برای برآوردن نیازهای مملکت و پیشبرد مرزهای علم است. این وظایف چنان گسترده و فراگیرند که همه سطوح و طبقات مملکت از فنی، اداری، اجرایی گرفته تا آموزشی، فرهنگی، نظامی، قضایی و سیاسی را دربرمیگیرند. اعضای هیأت علمی باید عالمترین و دانشمندترین و خبرهترین افراد در کار خود باشند تا بتوانند انسانهایی را تربیت کنند که قادر به تأمین امنیت، سلامت و سعادت ملت باشند. زیرا ذات نیافته از هستی بخش، کی تواند که شود هستی بخش.

این افراد با رسالت بزرگی که به عهده دارند، و نیز به سبب داشتن درجههای بالای تخصص و توانایی لازم برای احراز مقام علمی و سرمایه گذاری یی که برای به دست آوردن مدرک تحصیلی کرده اند طبیعی است که انتظار داشته باشند در زندگی اجتماعی از رفاه و امکانات معقول برخوردار شوند، تا بتوانند با فراغ خاطر وظیفه مهم خود را به نیکوترین وجه انجام دهند و از رفاه اقتصادی به گونه ای شایسته مقام علمی استادان دانشگاهها برخوردار باشند. از اینرو بررسی وضع اقتصادی و رفاهی اعضای هیأت علمی و مقایسه آن با همالان آنها در سایر مراکز به منظور درک حقایق جاری در جامعه و آگاهی از نقاط ضعف و قوت نظام دانشگاهی کشور، برای راضی نگهداشتن کارکنان خود، مسأله ای ضروری و در شرایط کنونی کشور امری الزامی است. زیرا به نظر می رسد که در سالهای اخیر به دلایل گوناگون، بین وضع رفاهی اعضای هیأت علمی دانشگاهها و متخصصان شاغل در خارج از دانشگاه تعادل برقرار نیست. این امر از دو جهت به زیان نظام آموزش عالی کشور است، اول فارغ التحصیلان جدید، انگیزه و تمایلی برای خدمت در دانشگاهها نشان نخواهند داد و در نهایت کشور از نیروی خلاقه آنها در آموزش و پژوهش محروم خواهد شد. دوم، فشارهای اقتصادی و نیاز به کسب درآمد بیشتر برای تأمین نیازهای ضروری خود و خانواده، اعضای هیأت علمی را به مشاغل دیگری می کشاند و آنها را از شغل اصلی خود که تدریس و تحقیق و پرورش دانشجویان است، باز می دارد.

در این پژوهش، قصد آن است که با بررسی عینی و به روش علمی به زمینه یابی پردازیم و

گمانهای مختلف را که به نظر می‌رسند ناشی از اوضاع و احوال کنونی هستند، به محک تجربه بزنیم و درستی یا نادرستی آنها را آشکار سازیم. به طور مشخص، می‌خواهیم موقعیت اقتصادی و رفاهی اعضای هیأت علمی دانشگاهها را در قیاس با همالان آنها در خارج از دانشگاه، که رقیب کاری آنها در صحنه کشور محسوب می‌شوند، بررسی و مقایسه کنیم.

۱-۲. طرح مسأله

مسأله مورد نظر در این پژوهش را می‌توان به شرح زیر تعریف و حدود آن را معین کرد. موقعیت اعضای هیأت علمی دانشگاهها و مؤسسات آموزش عالی کشور از حیث وضع اقتصادی و رفاهی، در مقایسه با دارندگان تخصصهای مشابه در سایر مراکز دولتی و غیر دولتی، چگونه است؟ دو گروه از افراد را در نظر بگیریم که دارای تحصیلات مشابه‌اند، گروهی از آنها در دانشگاهها و گروهی در خارج از دانشگاه و در یکی از مراکز اداری، پژوهشی، خدمات،... در بخش دولتی یا خصوصی کار می‌کنند. وضع اقتصادی این دو گروه را با هم مقایسه می‌کنیم تا مشخص شود که آیا این دو گروه از حیث وضع اقتصادی و رفاهی با هم تفاوت دارند یا خیر؟ اگر تفاوتی وجود دارد، کدام گروه از رفاه بیشتری برخوردار است؟ در صورت وجود تفاوت، در کدام جنبه‌ها تفاوت بارز است؟ چه عواملی موجب تفاوت شده است؟ و بالاخره آیا به اعضای هیأت علمی دانشگاهها، متناسب با تحصیلات و تخصص شان توجه می‌شود یا خیر؟

۱-۳. جامعه آماری

جامعه آماری تحت بررسی از دو زیر جامعه مجزا تشکیل شده است که آنها را با هم مقایسه می‌کنیم.

الف - جامعه اعضای هیأت علمی دانشگاهها و مؤسسات آموزش عالی کشور
منظور از عضو هیأت علمی دانشگاهها و مؤسسات آموزش عالی کسی است که شغل اصلی و سازمانی او در یکی از دانشگاهها و مؤسسات آموزش عالی، یا مراکز آموزش عالی رسمی وابسته به وزارت فرهنگ و آموزش عالی است. البته تعدادی مؤسسه آموزش عالی وابسته به سایر وزارتخانه‌ها در زمان انجام این پژوهش مشغول فعالیت بوده‌اند، که به سبب وضع خاص آنها و قوانین خاص استخدامی، اعضای آنها در این جامعه گنجانده نشده‌اند.

ب - جامعه همالان اعضای هیأت علمی شاغل در سایر مراکز دولتی و غیر دولتی.
این جامعه در بر دارنده متخصصانی است که دارای درجه کارشناسی ارشد یا بالاتر هستند

و در مراکزی غیر از دانشگاهها و مؤسسات آموزش عالی به کار مشغول اند. به بیان دیگر، این هیأت علمی دانشگاهها و مؤسسات آموزش عالی درآیند. از این رو از لحاظ تحصیلات همانند اعضای هیأت علمی دانشگاهها و مؤسسات آموزش عالی هستند ولی از حیث وظیفه و کاری که انجام می دهند، ممکن است به یکی از کارهای اداری، پژوهشی، آموزشی، یا ترکیبی از آنها بپردازند و در ساعات فراغت به تدریس در دانشگاهها و مؤسسات آموزش عالی مشغول باشند.

۱-۴. فرضیه‌ها

برای پاسخگویی به سؤالهای مذکور در بخش ۱-۲ فرضیه‌هایی را در نظر می‌گیریم که با رد یا قبول آنها می‌توان نتیجه گرفت که آیا وضع اقتصادی اعضای هیأت علمی دانشگاهها بهتر از وضع اقتصادی همالان آنها در سایر مراکز دولتی و غیر دولتی است. این فرضیه‌ها عبارتند از:

- با توجه به هزینه‌های جاری زندگی امروزی، و در قیاس با عامه مردم، اعضای هیأت علمی دانشگاهها، به طور کلی، وضع اقتصادی رضایت بخشی دارند.

- اعضای هیأت علمی دانشگاهها در مقایسه با دارندگان تخصصهای مشابه در سایر مراکز دولتی و غیر دولتی، دستمزدی متناسب با تحصیلات و تخصص خود دریافت می‌کنند.

- الگوی زندگی اعضای هیأت علمی و همالان آنها با یکدیگر تفاوت ندارد و این خود ناشی از میزان درآمد یکسان آنهاست.

- اعضای هیأت علمی برای جبران کمبود درآمد ناچار به اشتغال در مشاغل دیگر نیستند.

۱-۵. روش پژوهش

روش این پژوهش به صورت بررسی مقطعی مبتنی بر تحلیلهای آماری است. بررسی در مقطع زمانی سال تحصیلی ۱۳۷۲-۱۳۷۳ صورت گرفته است. پس از تهیه فهرست دانشگاهها و مؤسسات آموزش عالی و اعضای هیأت علمی آنها از لحاظ درجه و میزان تحصیلات، از بین آنها نمونه‌گیری به روش طبقه‌بندی به عمل آمد و با ایجاد همگروهی تصنعی از بین همالان هیأت علمی نیز نمونه‌گیری شد. پرسشنامه‌ای که برای گردآوری اطلاعات تنظیم شده بود، بین افراد نمونه توزیع گردید. به آمارگیران آموزش لازم درباره نحوه انجام مصاحبه و توضیح سؤالها و دریافت پاسخها داده شده بود، به طوری که درک همه مصاحبه شونده‌ها از مفهوم سؤالها واحد باشد و جوابها در مقیاسی واحد سنجیده شود، تا امکان جمع بستن آنها و مقایسه های لازم وجود داشته باشد. (مصاحبه‌ها از دی ۷۲ تا مرداد ۷۳ انجام شد.)

۱-۶. روش نمونه‌گیری

با توجه به خصوصیات افراد دو زیر جامعه، که در رشته‌های مختلف علمی تخصص دارند و با توجه به عرضه و تقاضای مربوط به تخصصها و نیز مکان زندگی آنها در تهران، مراکز استانها، یا شهرستانها که در برخی موارد امکان فعالیتهای جنبی را محدود می‌ساخت، مناسبترین طرح نمونه‌گیری برای انتخاب افراد نمونه، طرح نمونه‌گیری با طبقه‌بندی تشخیص داده شد. در طرح نمونه‌گیری با طبقه‌بندی هر اندازه، همگنی افراد درون طبقات بیشتر و همگنی افراد بین طبقات مختلف کمتر باشد، دقت بیشتری در برآورد صفات حاصل می‌شود. پس در این بررسی، به منظور انتخاب نمونه مسأله اساسی تعریف طبقات است. برای معرفی طبقات، باید چهارچوب کاملی از کلیه اعضای جامعه آماری و تعلق آنها به یکی از دو زیر جامعه در اختیار باشد. البته چنین چهارچوبی وجود نداشت و تهیه آن نیز با توجه به گستردگی رشته‌های تحصیلی و تفاوت عنوانهای ذکر شده در مدارک تحصیلی کشورهای گوناگون دنیا و اشتغال افراد به کاری غیر از رشته تحصیلی خود، کاری بس مشکل و احياناً ناممکن بود. در عوض از طبقه‌بندی دیگری استفاده شد که اشکالات فوق را نداشت، و مطابق با عرف معمول در طبقه‌بندی رشته‌های دانشگاهی بود. در این طبقه‌بندی کلیه رشته‌های دانشگاهی به پنج طبقه تقسیم می‌شوند:

- رشته‌های دانشگاهی علوم پایه از قبیل آمار، ریاضی، فیزیک، شیمی، زیست‌شناسی، زمین‌شناسی و شاخه‌های تخصصی آنها.
 - رشته‌های دانشگاهی علوم انسانی از قبیل ادبیات، روانشناسی، علوم تربیتی، اقتصاد، حقوق، تاریخ، جغرافیا، و شاخه‌های تخصصی آنها.
 - رشته‌های دانشگاهی هنر از قبیل معماری، نقاشی، موسیقی، فیلم و سینما، نمایش و شاخه‌های تخصصی آنها.
 - رشته‌های دانشگاهی کشاورزی، دامپروری، دامپزشکی و شاخه‌های تخصصی آنها.
 - رشته‌های دانشگاهی مهندسی از قبیل مهندسی برق، مهندسی شیمی و پتروشیمی، مهندسی راه و ساختمان، مهندسی مکانیک و شاخه‌های تخصصی آنها.
- درست است که در این طبقه‌بندی، برخی طبقات مانند علوم انسانی گستره وسیعی دارند و تقاضا برای خدمات اعضای آن از شاخه‌های مختلف نظیر اقتصاد، حقوق و ادبیات تفاوت

زیادی دارد، نظر به اینکه از معیار واحدی برای طبقه‌بندی دو زیر جامعه الف و ب استفاده شده است، در نهایت " این تفاوت تأثیر عمده‌ای بر نتایج نخواهد داشت. زیرا با ایجاد همگروهی تصنعی از هر رشته تخصصی که در جامعه الف نمونه داشته باشیم در جامعه ب نیز وجود خواهد داشت و چون تفاوت وضع رفاهی افراد دو جامعه برای ما مهم است، گستردگی طبقه و وجود ناهمگنی در آن چندان مهم نخواهد بود.

دلیل عمده برای استفاده از این طبقه‌بندی، وجود چهارچوب نسبتاً کامل و بهنگامی بود که مقارن با تدوین طرح نمونه‌گیری این تحقیق، از طرف اعضای گروه پژوهش‌های آماری وزارت فرهنگ و آموزش عالی در مؤسسه پژوهش و برنامه‌ریزی آموزش عالی در پنج مجلد انتشار یافت. این نشریات با عنوان شناخت ویژگیهای اعضای هیأت علمی دانشگاهها و مؤسسات آموزشی و پژوهشی در سال ۱۳۷۱ جدیدترین و بهنگامترین چهارچوب در دسترس بود و به همین دلیل طرح نمونه‌گیری چنان تدوین شد که بتوان از اطلاعات آن حداکثر بهره را برد (گروه پژوهش‌های آماری، ۱۳۷۲).

چهارچوب مشابهی برای افراد جامعه ب موجود نبود. تهیه آن نیز از یک گروه پژوهشی یا یک سازمان علمی بر نمی‌آمد، بلکه نیاز به همکاری کلیه مؤسسات علمی و اداری کشور داشت. از این رو برای انتخاب نمونه‌های رقیب از جامعه ب به تمهیدی پرداختیم که به همگروهی تصنعی مشهور است، در این روش از افراد نمونه جامعه الف خواسته شد که محل کار و در صورت امکان حداقل دو نفر از همالان خود را که در سازمانهای غیر دانشگاهی به کار اشتغال دارند معرفی کنند. برای کاستن از اربیبی حاصل از آشنایی افراد نمونه با همالان خود، از بین افرادی که در محل‌های معرفی شده به کار اشتغال داشتند، از بین معرفی شدگان، یک نفر را به قید قرعه به عنوان نمونه نهایی انتخاب کردیم.

آنچه بر اثر نمونه‌گیری تصادفی با طرح بالا از تخصصهای گوناگون به دست آمد از این قرار است: آمار (۱ نفر)، آموزش زبان (۳ نفر)، الهیات و معارف اسلامی (۳ نفر)، انسان‌شناسی (۱ نفر)، اقتصاد (۱ نفر)، برق و الکترونیک (۴ نفر)، بیوشیمی (۴ نفر)، تاریخ (۱ نفر)، جامعه‌شناسی (۶ نفر)، جنگلداری (۲ نفر)، حشره‌شناسی (۱ نفر)، حقوق (۲ نفر)، حسابداری (۱ نفر)، دامپزشکی (۲ نفر)، روانشناسی (۲ نفر)، ریاضی (۵ نفر)، زبان‌شناسی (۳ نفر)، زبان و ادبیات فارسی (۱ نفر)، زمین‌شناسی (۳ نفر)، ژنتیک (۱ نفر)، ژئوفیزیک (۱ نفر)، شیمی (۱ نفر)، علوم تربیتی (۶ نفر)، علوم

سیاسی (۱ نفر)، علوم کامپیوتر (۳ نفر)، فیزیک (۵ نفر)، کتابداری (۱ نفر)، کشاورزی (۷ نفر)، مدیریت صنعتی (۱ نفر)، متالوژی (۱ نفر)، معماری (۱ نفر)، مهندسی راه و ساختمان (۱ نفر)، مهندسی شیمی (۸ نفر)، مهندسی صنایع (۲ نفر)، مهندسی مخابرات (۱ نفر)، مهندسی مکانیک (۴ نفر)، مهندسی مواد (۱ نفر)، و هواشناسی (۱ نفر).

پیدا است که نمونه انتخابی از ویژگیهای جامعیت و معرف بودن کافی برخوردار است و باید به نتایج حاصل از آن بتوان اعتماد کرد. چون در هر نمونه گیری تعدادی بی پاسخ وجود دارد، حجم نمونه حدود ۱۰ درصد بیشتر از حجم مطلوب محاسبه شد.

۱-۷. متغیرهای مورد بررسی

در این بررسی، هدف آن است که وضع اقتصادی خانواده‌های اعضای هیأت علمی دانشگاهها و مؤسسات آموزش عالی را با همالان آنها در سایر مراکز دولتی و غیر دولتی مقایسه کنیم. طبیعی است که باید متغیرهایی را که نشانه رفاه اقتصادی و برخورداری از امکانات و شرایط زندگی متناسب با شأن و حرفه اعضای این جامعه‌هاست، برگزینیم. بدین لحاظ انواع درآمدها از قبیل حقوق و مزایای شغل، فوق‌العاده‌ها، حقوق ناشی از اضافه کاریها، مأموریتها، و کارانه‌ها و خلاصه هر نوع درآمد دیگر که ناشی از تصدی شغل باشد، مناسبترین متغیرها برای بررسی وضع رفاه اقتصادی عضو هیأت علمی و هم‌تای آن است. در این بررسی تنها درآمدهای ناشی از تصدی شغل مورد نظر است. درآمدهای موروثی و اتفاقی خارج از بحث است.

از آنجا که جامعه مورد بررسی با سوادترین و آگاه‌ترین قشر جامعه است، امیدوار بودیم که پاسخ دهندگان، پس از آگاهی از هدف پژوهش و عنایت به اینکه مسأله مورد بررسی مربوط به خود آنان است، در مورد درآمدهای خود پاسخهای صحیح بدهند. با این امید پرسشنامه‌ای تنظیم شد که در آن مستقیماً به اندازه‌گیری انواع درآمدها بی ذکر نام پاسخ دهنده ولی همراه با مشخصات فردی، سوابق کار، بعد خانوار، اشتغال همسر و وضعیت منزل مسکونی و مکان جغرافیایی محل خدمت او، پرداخته می‌شد. پس از آزمایش مقدماتی پرسشنامه بر روی نمونه‌ای کوچک از اعضای هیأت علمی و همالان آنها، داده‌های به دست آمده واقعیتی را آشکار ساختند که تا حدودی انتظار می‌رفت ولی به واسطه ویژگی جامعه مورد مطالعه خوشبین بودیم که چنین نشود. زیرا امیدوار بودیم که این قشر نخبه به طور استثنایی عمل می‌کنند و به سؤالهای مربوط به درآمدهای خود پاسخ درست می‌دهند. متأسفانه این امیدواری تحقق نیافت چرا که

ارقام مربوط به درآمدها غیر واقعی بود. در برخی موارد پاسخها به شوخی برگزار شده و برخی با ذکر معاذیری از پاسخ درست دادن طفره رفته بودند. با توجه به این مشکل، که تقریباً همه جادیده می شود و حتی در کشورهای پیشرفته دنیا نیز با آن مواجه اند، تصمیم گرفتیم استراتژی گردآوری اطلاعات را تغییر دهیم. از این رو سنجش هزینه خانوارها در جامعه تحت بررسی برای ادامه مطالعه انتخاب شد.

چون میزان هزینه های مختلف تا حدودی برای همه شناخته شده است و بر اساس قیمت های جاری تقریباً اطلاع کمابیش درستی از آنها در دست است، احتمال طفره رفتن در ذکر هزینه ها و اعلام ارقام غیر واقعی کمتر است.

مبنای نظری گزینش استراتژی فوق این است که هر هزینه ای که در خانوار صورت می گیرد باید متکی به اعتباری باشد که قبلاً تأمین شده است. در صورتی که هزینه ها به صورت نسیه یا اقساطی باشند، حتماً خانوار محل تأمین اعتبار آن را اندیشیده است و به هر حال برای آن چاره ای در نظر گرفته است. به طور کلی، می توان گفت که میزان هزینه های هر خانوار به خوبی شاخص درآمد خانوار است. در اکثر منابع نیز به این روش اشاره شده و اصولاً یکی از راه های تعیین و محاسبه تولید ناخالص ملی نیز سنجش هزینه ها است. نمونه بارز این نحوه سنجش در کشور ما " طرح آماری از هزینه و درآمد بودجه خانوار " است که زمانی بانک مرکزی ایران آن را اجرا می کرد و سال های اخیر این وظیفه به مرکز آمار ایران محول شده است. سابقه طولانی این طرح که از سال ۱۳۴۲ شروع شده و تا کنون هر ساله به اجرا درآمده است نشانه و دلیل اهمیت آن در تعیین وضع اقتصادی خانوارهاست. در این طرح هر خانوار ضمن تکمیل پرسشنامه ای مشتمل بر ۴۰ صفحه از سؤالات متنوع و متعدد و بسیار مفصل، درآمدها و هزینه های خود را اعلام می کند. نتایج این طرح که هر ساله، پس از تجزیه و تحلیل، منتشر می شوند، نشان می دهد که در اغلب موارد سنجش هزینه ها دقیقتر از سنجش درآمدها بوده است. مثلاً نتایج بررسی بودجه خانوار در مناطق شهری برای سال ۷۲ که در صفحه ۱۱ شماره ۲ ۱۵۳۰۲ روزنامه کیهان در تاریخ ۲۰ اسفند ۱۳۷۳ منتشر شد، حاکی از آن است که در سال ۱۳۷۲ " متوسط هزینه ناخالص سالیانه یک خانوار شهری برابر ۶۳۷۶۶۲۰ ریال (ماهانه معادل ۵۳۱۳۸۵ ریال) بوده است... و مجموع درآمد پولی و غیر پولی ناخالص یک خانوار شهری در سال ۱۳۷۲ برابر ۸۳۴۴۲ ریال " بوده است.

کسری درآمد ماهیانه، معادل ۸۳۴۴۲ ریال، برای هر خانوار فقط به این ترتیب توجیه کردنی است که خانوارها از ذکر درآمد واقعی خود طفره رفته‌اند و گرنه چگونه می‌توانند از عهده هزینه‌ای که اعتبار آن تأمین نشده است، برآیند؟ این کسری درآمد سالهای قبل نیز در همین حدود گزارش شده است و در عرف آمارگیریها پذیرفتنی محسوب می‌شود.

هم از این رو در پرسشنامه تحقیق، تجدید نظر به عمل آمد و این بار به سنجش هزینه‌ها پرداخته شد. در طراحی پرسشنامه جدید، به این مطلب توجه شد که ساختار پرسشنامه باید طوری باشد که پاسخ دهندگان در مدتی کمتر از یک ساعت آن را تکمیل کنند، و گرنه بر اثر خستگی و بی‌حوصلگی جوابهای قابل اعتمادی نخواهند داد. از این رو با الهام از مفاد پرسشنامه طرح آمارگیری هزینه و درآمد بودجه خانوار، پرسشنامه‌ای طراحی شد که ضمن خلاصه بودن، اقلام بسیار مهم و تعیین کننده وضع اقتصادی خانوار را نیز در بر داشت.

الف-متغیرهای مستقل

دسته متغیرهای مستقل، شامل ۱۸ متغیر است که مشخصات فردی و خانوادگی پاسخگو را بی ذکر نام و محل کار، درجه تحصیلی،... در اختیار می‌گذارند و عبارت اند از:
متغیر X_1 - متغیر شغل اصلی که به دو گروه دانشگاهی = ۱ و غیر دانشگاهی = ۲ تقسیم بندی شده است.

متغیر X_2 - متغیر جنسیت که به دو گروه زن = ۱ و مرد = ۲ تقسیم بندی شده است.

متغیر X_3 - وضعیت تأهل که به دو گروه مجرد = ۱ و متأهل = ۲ گروه بندی شده است.

متغیر X_4 - تعداد فرزندان.

متغیر X_5 - تعداد افراد تحت تکفل.

متغیر X_6 - متغیر سن.

متغیر X_7 - آخرین مدرک تحصیلی که به دو گروه دکتری = ۱ و فوق لیسانس = ۲ تقسیم بندی شده است.

شده است.

متغیر X_8 - رشته تحصیلی که به پنج گروه علوم انسانی = ۱، علوم پایه = ۲، فنی و مهندسی

= ۳، کشاورزی و دامپزشکی = ۴ و هنر = ۵ گروه بندی شده است.

متغیر X_9 - محل اخذ مدرک تحصیلی که به دو گروه داخل کشور = ۱ و خارج کشور = ۲

گروه بندی شده است.

متغیر X_{10} - مشاغل دیگر علاوه بر شغل اصلی، به پنج گروه شغل دیگری ندارد = ۱، یک شغل علاوه بر شغل اصلی دارد = ۲، دو شغل علاوه بر شغل اصلی دارد = ۳، سه شغل علاوه بر شغل اصلی دارد = ۴، و بیش از سه شغل علاوه بر شغل اصلی دارد = ۵ تقسیم‌بندی شده است. متغیر X_{11} - وضع استخدام که به سه گروه رسمی = ۱، پیمانی = ۲، قراردادی = ۳ تقسیم‌بندی شده است.

متغیر X_{12} - سابقه کار به سال.

متغیر X_{13} - آخرین مدرک تحصیلی همسر که به پنج گروه، دیپلم = ۱، فوق دیپلم = ۲، لیسانس = ۳، فوق لیسانس = ۴ و دکتری = ۵ تقسیم‌بندی شده است.

متغیر X_{14} - شغل همسر که به چهار گروه خانه‌دار = ۱، کارمند دولت = ۲، کارمند بخش خصوصی = ۳، و عضو هیأت علمی = ۴ گروه‌بندی شده است.

متغیر X_{15} - وضع مسکن که به چهار گروه شخصی = ۱، اجاره‌ای = ۲، دولتی یا سازمانی = ۳ و منزل بستگان = ۴ تقسیم‌بندی شده است.

متغیر X_{16} - نوع منزل مسکونی که به دو گروه آپارتمانی = ۱ و ویلایی = ۲ گروه‌بندی شده است.

متغیر X_{17} - مساحت زیر بنای منزل مسکونی.

متغیر X_{18} - محل جغرافیایی سکونت که به دو گروه تهران = ۱ و سایر شهرستانها = ۲ تقسیم‌بندی شده است.

ب- متغیرهای وابسته:

این دسته شامل ۸ متغیر است که اقلام عمده هزینه ماهیانه خانوار را تشکیل می‌دهند و عبارت‌اند از:

متغیر y_1 - میزان هزینه ماهیانه تحصیل رسمی فرزندان.

متغیر y_2 - میزان هزینه ماهیانه دارو و درمان خانواده.

متغیر y_3 - میزان هزینه ماهیانه تغذیه خانواده.

متغیر y_4 - میزان هزینه ماهیانه پوشاک خانواده.

متغیر y_5 - میزان هزینه ماهیانه خدمات فرهنگی خانواده از قبیل کلاس زبان، موسیقی،

خط، نقاشی، ورزش و نظایر آنها.

متغیر ۷۶- میزان هزینه ماهیانه تفریح خانواده از قبیل سینما، تئاتر، مسابقات ورزشی، گردش داخل شهر و پارکها، بازدید از موزه‌ها، نمایشگاهها و ...

متغیر ۷۷- میزان هزینه ماهیانه مسافرت خانواده.

متغیر ۷۸- میزان اقساط ماهیانه‌ای که بابت مسکن، اتومبیل و سایر لوازم و وسائل می‌پردازند.

در مدل رگرسیونی لوجستیک متغیر وابسته با جمع کردن متغیرهای وابسته فوق و محاسبه میانه آن به دست می‌آید، که به دو گروه مجموع هزینه‌های بالاتر از میانه = ۱ و مجموع هزینه‌های کمتر از میانه = ۰ تقسیم بندی شده است.

۱-۸. روشهای آماری مورد استفاده

در این بخش آن دسته از روشهای آماری را که برای مقایسه دو جامعه دانشگاهی و غیر دانشگاهی به کار خواهیم برد، به اختصار شرح می‌دهیم.

روشهای آماری مورد استفاده در این بررسی عموماً از نوع روشهای آماری چند متغیره‌اند. دلیل انتخاب این روشها برای انجام مقایسه‌های لازم بین دو زیر جامعه دانشگاهی و غیر دانشگاهی آن است که افراد این دو زیر جامعه دارای مشخصات متفاوتی هستند: در مکانهای جغرافیایی متفاوت کار می‌کنند که در نتیجه هزینه‌های زندگی آنها و همچنین درآمدهایی که کسب می‌کنند، به دلیل همین اختلاف در مکان زندگی، با هم متفاوت است. پس هنگام مقایسه باید این تفاوت مکانی در نظر گرفته شود؛ همین طور رشته تخصصی، جنسیت، وضعیت تأهل، بعد خانوار، و درجه تحصیلی و سایر متغیرهای مستقل. از این رو باید تمهیدی به کار بریم تا بتوانیم دو فردی را که انگار از جمیع جهات مشابه‌اند و از لحاظ شغل اصلی یکی در دانشگاه کار می‌کند و دیگری بیرون از دانشگاه، با هم مقایسه کنیم. چنین کاری تنها از عهده روشهای آماری چند متغیری بر می‌آید، که در آنها با تمهیدات آماری اثر متغیرهای مستقل بر متغیر وابسته مورد نظر کنار گذاشته می‌شوند. متغیر پاسخ مورد نظر به نحوی خالص سازی می‌شود که بتوان دو زیر جامعه را از لحاظ این متغیر و بی‌بیم از تأثیر متغیرهای مزاحم مقایسه کرد.

در مسأله ما، هزینه‌های هر خانوار کاملاً متأثر از متغیرهای مستقل است. فهرست این متغیرها را در آخر بخش ۱-۷ آورده‌ایم. بنابراین، هنگام مقایسه هزینه‌های دو فرد از دو جامعه دانشگاهی و غیر دانشگاهی باید تأثیر متغیرهای دخیل در این مقایسه را در نظر بگیریم. مثلاً،

فرض کنید می‌خواهیم شخصی را که دارای درجه دکتری زمین‌شناسی است و در تهران کار می‌کند با همتای او مقایسه کنیم. در این مقایسه همه عوامل دیگر که با انواع متغیرهای مستقل سنجیده می‌شوند باید برای این دو نفر یکسان باشند، جز اینکه یکی در دانشگاه کار می‌کند و دیگری خارج از دانشگاه. پیدا کردن چنین افرادی در نمونه‌ای محدود که بر حسب تصادف انتخاب شده باشد، بسیار مشکل است. اگر هم چنین افرادی وجود داشته باشند آن قدر کم هستند که نتیجه‌گیری از آنها از لحاظ آماری قابل اعتماد نیست. پس ناچار باید به روشهایی آماری که مقایسه‌ای با اوصاف فوق را میسر می‌سازند، متوسل شد. چنین کاری در عهده روشهای آماری چند متغیری است.

اکنون روشهایی را که برای تحلیل داده‌های گردآوری شده به کار خواهیم برد به اختصار معرفی می‌کنیم.

۱-۸-۱. روشهای مقایسه آماری زیر جامعه‌ها

الف- روشهای توصیفی- برای آشنایی با نمونه‌های به دست آمده ابتدا آماره‌های توصیفی را برای هر یک از متغیرها در دو زیر جامعه دانشگاهی و غیر دانشگاهی بررسی می‌کنیم. این آماره‌ها برای هر متغیر عبارت‌اند از: میانگین، خطای معیار، باتاچاریا و جانسون (۱۹۷۷).

ب- روشهای تحلیلی- از آنجا که نتایج به دست آمده از نمونه، تصادفی است برای کسب اطمینان از حد نهایی که از بررسی نمونه‌ها به دست می‌آیند، ناچاریم از روشهای تحلیلی آمار استفاده کنیم تا دریابیم که کدام دریافتهای واقعی‌اند و کدام دریافتهای و حدسها از خطاهای نمونه‌گیری ناشی می‌شوند. روشهای تحلیلی مورد استفاده در این بررسی را معرفی می‌کنیم. علت انتخاب این روشها را هنگام تشریح چگونگی استفاده از هر مدل برای آزمون فرضیه‌ها بیان خواهیم کرد. در اینجا برای جلوگیری از تکرار و طولانی شدن گزارش از ذکر آنها خودداری و به این نکته بسنده می‌کنیم که برای آزمون فرضیه‌های تحقیق از دو روش مختلف استفاده کرده‌ایم تا بتوانیم با اطمینان بیشتری به نتیجه‌گیری بپردازیم. روشهای مورد استفاده عبارت‌اند از: رگرسیون چند متغیره- با استفاده از این روش مشخص می‌شود که انواع هزینه‌های خانوار (متغیرهای وابسته که در آخر بخش ۱-۷ معرفی کردیم) به چه صورت و با چه شدتی از مشخصات رییس و اهالی خانوار (متغیرهای مستقل که در آخر بخش ۱-۷ معرفی کردیم)

پیروی می‌کنند. به این ترتیب نقش دانشگاهی یا غیر دانشگاهی بودن رئیس خانوار در میزان هزینه‌های مختلف آشکار می‌شود. همچنین این امکان فراهم می‌شود که در صورت مؤثر بودن شغل اصلی رئیس خانوار در انواع هزینه‌ها، میزان تفاوت بین دانشگاهیان و غیر دانشگاهیان، به شرط ثابت نگهداشتن سایر عوامل، مشخص شود (جانسون و ویچرن ۱۹۸۸).

رگرسیون لوجستیک چند متغیره - استفاده از این روش که در آن متغیر پاسخ دو یا چند حالت دارد و متغیرهای مستقل می‌توانند گسسته یا پیوسته باشند، بنابر ملاحظات نظریه‌ای آمار است. روش رگرسیونی که در بالا ذکر شد لازم می‌نماید که متغیر پاسخ (در مسأله ما هزینه‌ها) دارای توزیع نرمال باشند تا استنباطها از اعتبار لازم برخوردار شوند. اما تضمین چنین شرطی آسان نیست. به منظور احتراز از خطاهای بزرگ و گمراهی ناشی از مفروضات نادرست، روش رگرسیون لوجستیک را نیز به کار می‌بریم. در این روش ابتدا هزینه‌ها را به دو دسته هزینه‌های بالا و هزینه‌های پایین تقسیم می‌کنیم. سپس وابستگی هزینه بالا یا پایین را با متغیرهای مستقل می‌سنجیم. از لحاظ مفهومی، شباهت زیادی بین این روش و روش رگرسیون وجود دارد. اما از لحاظ تکنیکی و به دلیل آنکه این روش مفروضات محدود کننده‌ای را لازم ندارد، می‌تواند به عنوان گواه به کار آید. با کار بست این روش در می‌یابیم که چه متغیرهای مستقلی سبب بالا بودن هزینه خانوار می‌شوند، که این خود حاکی از درآمد بیشتر و رفاه نسبی بیشتر است. اگر متغیر شغل اصلی (دانشگاهی یا غیر دانشگاهی) از جمله عوامل مؤثر باشد به ما امکان می‌دهد که مقایسه‌ای عینی بین این دو زیر جامعه به عمل آوریم (هاسمر و لیمی شا ۱۹۸۹).

۲- مقایسه‌های مقدماتی

در این بخش ابتدا دو زیر جامعه تحت بررسی را از لحاظ متغیرهایی که اندازه‌گیری کرده‌ایم مقایسه می‌کنیم. برای مقایسه آنها از آماره‌های توصیفی و تحلیلی بهره می‌گیریم.

۲-۱. آماره‌های توصیفی

برای آگاهی از وضع دو زیر جامعه تحت بررسی از حیث متغیرهای مستقل و وابسته، در زیر آماره‌های خلاصه را درباره تک تک متغیرها می‌آوریم. در هر مورد مقایسه‌ای بین افراد دانشگاهی و غیر دانشگاهی به عمل خواهیم آورد.

از تعداد ۹۵ نفر نمونه‌ای که اسامی آنها از بین اعضای هیأت علمی دانشگاهها و مؤسسات آموزش عالی کشور انتخاب شده بودند، موفق شدیم به ۷۸ نفر دسترسی پیدا کنیم. بقیه یا خارج

از کشور بودند یا نشانی آنها نادرست بود، و از همالان آنها ۶۹ نفر را پیدا کردیم. بدین ترتیب پوشش نمونه گیری ما تقریباً ۷۷ درصد بوده است، که در پژوهشهای اجتماعی در حد قابل قبول است. اما چون پاسخگویان به کلیه سؤالات پاسخ نداده بودند، در مورد متغیرهای مختلف نمونه‌ها متفاوت خواهند بود. در زیر، خلاصه‌ای از وضع کل جامعه تحت بررسی را ارائه می‌کنیم:

از بین ۱۳۷ نفر نمونه مورد بررسی که به سؤالات پاسخ داده بودند، ۶۸ نفر دانشگاهی و بقیه غیر دانشگاهی بودند. از این تعداد ۱۶ نفر (۱۱/۷ درصد) زن و ۱۲۱ نفر (۸۸/۳ درصد) مرد بودند.

جدول ۲-۱. آماره‌های توصیفی و مقایسه متغیرهای مستقل در دو جامعه دانشگاهی و غیردانشگاهی

نتیجه	P	غیردانشگاهی			دانشگاهی			متغیر
		S.E	\bar{p} یا \bar{x}	n	S.E.	\bar{p} یا \bar{x}	n	
تفاوت ندارد	۰/۲۷۴۳	۰/۰۴۴	۰/۸۷	۶۹	۰/۰۳۶۴	۰/۹۰	۶۸	درصد مردان
تفاوت ندارد	۰/۰۱۹۲	۰/۰۴۵۲	۰/۸۳	۶۹	۰/۰۲۸۸	۰/۹۴	۶۸	درصد متأهلان
تفاوت ندارد	۰/۳۷۴۹	۱/۷۰	۱/۷۷	۶۹	۱/-	۲/-	۶۵	تعداد فرزندان
تفاوت ندارد	۰/۳۹۳۶	۱/۸۱	۲/۳۱	۶۷	۱/۴۱	۲/۷۵	۶۷	تعداد تحت تکفل
تفاوت ندارد	۰/۱۱۹۰	۸/۵۴	۴۰/۸	۷۶	۸/۵۳	۴۲/۵	۶۷	سن
تفاوت ندارد	۰/۰۰۰۱	۰/۰۵۶۰	۰/۳۱	۶۸	۰/۵۶۸	۰/۶۸	۶۸	درصد دکترا
تفاوت ندارد	۰/۰۰۰۱	۰/۰۵۸۳	۰/۳۶	۶۹	۰/۰۵۳۹	۰/۷۲	۶۸	درصد فارغ التحصیلان
تفاوت ندارد	۰/۱۴۸۰	۱/۲۴	۱/۵	۶۸	۱/۱۸	۱/۵۸	۶۷	مشاغل اضافی
تفاوت ندارد	۰/۰۳۱۴	۰/۰۵۴۹	۰/۷۱	۶۸	۰/۰۴۶۹	۰/۸۴	۶۱	درصد استخدام رسمی
تفاوت ندارد	۰/۰۲۹۴	۹/۳	۱۵/۴	۷۰	۸/۸	۱۲/۵	۷۰	سابقه کار
تفاوت ندارد	۰/۰۰۱	۰/۰۵۸۸	۰/۶۳	۶۷	۰/۵۹۷	۰/۴۱	۶۸	داشتن مسکن
تفاوت ندارد	۰/۱۴۰۱	۸۲/۵	۱۶۳/۱	۷۵	۱۰۸/۲	۱۸/۶	۶۵	زیربنای مسکن
تفاوت ندارد	۰/۰۹۱۸	۰/۰۵	۰/۷۷	۶۹	۰/۱۴	۰/۵۷	۶۸	درصد ساکن تهران

این نتایج، ترکیب جنسی کل جامعه مورد بررسی را نشان می‌دهد. ترکیب جنسی افراد نمونه در زیر جامعه دانشگاهیان عبارت است از ۷ زن (۱۰ درصد) و ۶۱ مرد (۹۰ درصد) در حالی که در بین غیر دانشگاهیان ۹ زن (۱۳ درصد) و ۶۰ مرد (۸۷ درصد) است. در بین غیر دانشگاهیان نسبت زنان بزرگتر از نسبت آنها در بین دانشگاهیان است ولی تفاوت معنی داری نیست.

جدولهای ۱-۲ و ۲-۲ مقایسه دانشگاهیان و غیر دانشگاهیان را از لحاظ متغیرهای تحت بررسی نشان می‌دهد. یادآور می‌شویم که این مقایسه‌ها تنها از حیث یک متغیر است، یعنی تأثیر سایر متغیرها بر متغیر مورد مقایسه نادیده گرفته شده است. مقایسه‌هایی را که در آنها اثر متغیر مورد مقایسه حذف شده است در بخشهای بعد ارائه خواهیم کرد. در جدولهای زیر n نشانه حجم نمونه، \bar{X} و \bar{Y} میانگین نمونه، P نسبت یک صفت در نمونه، SE خطای معیار آماره است، و P سطح معنی داری مقایسه‌ها را نشان می‌دهد.

جدول ۲-۲. خلاصه آماره‌های توصیفی و مقایسه مقادیر تبدیل یافته متغیرهای وابسته در دو زیر جامعه دانشگاهی و غیر دانشگاهی

p-value	t	غیر دانشگاهی			دانشگاهی			متغیر
		S.E.	\bar{y}	n	S.E.	\bar{y}	y	
۰/۰۴	۲/۰۳۴	۱۱۹۴۶۹	۷۸۴۸۰	۴۲	۲۲۸۹۹۶	۱۴۰۹۲	۵۲	تحصیل فرزندان
۰/۷۰	۰/۰۹۹۸	۳۰۷۱۲	۱۹۹۷۰	۵۶	۲۸۰۷۱	۱۹۴۶۸	۵۸	دارو و درمان
۰/۲۵	۱/۰۰۸	۱۴۵۹۱۶	۲۹۲۰۱۰	۶۷	۱۹۲۰۷۱	۳۲۱۴۱۲	۶۱	تغذیه
۰/۲۰	۱/۲۰۷	۱۴۰۸۱۹	۸۵۶۹۸	۶۵	۵۳۰۰۴	۶۵۲۳۳	۶۲	پوشاک
۰/۵۰	۰/۲۷۶	۳۴۶۰۹	۲۳۲۶۶	۴۷	۳۴۵۵۸	۲۱۶۳۳	۵۳	فرهنگی
۰/۲۵	۰/۱۰۷	۶۲۶۲	۶۷۹۳۹	۵۴	۱۱۶۶۱۲	۶۶۰۶۴	۵۱	تفریح
۰/۱۵	۰/۱۲۶	۲۴۴۸۱۸	۱۱۲۳۲۳	۵۶	۲۱۸۸۹۹	۱۰۷۳۲۹	۵۲	مسافرت
۰/۲۰	۰/۸۵۲	۵۴۳۷۰۸	۶۷۳۹۳۷	۶۸	۴۶۳۵۴۰	۷۴۷۴۲۰	۶۷	اقساط
۰/۴۰	۰/۳۳۹	۷۲۵۸۰۱	۱۸۰۸۳۱	۴۶	۱۵۵۲۳۷	۱۵۰۵۴۴	۴۶	مجموع

وقتی ارقام هزینه‌ها را به طور جداگانه برای دو جامعه مقایسه می‌کنیم، در اکثر ارقام تفاوتی بین میانگینهای دو جامعه دیده نمی‌شود. این امر ناشی از ناهمگنی زیاد افراد دو جامعه است و دلیل آن وجود واریانسهای بسیار بزرگ برای هریک از ارقام هزینه است. برای مقایسه دقیقتر باید نقش متغیرهای مستقل را در ایجاد پراکندگی و ناهمگنی بین افراد هر جامعه در نظر بگیریم. در این خصوص سؤال این است که با توجه به مشخصات متفاوت افراد موجود در هر نمونه و با توجه به اینکه الگوی مصرف خانوارها با هم تفاوت دارند، آیا به طور متوسط هزینه‌های دو نفر از این دو زیر جامعه که دارای شرایط مشابه و ویژگیهای همانند هستند و تنها از لحاظ شغل اصلی با هم تفاوت دارند، یکسان است یا با هم تفاوت دارند؟ اگر تفاوتی وجود دارد به دلیل کدام ویژگی است؟ در زیر به پاسخ این سؤالها می‌پردازیم:

۳- مدل‌های آماری برای مقایسه دو زیر جامعه

در این بخش می‌خواهیم با استفاده از مدل‌های آماری که در بخش ۱-۸-۱ آورده‌ایم به مقایسه ارقام هزینه خانوارهای دانشگاهیان و غیر دانشگاهیان پردازیم. در این مقایسه‌ها با توجه به مدل‌های مورد استفاده در واقع افرادی را که از هر لحاظ شبیه هم هستند و تنها از نظر شغل اصلی با هم تفاوت دارند، از بابت هزینه‌هایی که تقبل کرده‌اند با هم مقایسه می‌کنیم. در این مدل‌ها به کمک روشهای آماری، میزان تأثیر هر یک از متغیرهای مستقل به جز متغیر شغل اصلی را بر هریک از ارقام هزینه (و مجموع هزینه‌ها) برآورد می‌کنیم و کنار می‌گذاریم، و سپس خالص هزینه‌ها را که مربوط به دو گروه شغل اصلی هستند، با هم مقایسه می‌کنیم. البته این عمل به طور ضمنی و با استفاده از فرمولهای آماری صورت خواهد گرفت و نه به صورت صریح. بنابراین لازم است که در هر مدل فرضیه‌های مورد آزمون را به نحوی با پارامترهای هر مدل معرفی کنیم. برقراری این تناظر بین فرضیه‌ها و پارامترهای مدل آماری امکان می‌دهد که با انجام آزمون آماری بر روی پارامترها و مخصوصاً آزمون برابری آنها در دو گروه شغلی (پس از حذف اثر سایر متغیرها بر روی هزینه‌ها) به هدف خود که مقایسه دو زیر جامعه، فارغ از تداخل ناهماهنگیهای مربوط به مشخصات فردی افراد جامعه است، برسیم. در زیر مدل‌های گوناگون و نتایج محاسبات آماری مربوط به برازش آنها به داده‌های تحقیق را می‌آوریم.

۳-۱. برازش مدل رگرسیون خطی چند متغیره

مدل رگرسیون خطی چند متغیره، مدلی است که وابستگی مقدار متوسط متغیر پاسخ را به چند متغیر مستقل نشان می‌دهد. متغیرهای مستقل را متغیرهای پیشگو نیز می‌گویند، زیرا با در دست داشتن مدل رگرسیون چند متغیره و دانستن مقادیر متغیرهای پیشگو، می‌توان مقدار

متوسط یا انفرادی متغیر پاسخ را با درجه معینی از اطمینان پیشگویی کرد. این مدل که شایعترین مدل برای تعیین میزان وابستگی یک متغیر به متغیرهای دیگر است، به صورت:

$$y = b_0 + b_1x_1 + \dots + b_px_p + e$$

بیان می‌شود که در آن ضرایب b میزان تأثیر هر متغیر مستقل را، به فرض ثابت ماندن متغیرهای مستقل دیگر، بر y نشان می‌دهد. به کمک تحلیل‌های آماری باید از بین کلیه متغیرهای مستقلی که ممکن است بر y مؤثر باشند، کوچکترین مجموعه ممکن را برگزینیم. در آن توانسته‌ایم متغیرهای مؤثر را جدا و میزان تأثیر هر یک را بر y تعیین کنیم. به منظور آزمون فرضیه‌هایی از قبیل درستی مدل، درستی مفروضاتی که مدل بر اساس آنها تعیین می‌شود، شدت تأثیر هر یک از متغیرهای مستقل بر y ، دسته‌ای از آزمونهای آماری ابداع شده‌اند که به کمک آنها می‌توان پاسخ سؤالیهای زیر را داد. آیا مدل انتخاب شده کلاً مدل درستی است؟ به فرض درست بودن مدل، آیا مفروضات زیر بنایی از قبیل نرمال بودن داده‌ها، استقلال مشاهدات،... برقرارند؟ به فرض درست بودن مدل و مفروضات آن، کدام یک از متغیرهای مستقل، با در نظر گرفتن اثر سایر متغیرها، بر متغیر پاسخ اثر دارد؟ جواب این سؤالیها را نیز همراه با بهترین مدل انتخاب شده در هر مورد ذکر کرده‌ایم. بدین معنی که ابتدا نرمال بودن داده‌ها را آزمون کرده‌ایم و به دلیل نرمال نبودن آنها به تبدیل داده‌ها پرداخته‌ایم. یعنی به جای متغیر خام y از تبدیل یافته آن y که نرمال است استفاده کرده‌ایم. در هر مورد پس از برازش مدل و یافتن مؤثرترین متغیرهای مستقل، درستی مدل را با انجام آنالیز واریانس و محاسبه ضریب تعیین R^2 نشان داده‌ایم. همچنین با آزمون "دوربین-واتسن" درستی فرض استقلال داده‌ها را آزموده‌ایم. ملاک R^2 نشان می‌دهد که متغیرهای مستقل موجود در مدل رگرسیونی چند درصد از تغییرات موجود در متغیر پاسخ را می‌توانند تبیین کنند. بدیهی است که هر قدر این درصد بیشتر و نزدیک به ۱۰۰ باشد، مدل توانایی بیشتری برای پیشگویی دارد. مدلهای به دست آمده دارای ضریب تعیین بیش از ۸۰ درصد و در اغلب موارد حدود ۹۰ درصد هستند. بنابراین قدرت پیشگویی آنها بسیار بالاست (سیرل ۱۹۷۱).

از آنجا که در بیشتر روشهای آماری، نرمال بودن داده‌ها از اهمیت بسزائی برخوردار است و برقرار نبودن این فرض نتایج حاصل از کاربست آن روش را بی اعتبار می‌کند، در این بررسی نیز بی‌نیاز از آزمون این فرض در مورد داده‌های به دست آمده نبودیم.

روشهای گوناگونی برای بررسی نرمال بودن داده‌ها وجود دارد. با توجه به توان آزمون "کولموگروف" از این روش برای آزمون نرمال بودن داده‌ها استفاده شد. نتایج حاصل نشان

داد که در همه داده‌های موجود فرض نرمال بودن رد می‌شود. بنابراین از روش "باکس-کاکس" برای نرمال کردن داده‌ها استفاده شد.

این روش به شرح زیر خلاصه می‌شود: اگر داده‌های اصلی با X نشان داده شوند و نرمال نباشند، متغیر جدید Y که با تبدیل زیر محاسبه شده باشد، دارای توزیع نرمال می‌شود.

$$Y = Y(x) = \begin{cases} (x^b - 1)/b, & b \neq 0 \\ \log x, & b = 0 \end{cases} \quad (1-1-3)$$

که در آن:

$$\lim_{b \rightarrow 0} (x^b - 1)/b = \log x$$

مقدار b به کمک کامپیوتر و با روش سعی و خطا دست آورده می‌شود.

۳-۱-۲. چگونگی برازش مدل

رگرسیون هر یک از متغیرهای نرمال شده Y را بر حسب متغیرهای مستقل مورد بررسی قرار می‌دهیم. پس از برازاندن مدل (با استفاده از روش گام به گام پسرو) در نهایت تعدادی متغیر در مدل باقی می‌ماند. برای هر یک از متغیرها برآورد ضرایب، خطای معیار، آماره t و سطح معنی داری محاسبه می‌شود.

۱- مدل رگرسیونی هزینه تحصیل فرزندان بر حسب متغیرهای مستقل به صورت (۳-۱-۲) محاسبه شده است. همه ضرایب این مدل در سطح کمتر از ۲ درصد معنی دارند و ضریب تعیین آن، یعنی $R^2 = 0.7977$ نشان می‌دهد که حدود ۸۰ درصد تغییرات داده‌ها را با این مدل می‌توان تبیین کرد. پس مدل به داده‌ها به خوبی می‌بrazد.

$$Y_1 = -1/6268 \times 10^4 X_2 + 8473/35 X_5 + 23674/308 X_{10}$$

$$+ 1069/812 X_{12} + 40460/417 X_{18} + 43/894 X_{17}$$

$$- 1/279 \times 10^4 X_1 X_{18} - 723/307 X_7 X_{12} - 1/813 \times 10^4 X_{10} X_{18}$$

۲- مدل رگرسیونی هزینه دارو و درمان بر حسب متغیرهای مستقل در (۳-۱-۳) ارائه شده است. همه ضرایب در سطح کمتر از ۳ درصد معنی دارند و $R^2 = 0.8936$.

$$Y_2 = 996/27 X_6 + 22566/003 X_7 - 1147 \times 10^4 X_{16} - 179/64 X_{17} \quad (3-1-3)$$

$$+ 3232/078 X_1 X_{16} - 558/848 X_6 X_7 - 278/304 X_6 X_{18}$$

$$- 2231/374 X_8 X_{11} + 2839/136 X_8 X_{18} + 1807/878 X_9 X_{15}$$

$$+ 20/892 X_{14} X_{17} + 81/601 X_{16} X_{17}$$

۳- مدل رگرسیونی هزینه تغذیه بر حسب متغیرهای مستقل در (۳-۱-۴) نشان داده شده است. ضرایب این مدل همگی در سطحی کمتر از ۲ درصد معنی دارند و $R^2 = 0.7926$.

$$Y_3 = 2/225 \times 10^5 X_7 + 540930/0.48 X_{10} + 2/987 \times 10^5 X_{11} \quad (4-1-3)$$

$$-4/0.17 \times 10^5 X_1 \cdot X_{14}^4 + 266/566 X_1 \cdot X_{17} - 1/449 \times 10^5 \cdot X_7^5 X_{11} \\ - 2/375 \times 10^4 X_8 \cdot X_{11} + 19850/59 X_8 \cdot X_{14} + 49138/907 X_9 \cdot X_{14} \\ - 3/533 \times 10^4 \cdot X_{18} X_{10}$$

۴- مدل رگرسیونی هزینه پوشاک بر حسب متغیرهای مستقل به شرح (۳-۱-۵) است که در آن همه ضرایب در سطح کمتر از ۲ درصد معنی دارند و $R^2 = 0.9929$.

(۳-۱-۵)

$$Y_4 = -4/128 \times 10^4 X_1 - 1/682 \times 10^5 X_2 - 1/579 \times 10^4 X_6 + 4/464 \times 10^5 X_7 + 38925/0.07 X_{10} \\ + 48600/661 X_{11} + 5753/79 X_{12} + 404/44 X_1 \cdot X_{17} + 6489/184 X_2 \cdot X_6 - 2/34 \times 10^5 X_2 \cdot X_7 \\ + 12891/0.23 X_2 \cdot X_8 + 4087/390 X_4 \cdot X_{12} - 4/0.72 \times 10^4 X_4 \cdot X_{16} - 3459/20 X_5 \cdot X_{12} \\ + 61859/954 X_5 \cdot X_{16} - 2/149 \times 10^4 \times X_5 \cdot X_{18} - 16/825 X_6 \cdot X_{17} + 3039/47 X_6 \cdot X_{18} \\ - 3857/508 X_8 \cdot X_{13} + 5402/135 X_8 \cdot X_{14} - 151/0.67 X_8 \cdot X_{17} + 451/167 X_9 \cdot X_{17} \\ - 4/559 \times 10^4 X_9 \cdot X_{18} - 3/10.1 \times 10^4 X_{10} \cdot X_{18} - 9899 \times 10^4 X_{11} \cdot X_{16} + 64184/82 X_{11} \cdot X_{18} \\ - 3500/74 X_{12} \cdot X_{18}$$

۵- مدل رگرسیونی هزینه خدمات فرهنگی بر حسب متغیرهای مستقل به قرار (۳-۱-۶) به دست آمده است که در آن همگی ضرایب در سطحی کمتر از ۲ درصد معنی دارند و $R^2 = 0.8507$.

(۳-۱-۶)

$$Y_5 = 24638/433 X_1 + 35136/0.17 X_2 + 54136/565 X_4 - 3/667 \times 10^4 X_5 - 1861/0.74 X_6 \\ + 10783/155 X_7 + 1395/270 X_{12} - 4531/57 X_{13} - 6848/623 X_{14} - 3/3035 \times 10^4 X_{16} \\ - 7301/589 X_1 \cdot X_7 - 2576/904 X_1 \cdot X_{10} - 6797/794 X_1 \cdot X_{11} + 21/303 X_1 \cdot X_{17} \\ + 21/303 X_1 \cdot X_{17} - 3893/14 X_1 \cdot X_{18} - 1/926 \times 10^4 X_2 \cdot X_4 + 4214/764 X_2 \cdot X_5 \\ - 723/0.39 X_4 \cdot X_6 + 10945/382 X_4 \cdot X_{16} + 810/0.74 X_5 \cdot X_6 - 442889 X_5 \cdot X_{12} \\ + 340/301 X_6 \cdot X_9 + 651/637 X_6 \cdot X_{11} - 1/224 \times 10^4 X_9 \cdot X_{11} + 3446/415 X_{10} \cdot X_{16} \\ + 2100/282 X_{13} \cdot X_{14}$$

۶- مدل رگرسیونی هزینه تفریح بر حسب متغیرهای مستقل به صورت (۳-۱-۷) حاصل شده است. در این مدل همه ضرایب در سطحی کمتر از ۸ درصد معنی دارند، $R^2 = 0.90439$ که حاکی از برازش فوق العاده مدل به داده‌هاست.

$$Y_6 = 55988/884 X_1 - 7/923 \times 10^4 X_5 + 49792/63 X_7 - 3/43 \times 10^4 X_1 \cdot X_7 \\ + 43310/308 X_1 \cdot X_{16} - 4/825 \times 10^4 X_1 \cdot X_{18} + 24556/244 X_2 \cdot X_5$$

$$1028/48X_5X_6 - 593/455X_5X_{12} - 11844X_6X_{17} - 1/098 \times 10^4 X_8X_{16}$$

$$+ 12277/159X_8X_{18} - 1191/784X_9X_{12} + 117/405X_9X_{17}$$

$$+ 2085/867X_{12}X_{18} - 1/935 \times 10^4 X_{16}X_{18} + 203/167X_{17}X_{18}$$

۷- مدل رگرسیونی هزینه مسافرت بر حسب متغیرهای مستقل در (۳-۱-۸) داده شده است که در آن همه ضرایب در سطحی کمتر از ۵ درصد معنی دار شده‌اند و $R^2 = 0.9247$. باز هم برازش بسیار خوب است.

(۳-۱-۸)

$$Y_7 = -2/985 \times 10^4 X_5 + 47699/912X_8 - 5/065 \times 10^4 X_{14} + 1060/031X_{17}$$

$$+ 37978/01X_1X_2 + 305/53X_1X_{17} - 7/673 \times 10^4 X_1X_{18} + 10679/245X_2X_9$$

$$+ 12483/223X_4X_{16} - 913/349X_5X_{12} + 10409/762X_5X_{13} - 42/453X_6X_{17}$$

$$+ 3902/363X_6X_{18} - 2/878 \times 10^4 X_8X_{18} + 18/397X_{12}X_{17} - 2/217 \times 10^4 X_{13}X_{18}$$

$$+ 42537/34X_{14}X_{18}$$

۸- مدل رگرسیونی هزینه اقساط ماهانه بر حسب متغیرهای مستقل در (۳-۱-۹) آمده است در این مدل نیز همه ضرایب در سطحی کمتر از ۵ درصد معنی دار شده‌اند و $R^2 = 0.8510$ پس برازش خیلی خوب است.

$$Y_8 = 7893/90X_6 - 6619/303X_{12} - 6/153 \times 10^4 X_{14} - 439/7493X_{17}$$

$$+ 69791/791X_{18} - 4996/206X_1X_6 + 565253/73X_1X_{14} + 81274/664X_1X_{16}$$

$$+ 10333/284X_4X_{14} - 6732/885X_5X_7 + 2386/153X_8X_{12} - 3/285 \times 10^4 X_8X_{16}$$

$$+ 86/603X_{13}X_{17} - 192/32X_{14}X_{17} + 174/911X_{15}X_{17} - 2/291X_{15}X_{18}$$

۲-۳. برازش مدل رگرسیونی لوجستیک

مدل رگرسیونی لوجستیک زمانی به کار می‌رود که متغیر پاسخ دو یا چند حالتی باشد. در این مدل قیود چندانی برای متغیر پاسخ از قبیل نرمال بودن، همواریانسی،... که در رگرسیون خطی چند متغیری دیدیم وجود ندارد. از این رو اطمینان خاطر بیشتر در انتخاب مدل درست، وجود دارد. چون میزان هزینه‌ها متغیرهایی پیوسته و در دامنه‌ای گسترده‌اند، هنگام اندازه‌گیری آنها فقط به حافظه پاسخگویان اعتماد شده است. شاید محافظه کارانه‌تر آن باشد که هزینه‌های انجام شده را به دو دسته "کم" و "زیاد" تقسیم کنیم. یعنی، بپذیریم که افراد نمونه با ارقامی که بیان کرده‌اند، فقط خواسته‌اند بگویند که جزو کدام دسته‌اند، نه آنکه دقیقاً میزان هزینه خود را در هر مورد اعلام کرده باشند. زیرا چنین تقاضایی بی‌مراجعه به سوابق یا حساب و کتاب مخارج خانواده‌ها امری تقریباً محال است.

با توجه به استدلال فوق، مجموع هزینه‌ها را به دو دسته تقسیم کرده‌ایم. دسته بالاتر از میانه را "زیاد" و با کُد ۱ و پایینتر از میانه را با کُد صفر مشخص کرده‌ایم. حال سؤال این است: کدام متغیرهای مستقل باعث می‌شوند که فردی دارای هزینه بالاتر یا پایینتر از میانه شود؟ آیا متغیر شغل اصلی از این گونه متغیرهاست؟ آیا پس از در نظر گرفتن و به حساب آوردن اثر سایر متغیرها، اثر شغل اصلی به قوت خود باقی می‌ماند و جزو متغیرهای فارق محسوب می‌شود؟ برای پاسخگویی به این سؤالها با استفاده از مدل رگرسیون لجستیک، لازم است که با اساس منطقی آن آشنا شویم. اگر احتمال قرار گرفتن در دسته پر هزینه را به عنوان تابعی از متغیرهای مستقل به صورت $\rho(x)$ در نظر بگیریم، روشن است که $\rho(x)$ باید عددی بین صفر و یک باشد. از جمله توابع مورد استفاده برای این منظور که خواص جالبی دارد، تابع لجستیک است:

$$\rho(x) = [e^{b_0 + b_1x_1 + \dots + b_px_p} / (1 + e^{b_0 + b_1x_1 + \dots + b_px_p})]$$

برای این صورت خاص از تابع $\rho(x)$ ، اگر احتمال تعلق به دسته ۱ را به احتمال تعلق به

دسته ۰ تقسیم کنیم، نسبت بختها به دست می‌آید:

$$\phi(x) = \frac{\rho(x)}{1 - \rho(x)} = e^{b_0 + b_1x_1 + \dots + b_px_p}$$

مقدار این نسبت در صورتی که بزرگتر از یک باشد، حاکی از آن است که شخصی که اندازه متغیرهای مستقل برای او برابر (X_1, X_2, \dots, X_p) است، بیشتر احتمال دارد در دسته با هزینه "زیاد" قرار گیرد تا در دسته "کم" هزینه. بر عکس اگر نسبت بختها کوچکتر از یک باشد، دارا بودن مقدار (X_1, X_2, \dots, X_p) سبب می‌شود که این فرد، با احتمال بیشتری در دسته "کم" هزینه قرار گیرد. از اینجا نقش x ها در تعیین میزان هزینه روشن می‌شود. اما برای سادگی اعمال ریاضی، به جای نسبت بختها با لگاریتم طبیعی آن یعنی،

$$g(x) = \log \phi(x) = \log \left(\frac{\rho(x)}{1 - \rho(x)} \right) = b_0 + b_1x_1 + \dots + b_px_p$$

کار می‌کنیم. تابع $g(X)$ را تابع لجیت می‌گویند که نظیر رگرسیون خطی است و وابستگی

y را به x ها از طریق نسبت بختها نشان می‌دهد.

دوباره این سؤالها مطرح می‌شود: آیا اصولاً متغیرهای مستقل در تعیین نسبت بختها نقشی دارند یا خیر؟ کدام یک از متغیرهای مستقل در این نقش مؤثرند و میزان تأثیر آنها چقدر است؟ جواب آن است که پس از برآوردن مدل و آزمونهای لازم آن تعداد از متغیرهای مستقل که در مدل باقی بمانند و حذف نشوند، تأثیر متغیرهای مستقل را بر نسبت بختها و به عبارت دیگر در بالا

یا پایین بودن هزینه نشان می‌دهند. ضرایب متغیرها میزان شدت تأثیر آنها را نشان می‌دهند. در زیر نتایج برازش مدل را به داده‌ها ذکر می‌کنیم. در این مورد متغیر پاسخ مجموع هزینه‌های خانوار است.

در برازش مدل لوجستیک ابتدا کلیه متغیرهای مستقل را در مدل گنجانده‌ایم. سپس تعیین کرده‌ایم که با احتمال خطای نوع اول ۲۰ درصد کدام متغیرها را می‌توان از مدل حذف کرد. این کار را به صورت گام به گام پسرو انجام داده‌ایم. در هر گام پس از حذف یک متغیر مستقل، دوباره به گنجاندن متغیرهای مستقلی که در گامهای قبلی از مدل حذف شده‌اند، اقدام کرده‌ایم. برای گنجاندن متغیری در مدل از احتمال خطای نوع اول ۱۵ درصد استفاده شده است. مقادیر فوق، مقادیری هستند که برای مرحله مقدماتی مدل سازی توصیه شده‌اند (هاسمر و لیمی شا ۱۹۸۹). مدل نهایی برای رگرسیون لوجستیک چند متغیره به صورت معادله (۱-۲-۳) حاصل شده است. برای رسیدن به این مدل ۲۰ گام، یعنی ۲۰ نوع مدل مختلف برآزنده و آزمون شده‌اند. مشخصات مدل در جدولهای ۲-۲-۳ و معادله ۱-۲-۳ درج شده‌اند. در جدول ۲-۲-۳، ملاحظه می‌کنیم که مدل موصوف می‌تواند ۸۴ درصد از موارد را صحیح پیش‌بینی کند. یعنی با در دست داشتن متغیرهای مستقل افراد جامعه را در دسته‌های کم هزینه و پر هزینه دسته‌بندی کند و حاصل دسته‌بندی از روی مدل با واقعیت مشهود در نمونه ۸۴ درصد تطبیق دارد. ضمناً ضرایب اثرهای اصلی متغیرهای مستقل همگی در سطح ۵ درصد معنی دارند. در واقع جدول ۲-۲-۳ همان مطلبی را بیان می‌کند که ضریب تعیین در مدل رگرسیون خطی چند متغیره، بیان می‌کرد. بنابراین مدل ما مدل موفقی است.

جدول ۲-۲-۳ جدول همخوانی مقادیر مشاهده شده با مقادیر پیش‌بینی شده

مقدار پیش‌بینی شده	مقادیر مشاهده شده	درصد همخوانی
۰	۱	
۰	۳۷	۸۰/۴۳٪
۱	۷	۸۷/۰۴٪
		۸۴٪

$$\begin{aligned}
 g(x) = & 34/9249 + 3/1844x_1(1) - 0/1662x_{10}(1) + 2/2166x_{10}(2) + (1-2-3) \\
 & 0/1898x_{10}(3) - 0/1539x_{10}(4) + 5/1326x_{15}(1) + 16984x_{15}(2) \\
 & - 6/9718x_{15}(3) + 0/1899x_{17} - 4/9161x_5 + 2/6576x_9(1) \\
 & - 5/4644(x_{17})^2 - 9/3488(x_5)^2 - 1/0799x_1(1)x_4 + \\
 & 0/6452x_{18}(1)x_4 - 2/8742x_9(1)(x_5)^2 - 1/1018x_{16}(1)x_{18}(1) \\
 & - 1/2835x_1(1)x_1
 \end{aligned}$$

۴- نتیجه گیری و پیشنهاد

۴-۱. مقدمه

در این بخش می‌خواهیم نتایجی را که از کاربست روشهای آماری به دست آمده‌اند بیان و تفسیر کنیم، به طوری که در نهایت بتوانیم فرضیه‌های مذکور در بخش ۱-۴ را قبول یا رد کنیم و به این ترتیب یافته این تحقیق را به خوانندگان عرضه کنیم.

برای نیل به مقصود فوق لازم است که ابتدا فرضیه‌های مذکور درباره زیر جامعه‌ها را به فرضیهایی درباره پارامترهای مدلهای مورد استفاده برگردانیم. چون هدف اصلی ما مقایسه دانشگاهیان و همالان آنها از حیث رفاه و وضع اقتصادی است مهمترین متغیر مستقل، همان متغیر X_1 یعنی شغل اصلی است. سایر متغیرهای مستقل نقش کنترل کننده دارند و غرض از اندازه‌گیری آنها، به حساب آوردن سهم آنها در وضع رفاهی خانواده‌هاست، به نحوی که مقایسه دانشگاهیان و همالان آنها در شرایط یکسانی از تحصیلات، سابقه خدمت، بعد خانوار، مکان جغرافیایی و ... صورت گیرد. بنابراین پارامترهایی که اثر متغیر مستقل شغل اصلی را بر اقلام هزینه یا مجموع هزینه‌ها نشان می‌دهند، پارامترهای مورد نظراند.

در مدل رگرسیون چند متغیری برای هر یک از اقلام هزینه، تناظر زیر را بین فرضیه‌های مذکور در بخش ۱-۴ و فرض آماری منظور برای ضریب شغل اصلی مشاهده می‌کنیم:

در صورتی که ضریب شغل اصلی معنی دار باشد و صفر نباشد به معنی آن است که دو جامعه دانشگاهی و غیر دانشگاهی در هر قلم از هزینه‌ها با هم تفاوت دارند و این تفاوت به واسطه شغل اصلی آنهاست. در محاسبات، دانشگاهیان با $X_1=1$ و غیر دانشگاهیان با $X_1=2$ مشخص شده‌اند. در صورتی که متغیر X_1 در هر یک از معادله‌ها تنها ظاهر شود، نشانه اثر مستقیم آن بر متغیر پاسخ است. در این صورت، اگر ضریب X_1 مثبت باشد، نشانه آن است که گروه غیر دانشگاهیان در آن قلم از هزینه‌ها، مبلغ بیشتری را هزینه می‌کنند، بر عکس، اگر ضریب X_1 منفی

باشد به معنی آن است که دانشگاهیان برای آن قلم بیشتر هزینه می‌کنند. هنگامی که X_1 در ترکیب با سایر متغیرها در معادله ظاهر می‌شود، نشانه تأثیر غیر مستقیم آن بر پاسخ است، در این صورت باید ترکیبهای متفاوت آن را در نظر گرفت و مقایسه را انجام داد.

پس از تعیین اثر شغل اصلی در ارقام هزینه، می‌توان گفت که کدام گروه در وضع رفاهی بهتری است. زیرا کسی می‌تواند از عهده این هزینه‌ها برآید که درآمد بیشتری داشته باشد. اکنون که تناظر بین فرضیه‌های تحقیق را با پارامترهای مدلها روشن کرده‌ایم، به بررسی مدل‌های به دست آمده می‌پردازیم.

۲-۴. تفسیر مدل‌های رگرسیونی

۱- از مدل (۳-۱-۲) چنین بر می‌آید که شغل اصلی بر هزینه تحصیل رسمی فرزندان مستقیماً اثر ندارد. بنابراین از لحاظ این قلم هزینه دو جامعه دانشگاهی و همالان آنها با هم تفاوت ندارند. این موضوع به نظر می‌رسد بیشتر به این دلیل باشد که در کشور ما تحصیلات رسمی اکثراً رایگان است. تنها در سالهای اخیر است که مدارس و دانشگاههای غیرانتفاعی (مربوط به بخش خصوصی) دایر شده‌اند و این امر هنوز به صورتی نیست که شاخص تمایز بین این دو جامعه بشود. متغیر شغل اصلی، با متغیرهای زیربنای منزل مسکونی و مکان جغرافیایی سکونت اثر متقابل دارد. یعنی شغل اصلی به طور غیر مستقیم و در تأثیر متقابل با دو متغیر فوق‌الذکر بر هزینه تحصیل اثر دارد. از آنجا که خود متغیر X_1 در معادله به طور مستقیم ظاهر نمی‌شود، نمی‌توان برای اثرهای متقابل آن اهمیت چندانی قائل شد. اما اگر بخواهیم موشکافی کنیم چنین نتیجه خواهیم گرفت:

تفاوت هزینه تحصیل برای فرزندان فرد دانشگاهی ساکن تهران، که از منزلی با زیر بنای متوسط استفاده می‌کند، و فردی غیر دانشگاهی، که ساکن تهران است و از منزلی با زیر بنای متوسط استفاده می‌کند، تقریباً برابر است با ۶۴۱۷ واحد در ماه. همین تفاوت برای دانشگاهی و غیر دانشگاهی ساکن شهرستانها برابر است با ۱۹۲۱۴ واحد در ماه. بنابراین هزینه تحصیل فرزندان برای دانشگاهیان بالاتر است. دلیل ذکر واحد برای تفاوتها آن است که هزینه‌ها تبدیل شده‌اند و واحد آنها ریال نیست. اگر بخواهیم تفاوت را به ریال به دست آوریم باید عکس تبدیل (۳-۱-۱) را به کار ببریم. اما چون قصد مقایسه است چنین کاری لزومی ندارد.

۲- مدل رگرسیونی هزینه دارو و درمان بر حسب متغیرهای مستقل در (۳-۱-۳) ارائه شده است. بر طبق این مدل باز هم شغل اصلی اثر مستقیم بر هزینه دارو و درمان خانواده ندارد.

متغیر شغل اصلی با متغیر نوع منزل مسکونی اثر متقابل دارد. پس در بین کسانی که در آپارتمان زندگی می‌کنند، غیر دانشگاهیان بیشتر هزینه دارو و درمان می‌پردازند تا دانشگاهیان و تفاوت برابر است با ۵۳۲۳/۱ واحد در ماه و در بین ویلا نشینان نیز روال به همین ترتیب است منتها تفاوت بیشتر و برابر ۶۴۶۴/۲ واحد در ماه است. پس، هزینه دارو و درمان خانواده غیر دانشگاهیان بیشتر از دانشگاهیان است.

۳- مدل مربوط به وابستگی هزینه تغذیه به متغیرهای مستقل در (۳-۱-۴) عرضه شده است. بر پایه این مدل، شغل اصلی مستقیماً در هزینه تغذیه اثر ندارد. بلکه در تأثیر متقابل با متغیر شغل همسر و زیر بنای منزل مسکونی است. چون متغیر شغل همسر ۴ مقدار دارد، تعداد ترکیبها و حالت‌های گوناگون زیاد می‌شوند. از این رو دو وضع انتهایی را در نظر می‌گیریم و افراد دانشگاهی و غیر دانشگاهی را در این دو وضع با هم مقایسه می‌کنیم. حالت‌های بینابین را می‌توان از روی این دو حالت حدس زد.

دانشگاهیان و غیر دانشگاهیانی را در نظر بگیریم که شغل همسر آنها خانه‌دار یا عضو هیأت علمی است، در این دو حالت انتهایی، تفاوت هزینه خانوارهای دانشگاهی و غیر دانشگاهی برابر است با ۱۵۱۳/۳ واحد در ماه. در انتهای دیگر مقیاس که شغل همسر عضو هیأت علمی است، این تفاوت برابر است با ۱۲۲۰۲۴/۲ واحد در هر ماه. هر دو عدد نشانگر آنند که هزینه تغذیه دانشگاهیان بالاتر است.

۴- درباره هزینه پوشاک مدل (۳-۱-۵) را به دست آورده‌ایم. از این مدل پیداست که متغیر شغل اصلی مستقیماً در هزینه پوشاک تأثیر دارد و به صورت غیر مستقیم نیز با متغیر زیر بنای منزل مورد استفاده در تأثیر متقابل است. ضریب منفی X_1 دال بر آن است که دانشگاهیان هزینه پوشاک بیشتری دارند ولی بزرگی ضریب اثر متقابل X_1 و X_{17} که صد برابر ضریب X_1 و مثبت است، این تصویر را وارونه می‌کند. تفاوت هزینه‌های پوشاک برای دانشگاهیانی که از متوسط زیر بنای منزل مسکونی زیر جامعه خود استفاده می‌کنند و غیر دانشگاهیانی که از متوسط زیر بنای مسکونی زیر جامعه غیر دانشگاهیان استفاده می‌کنند، برابر است با ۵۸۶۳۴- واحد در ماه. پس غیر دانشگاهیان هزینه پوشاک بیشتری دارند.

۵- در مورد هزینه خدمات فرهنگی، مدل در (۳-۱-۶) درج شده است. در این مدل متغیر شغل اصلی هم به صورت مستقیم و هم در تأثیر متقابل با متغیرهای مستقل آخرین مدرک تحصیلی، مشاغل اضافی، وضع استخدام، زیر بنای منزل مسکونی، و مکان جغرافیایی خدمت،

بر هزینه خدمات فرهنگی اثر دارد. چون حالت‌های ممکن برای اثرهای متقابل بسیار زیاد است، تفسیر نتایج قدری مشکل است. ولی به طور کلی با توجه به ضریب اثر مستقیم X_1 می‌توان گفت که هزینه خدمات فرهنگی غیر دانشگاهیان بیشتر از آن دانشگاهیان است.

۶- برای هزینه تفریح، مدل (۳-۱-۷) حاصل شده است. در این مدل نیز وضع تا حدودی شبیه مورد قبلی است. متغیر شغل اصلی به طور مستقیم بسیار مؤثر است و به طور غیر مستقیم نیز در تأثیر متقابل با آخرین مدرک تحصیلی، نوع منزل مسکونی، و مکان جغرافیایی خدمت است. با توجه به ضرایب اثر اصلی و اثرهای متقابل باز هم می‌توان نتیجه گرفت که به طور کلی هزینه تفریح غیر دانشگاهیان بیشتر است.

۷- مدل رگرسیونی هزینه مسافرت در معادله (۳-۱-۸) ارائه شده است. متغیر شغل اصلی اثر مستقیم بر هزینه مسافرت ندارد و تنها به طور غیر مستقیم از طریق متغیرهای جنسیت، زیربنای منزل مسکونی، و مکان جغرافیایی خدمت بر هزینه مسافرت اثر می‌کند. با توجه به بزرگی ضریب $X_1 * X_2$ ، یعنی اثر متقابل شغل اصلی و جنس بر هزینه مسافرت که حدوداً پنج‌هزار برابر اثر متقابل $X_1 * X_{18}$ است، می‌توان مقایسه را بر پایه دو اثر متقابل $X_1 * X_2$ و $X_1 * X_{17}$ بنا کرد. برای این منظور دانشگاهیان و غیردانشگاهیان مرد را که از متوسط زیربنای مسکونی استفاده می‌کنند بایکدیگر و زنهایی را که در شرایط فوق هستند باهم مقایسه می‌کنیم. تفاوت هزینه مسافرت دانشگاهیان و غیردانشگاهیان مرد، که در خانه‌هایی برابر متوسط زیر جامعه خود زندگی می‌کنند، برابر است با ۱۲۰۱۸۱- واحد در ماه که نشان می‌دهد هزینه مسافرت غیردانشگاهیان مرد خیلی بیشتر از دانشگاهیان مرد است. در مورد زنان این تفاوت برابر می‌شود با ۸۲۲۰۳- واحد در ماه. پس غیردانشگاهیان هزینه مسافرت بیشتری از دانشگاهیان دارند.

۸- مدل رگرسیونی اقساط ماهانه که در معادله (۳-۱-۹) نشان داده شده است، حاکی از آن است که شغل اصلی مستقیماً بر میزان اقساط ماهانه‌ای که خانواده می‌پردازد، اثر ندارد. اما شغل اصلی به صورت اثر متقابل با متغیرهای مستقل سن، شغل همسر، و نوع منزل مسکونی بر مبلغ اقساط ماهانه اثر می‌کند. در اینجا نیز با توجه به علامت جبری ضریبها و قدر مطلق آنها به راحتی می‌توان نشان داد که مبلغ اقساط ماهانه‌ای که غیردانشگاهیان می‌پردازند بیشتر از آن دانشگاهیان است. مثلاً تفاوت اقساط ماهانه یک فرد دانشگاهی، که متوسط سن گروه خود را دارد و همسرش خانه‌دار است و در آپارتمان زندگی می‌کند، با فرد غیردانشگاهی، که متوسط

سن گروه خود را دارد و همسرش خانه‌دار است و در آپارتمان زندگی می‌کند، با فرد غیر دانشگاهی مشابه، برابر است با ۵۹۵۴۲ - واحد در ماه. برای ترکیبهای دیگر از این سه متغیر نیز می‌توان به همین نحو نشان داد که اقساط دانشگاهیان کمتر از غیر دانشگاهیان است. بنابراین، به طور کلی نیز دانشگاهیان قسط کمتری نسبت به غیر دانشگاهیان می‌پردازند.

اگر بخواهیم یافته‌های فوق را یکجا و به طور خلاصه عرضه کنیم می‌توانیم جدول زیر را ارائه دهیم. در این جدول علامت (+) نشانه آن است که دانشگاهیان هزینه بیشتری نسبت به غیر دانشگاهیان، در هر یک از اقلام مورد بحث، دارند. علامت (-) نشانه آن است که هزینه دانشگاهیان در قلم مربوط، کمتر از همالان ایشان است و علامت (=) بدان معنی است که تفاوتی بین آن دو گروه وجود ندارد.

جدول ۴-۱. مقایسه هزینه‌های دانشگاهیان و غیر دانشگاهیان برای اقلام مختلف هزینه

جامعه	تحصیل	درمان	تغذیه	پوشاک	فرهنگی	تفریح	مسافرت	اقساط
دانشگاهیان	=	-	+	-	-	-	-	-
غیردانشگاهیان	=	+	-	+	+	+	+	+

نتیجه کلی مدل‌های رگرسیونی:

از بین ۸ قلم هزینه‌های عمده خانوار، در ۶ قلم دانشگاهیان هزینه‌ای کمتر از غیر دانشگاهیان دارند، در یک مورد هزینه بیشتر دارند و در یک مورد تفاوت وجود ندارد. بنابراین، غیر دانشگاهیان در رفاه نسبی بیشتری هستند.

۳-۴. تفسیر رگرسیون لوجستیک

در بخش ۲-۳، بهترین مدل لوجیت را به دست آوردیم که به ترتیب در بر دارنده متغیرهای مستقل شغل اصلی، تعداد مشاغل اضافه بر شغل اصلی، وضع مسکن، مساحت زیر بنای ساختمان مسکونی، تعداد افراد تحت تکفل، محل اخذ مدرک تحصیلی، تعداد فرزندان، مکان جغرافیایی خدمت، نوع منزل مسکونی و آخرین مدرک تحصیلی، چه به صورت انفرادی و چه به صورت ترکیب با دیگر متغیرهای مستقل است. در زیر به تفسیر این مدل می‌پردازیم: چنانکه در مورد لگاریتم نسبت بختها یا لوجیت در بخش ۲-۳، توضیح دادیم ضریب هر متغیر موجود در مدل و علامت آن نشان دهنده آن است که آیا با افزایش این متغیر، شانس قرار

گرفتن فرد در گروه پر هزینه نسبت به شانس واقع شدن او در گروه کم هزینه، زیاد می شود یا کم؟ در صورتی که متغیری دو حالتی و تنها دارای اثر اصلی باشد، یعنی به طور انفرادی در مدل ظاهر شود، ضریب مثبت آن باعث افزایش شانس برای قرار گرفتن در گروه پر هزینه می شود. به عبارت دیگر، خواهیم گفت که "باعث افزایش هزینه" می شود. برعکس ضریب منفی در این حالت موجب آن می شود که شانس قرار گرفتن فرد در گروه کم هزینه زیاد شود. اما از آنجا که این گونه ضرایب مقادیر برآورد شده هستند، بهتر است که به فاصله اطمینان ۹۵ درصد ضرایب واقعی معادله توجه کنیم. این فاصله ها نشان دهنده آن هستند که با اطمینان ۹۵ درصد، ضرایب واقعی (نه برآورد شده) در چه دامنه ای قرار دارند. در صورتی که این دامنه عدد صفر (لگاریتم عدد یک) را در بر داشته باشد، به معنی آن است که ضریب واقعی می تواند برابر صفر باشد و شانس قرار گرفتن فرد در هر یک از دو گروه پر هزینه و کم هزینه برابر است، یعنی نسبت دو شانس برابر یک و لگاریتم آن صفر است. بدین ترتیب امکان مقایسه دو نفر که دارای متغیرهای مستقل دو حالتی با مقادیر متفاوت هستند، آسان است. زمانی که متغیر مستقلی دارای چند حالت باشد، باید یکی از مقادیر آن، معمولاً کمترین مقدار را، به عنوان مرجع بپذیریم و سایر حالتها را، مثل مورد دو حالتی، هر بار یکی از چند حالت را با مرجع مقایسه کنیم. در مورد متغیرهای مستقل پیوسته که در واقع بینهایت مقدار دارند، می توانیم تغییرات لوجیت را به ازای مقدار مشخصی افزایش در متغیر مستقل بررسی کنیم. در این صورت مسأله را به صورتی که انگار متغیر مستقل دو حالت دارد، تبدیل کرده ایم. بدین ترتیب، وابستگی لوجیت را به مقدار افزایش متغیر مستقل به دست می آوریم که نشان می دهد، به ازای افزایش یک واحد در متغیر مستقل، لوجیت چقدر تغییر می کند.

وقتی بخواهیم وابستگی لوجیت را فقط به یکی از متغیرهای مستقل بررسی کنیم، فرض می کنیم که سایر متغیرها در سطح ثابتی نگه داشته می شوند. به عبارت دیگر، دو فرد که صرفاً از لحاظ یکی از متغیرها با هم تفاوت دارند، مقایسه می شوند. در این بررسی چون هدف ما مقایسه دانشگاهیان و غیر دانشگاهیان به نحوی است که از لحاظ سایر متغیرهای مستقل شبیه باشند، کافی است که لوجیت را برای یک فرد دانشگاهی و یک فرد غیر دانشگاهی که سایر متغیرهای آنها مثل هم هستند، بنویسیم و تفاوت این دو مقدار را بیابیم. در مورد متغیرهای دو حالتی که فقط اثر اصلی آنها در مدل ظاهر می شوند، تفاوت بالا برابر ضریب مربوطه خواهد شد. در مورد متغیرهایی که اثر متقابل با سایر متغیرها دارند، باید ترکیبهای خاص متغیر مورد

نظر را با متغیرهایی که با آنها اثر متقابل دارد، بررسی کنیم.

اکنون، همان گونه که در مدل (۱-۲-۳) دیده می‌شود، متغیر شغل اصلی در معادله لوجیت دارای اثر اصلی با ضریب $3/1844$ برای سطح ۱ است. همچنین، این متغیر دارای اثر متقابل با تعداد فرزندان، و آخرین مدرک تحصیلی است. تعداد فرزندان متغیری حقیقی است و آخرین مدرک تحصیلی دارای دو سطح است. بنا به ضرورت برنامه‌های کامپیوتری، لازم است که هر یک از سطوح را با کد یا متغیری ظاهری معرفی کنیم.

پس از مساوی انگاشتن سایر متغیرها، به جز شغل اصلی و آخرین مدرک تحصیلی و ساده کردن معادله، تفاضل لوجیت برای افراد دانشگاهی دارای درجه دکتری با تعداد X فرزند و افراد غیر دانشگاهی دارای درجه دکتری با X فرزند، به صورت $1/9 - 1/08X$ در می‌آید. از این عبارت چنین نتیجه می‌گیریم که اگر X صفر یا یک باشد، دانشگاهیان هزینه بیشتری دارند. یعنی خانوارهای بدون فرزند یا یک فرزندی دانشگاهیان بیشتر از خانوارهای مشابه غیر دانشگاهیان هزینه دارند. اگر X از $1/76$ بیشتر باشد، وضعیت وارونه می‌شود و غیر دانشگاهیان هزینه بیشتری دارند. چون متوسط تعداد فرزند برای دانشگاهیان ۲ است تفاضل لوجیت بالا منفی است و در نتیجه در سطح دکتری به طور متوسط دانشگاهیان کمتر از غیر دانشگاهیان هزینه می‌کنند.

به همین نحو اگر مقایسه را در سطح دارندگان درجه فوق لیسانس مقایسه کنیم، نتیجه‌ای مشابه می‌گیریم، منتهی تفاوت چندان چشمگیر نیست.

۴-۴. نتیجه گیری کلی

با توجه به نتایج به دست آمده از آمارهای توصیفی، تحلیلی و بررسی رگرسیون خطی در مورد اقلام هزینه‌ها و بررسی رگرسیون لوجستیک در مورد کل هزینه‌ها، و با آزمونهایی که انجام دادیم، می‌توان درباره فرضیه‌های مذکور در بخش ۱-۴ به شرح زیر قضاوت کرد:

چنانکه در بخش ۱-۲-۲ دیدیم، هزینه‌های اعلام شده دانشگاهیان برای یک ماه بسیار کمتر از هزینه متوسط خانوارهای شهری است که مرکز آمار ایران منتشر کرده است. یعنی در حدود نصف آن است. در تفسیر این نتیجه باید گفت که رقمی که مرکز آمار ایران منتشر کرده از پرسشنامه‌های بسیار مفصلتر به دست آمده است که در بردارنده جزئیترین هزینه‌های خانوارهاست. در حالی که نتیجه ما (۲۲۲۵۴۳ ریال در ماه)، تنها مجموع ۸ قلم هزینه مشخص است. حال اگر هزینه‌هایی را که نپرسیده‌ایم، در حدود مجموع این ۸ قلم به شمار آوریم. باز هم از

رقم (۵۳۱۳۸۵ ریال در ماه)، که در گزارش مرکز آمار آمده است، کمتر خواهد بود. این امر نشانه آن است که به طور کلی دانشگاهیان چنان که باید از حقوق و دستمزد مناسب بهره‌مند نیستند. این حکم تا حدودی در مورد غیر دانشگاهیان هم‌رتبه اعضای هیأت علمی نیز صادق است. زیرا تفاوت میانگینهای دو زیر جامعه معنی دار نیست.

اگر اختلافات موجود بین مشخصات فردی اعضای هیأت علمی و همالان آنها را در نظر بگیریم، بین میانگینهای مجموع هزینه‌ها برای اعضای هیأت علمی و همالان آنها در خارج از دانشگاه تفاوت وجود ندارد. پس از لحاظ کردن مشخصات فردی، چون میزان هزینه‌های اعضای هیأت علمی کمتر از همالان آنها در خارج از دانشگاه است، در صورتی که غیر دانشگاهیان را نمونه‌ای از قشر تحصیل کرده کشور بدانیم، وضع اقتصادی اعضای هیأت علمی به طور نسبی بدتر از سایر تحصیل کرده‌هاست. با آنکه اعضای هیأت علمی اغلب افراد نخبه‌ای هستند که به امر خطیر آموزش و پرورش و تربیت نیروی انسانی مورد نیاز کشور در همه رشته‌ها و زمینه‌ها می‌پردازند، بنابر یافته‌های این تحقیق دستمزدی متناسب با این مسؤولیت خطیر دریافت نمی‌کنند.

اولویت هزینه‌های زندگی برای اعضای هیأت علمی و همالان آنها در خارج از دانشگاه الگوی یکسانی ندارد. چنانکه دیدیم اعضای هیأت علمی به غیر از هزینه تغذیه، در سایر موارد کمتر از همالان خود هزینه می‌کنند و در مورد هزینه تحصیل رسمی فرزندان با همالان خود یکسان عمل می‌کنند. این واقعیت بیانگر آن است که تغذیه مناسب برای اعضای هیأت علمی در قیاس با همالان آنها، مهمتر از هزینه‌هایی مانند تفریح، مسافرت و غیره است.

۴-۵. پیشنهاد

در حین اجرای این تحقیق به نکاتی برخوردیم که بی‌مناسبت نیست برای بهره‌گیری آیندگان و کسانی که بخواهند در این راه به تحقیق بپردازند، به آنها اشاره کنیم. اکثر خانوارهای ایرانی و از آن جمله خانواده‌های اعضای هیأت علمی، سابقه‌ای مکتوب و دقیق از هزینه‌های خود ندارند. از این رو سنجش هزینه‌ها علی‌الاصول متکی به حافظه پاسخگویان است که نمی‌تواند به حد کافی دقیق باشد. بررسی آمارها نشان می‌دهد که گاهی ارقام غیر واقعی به دلیل اتکا به حافظه گزارش شده‌اند. از این رو پیشنهاد می‌شود که در صورتی که بخواهیم این بررسی را به طور عینی‌تر و متکی به سنجشهای دقیقتر اجرا کنیم، لازم است گروهی از دانشگاهیان و غیردانشگاهیان را که به عنوان نمونه انتخاب می‌شوند، با اهداف و مشخصات طرح آشنا کنیم.

در قبال نوعی پاداش یا جایزه که به آنها داده خواهد شد از ایشان بخواهیم که برای مدتی مثلاً دو تا سه سال با مجریان طرح همکاری کنند. در این صورت می توان دفترچه هایی را در اختیار آنان گذاشت و آنها را طوری طراحی کرد که خانوارها بتوانند سابقه ای مکتوب و دقیق از هزینه های خود نگهداری کنند. در پایان هر ۶ ماه، می توان نتیجه اندازه گیریها را گردآوری و تجزیه و تحلیل کرد تا وقتی به پایان دوره می رسیم بتوان تصویری کامل از وضع رفاهی اعضای هیأت علمی و مقایسه آن با همالان آنها عرضه داشت.

مآخذ

○ اعضای گروه پژوهشهای آماری وزارت فرهنگ و آموزش عالی-شناخت ویژگیهای اعضای هیأت علمی دانشگاهها ومؤسسات آموزشی و پژوهشی در سال ۱۳۷۱، مؤسسه پژوهش و برنامه ریزی آموزش عالی، تهران ۱۳۷۲.

○ Madansky, A. (1988). *Prescriptions for Working Statisticians*.

Springer-Verlag, New York.

○ Bhattacharyya, G. K. and R. A. Johnson, (1977). *Statistical Concepts and Methods*. New York, John Wiley and Sons.

○ Hosmer, Jr. David and S. Lemeshow (1989). *Applied Logistic Regression*. New York, Wiley.

○ Johnson, R. A. and D. W. Wichern (1988). *Applied Multivariate Statistical Analysis*, (2nd Ed.). New Jersey, Prentice Hall Inc.

○ Press, S. James (1972). *Applied Multivariate Analysis*. New York, Holt, Rinehart and Winston, Inc.

○ Searle, S.R. (1971). *Linear Models*. New York. John Wiley & Sons, Inc. (1971).