

## برآورد میزان هزینه‌های اقتصادی بیماری‌های عملکردی دستگاه گوارش با استفاده از مدل رگرسیون پارامتری دو بخشی

محدثه شجاعی شاهرخ آبادی: کارشناسی ارشد آمار زیستی، گروه آمار زیستی، دانشکده پزشکی، دانشگاه تربیت مدرس، تهران، ایران. m\_shojai@modares.ac.ir  
 \* دکتر انوشیروان کاظم نژاد: استاد آمار زیستی، گروه آمار زیستی، دانشکده پزشکی، دانشگاه تربیت مدرس، تهران، ایران. (\*نویسنده مسئول). kazem\_an@modares.ac.ir  
 دکتر فرید زایری: استادیار آمار زیستی، دانشگاه علوم پزشکی شهید بهشتی، دانشکده پیراپزشکی، گروه آمار زیستی، تهران، ایران. fzayeri@yahoo.com  
 محسن واحدی: کارشناس ارشد آمار زیستی، مرکز تحقیقات بیماری‌های گوارش و کبد، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران. mohsenvahedi540@gmail.com

تاریخ دریافت: ۹۱/۴/۳ تاریخ پذیرش: ۹۱/۷/۱۶

### چکیده

**زمینه و هدف:** برآورد دقیق هزینه‌های پزشکی یکی از اهداف سیاست‌گذاران مراقبت‌های بهداشتی است. با توجه به ویژگی داده‌های مطالعات بهداشتی و پیچیدگی تحلیل آن‌ها، استفاده از مدل‌های معمولی مناسب نیست. این مطالعه با هدف برآورد میانگین هزینه‌ی اختلالات دستگاه گوارش و بررسی برخی از عوامل تاثیرگذار بر آن با استفاده از رگرسیون پارامتری دویخشی صورت گرفته است.

**روش کار:** طی سال‌های ۸۵-۸۶ اطلاعات مربوط به ۱۹۰۷ فرد مبتلا به بیماری‌های عملکردی دستگاه گوارش در مرکز تحقیقات بیماری‌های گوارش و کبد دانشگاه علوم پزشکی شهید بهشتی تهران به صورت مقطعی گذشته‌نگر جمع‌آوری گردید. برای مدل‌بندی هزینه‌ی بیماری‌های عملکردی دستگاه گوارش، مدل دویخشی استفاده شد. به این صورت که در بخش اول مدل، روی هزینه‌های دو حالتی [صفر و غیر صفر] رگرسیون لوجستیک و در بخش دوم روی هزینه‌های غیر صفر رگرسیون خطی چند متغیره برازش داده شد و تاثیر متغیرهای فردی، میزان توانایی انجام کار، دفعات مراجعه به پزشک، تعداد آزمون‌های تشخیصی، بیمه، تعداد روزهای غیبت از کار یا کاهش کارایی و بستری در بیمارستان بر هزینه سنجیده شد.

**یافته‌ها:** هزینه‌ی بیماری‌های دستگاه گوارش ۷۵/۹۳ دلار با انحراف معیار ۱۲۲/۲۹ دلار، حداقل ۲/۸۹ و حداکثر ۱۳۹۴/۳۲ دلار برآورد شد. درحالی‌که میانگین هزینه‌ی واقعی در نمونه‌ی مورد نظر ۷۸/۳۵ دلار با انحراف معیار ۲۲۲/۳۶، حداقل و حداکثر صفر و ۵۱۸۳/۸۱ دلار بوده است. نتایج حاصل از مدل دویخشی نشان داد، برخی متغیرهای فردی، تعداد روزهای غیبت از کار یا کاهش کارایی و دفعات مراجعه به پزشک در داشتن هزینه و متغیرهای تعداد روزهای غیبت از کار یا کاهش کارایی، دفعات مراجعه به پزشک و بستری در بیمارستان روی هزینه‌های غیر صفر تاثیرگذار است.

**نتیجه‌گیری:** هزینه‌ی برآورد شده از این روش به مقدار واقعی نزدیک است. با توجه به پایین‌تر بودن انحراف معیار برآوردهای مدل نسبت به مقادیر واقعی، برآوردهای حاصل از این مدل از دقت بالایی برخوردار هستند. به علاوه این مدل در حداکثر ساختن هزینه‌های صفر که یکی از اهداف ما بوده است، خوب عمل کرده است.

**کلیدواژه‌ها:** مدل دویخشی، هزینه‌های اقتصادی بیماری، رگرسیون پارامتری دویخشی، بیماری‌های دستگاه گوارش.

### مقدمه

بیمارستان‌ها، شرکت‌های بیمه و سلامت ثبت شده‌اند. ترکیب این اطلاعات با تکنولوژی‌های جدید تجزیه و تحلیل آماری می‌تواند به‌طور گسترده‌ای برای بالا بردن درک تأثیر عوامل مختلف بر روی هزینه به‌کار برده شود (۲). از طرفی اختلالات عملکرد دستگاه گوارش در بسیاری از جوامع مشترک است، از این رو انتظار می‌رود که بار اقتصادی قابل توجهی بر جامعه داشته باشند. در بیشتر کشورهای توسعه یافته و در حال توسعه هزینه‌های ناشی از ابتلا به بیماری‌های گوارشی بار سنگینی را به مسوولین بهداشتی تحمیل می‌کند (۳). این بیماری‌ها نه تنها

سیاست‌گذاران مراقبت‌های بهداشتی به دنبال استنباط‌های واقعی و برآوردهای دقیق از هزینه‌های پزشکی آینده‌ی جامعه هستند (۱). بحث اصلاح مراقبت‌های بهداشتی و علاقه‌مندی به نحوه‌ی توسعه‌ی برنامه‌های ملی مراقبت، انگیزه‌ی ما برای فهمیدن تأثیر نسبی شاخص‌های جمعیتی، سطوح مختلف بیماری و تغییر در نوع خدمات ارائه شده بر هزینه‌ی بیماری در الگوهای عملی هستند. با استفاده از تکنیک‌های پیشرفته‌ی مدیریت و روش‌های پیشرفته‌ی جمع‌آوری داده‌ها، اطلاعات هزینه‌های پزشکی به صورت منظم توسط

سال هیچ هزینه‌ای پرداخت نکنند، در حالی که به طور مشابه ممکن است افراد مبتلا به یک بیماری خاص (مثلاً سرطان) با پیشرفت بیماری، هزینه‌های پزشکی فوق‌العاده‌ای پرداخت کنند. علاوه بر این، تفاوت محسوسی بین هزینه‌های بیماران سرپایی و بیماران بستری شده در بیمارستان وجود دارد. همه‌ی این مشاهدات نشان می‌دهند که هیچ توزیع پارامتری ساده‌ای برای توصیف چنین داده‌های نیمه پیوسته‌ای مناسب نیست و وقتی هدف یافتن میانگین هزینه‌ی جامعه‌ی بیماران برای یک بیماری خاص است، بدون در نظر گرفتن این ویژگی‌ها، برآوردها و تحلیل‌های آماری غیر واقعی خواهند بود. نیاز به شناخت طبیعت نیمه پیوسته‌ی داده‌های هزینه با توسعه‌ی مدل دو بخشی آغاز شد. برای اولین بار در اقتصاد سلامت و به صورت خاص در مطالعات هزینه‌های بهداشتی، مدل‌های چند بخشی توسط دوان مطرح شدند. در واقع مدل دو بخشی با انعطاف‌پذیری، اجازه می‌دهد هزینه‌های صفر و غیر صفر توسط دو فرایند مجزا تولید شوند، همچنین مشکلات ناشی از مقادیر صفر زیاد هزینه‌ها در مدل یک بخشی را با جدا کردن رفتار مدل به دو قسمت تصحیح می‌کند: تصمیم در مورد داشتن هزینه‌های غیر صفر و در گام بعد تصمیم در مورد سطوح هزینه‌های غیر صفر. برآوردهای به‌دست آمده از مدل دو بخشی کارا، دقیق و سازگار هستند. به علاوه این مدل‌ها به مقادیر نهایی حساس نیستند. برآوردهای حاصل از مدل دو بخشی بسیار دقیق‌تر از مدل یک بخشی‌اند، زیرا برازش حاصل از این مدل بسیار نزدیک به توزیع اصلی داده‌ها است. خواص و ویژگی‌های مدل‌های چند بخشی و کاربرد آن‌ها در داده‌های پزشکی طی سال‌های متمادی مورد بررسی قرار گرفته‌اند (۲ و ۱۰-۱۲).

در این پژوهش سعی شده است برای اولین بار در ایران با برازش مدل دو بخشی پارامتری بر اطلاعات ۲۹۲۹ فرد مبتلا به بیماری‌های عملکردی دستگاه گوارش (تمام بیماری‌هایی که ناشی از اختلال در دستگاه گوارش می‌باشند مانند رفلکس معده به مری، یبوست، اسهال، سندرم

باعث استفاده‌ی بیشتر از مراقبت‌های بهداشتی می‌شوند، بلکه بر کار افراد و میزان بهره‌وری آن‌ها نیز تأثیر دارند. ماهیت مزمن بیماری‌های گوارشی و نقاط مبهم تشخیص و درمان آن‌ها می‌تواند هزینه را افزایش دهد (۴). بر اساس مطالعات مختلفی که به‌طور همزمان در نقاط مختلف جهان انجام شده است، سوءهاضمه بین ۱۱ تا ۲۷ درصد، سندرم رودی تحریک‌پذیر بین ۱۰ تا ۱۵ درصد و رفلکس معده به مری بین ۱۸ تا ۴۰ درصد شیوع داشته‌اند. به علاوه، بنابر گزارش‌های منتشر شده، مبتلایان به این بیماری‌ها به‌طور متناسب از منابع مراقبت‌های بهداشتی استفاده نمی‌کنند. شیوع بالای این بیماری‌ها و هزینه‌ی بالای مراقبت‌های بهداشتی برای هر بیمار باعث می‌شود که این بیماری‌ها به‌عنوان منبع قابل توجه‌ای برای هزینه‌های مراقبت‌های سلامت و هدفی بالقوه برای کاهش این هزینه‌ها شناخته شوند (۵-۷). گرچه هزینه‌های اقتصادی بیماری‌های دستگاه گوارشی در بسیاری از کشورهای توسعه یافته مورد مطالعه قرار گرفته است، اما هنوز هم از کشورهای در حال توسعه مانند ایران داده‌های کمی در این زمینه وجود دارد (۵). در این زمینه روش‌های تجزیه و تحلیل هزینه‌ی اختلالات روده‌ای در ایران پرداخته است و هزینه‌ی مستقیم و غیر مستقیم را برای دو گروه با مشاوره و بدون مشاوره در سندرم روده تحریک‌پذیر، اختلالات ناشناخته‌ی روده‌ای، یبوست و نفخ شکم گزارش کرده است (۸و۵). در تحقیقی دیگر به مطالعه‌ی هزینه‌ی بیماری سوءهاضمه و رفلکس معده به مری پرداخته شده است (۹).

از طرفی داده‌های مطالعات بهداشتی دارای ویژگی‌هایی هستند که باعث پیچیدگی تحلیل آن‌ها می‌شود. توزیع داده‌های هزینه بسیار چوله است زیرا بعضی از بیماران هزینه‌های بسیار زیادی متحمل می‌شوند؛ همچنین هزینه‌های پزشکی بسیاری از افراد در یک بازه زمانی خاص صفر است. در واقع داده‌های هزینه‌های پزشکی غالباً چوله به راست و شامل مقادیر قابل توجهی صفر هستند و ممکن است به صورت ناهمگن پراکنده شده باشند. برای مثال ممکن است افراد سالم در

آزمون‌های تشخیصی، بیمه (نداشته‌اند، بیمه دولتی و بیمه تکمیلی)، بستری شدن در بیمارستان (بستری شدن یا درمان سرپایی) و هزینه‌ی بیماری هر فرد استخراج شدند تا در مرحله تحلیل از آن‌ها استفاده شود.

در این مطالعه مجموع هزینه‌های مستقیم و غیر مستقیم بیماری‌های دستگاه گوارش به‌عنوان هزینه‌ی کل هر فرد در نظر گرفته شد. هزینه‌های مستقیم، همان هزینه‌هایی هستند که فرد آن‌ها را مستقیماً پرداخت کرده است مانند هزینه‌ی ویزیت، دارو، آزمایش و غیره. این هزینه‌ها برای هر فرد به دلار ثبت شده‌اند. هزینه‌های غیر مستقیم هزینه‌هایی را در بر می‌گیرند که نمی‌توان آن‌ها را به صورت هزینه‌های مستقیم به حساب آورد که شامل درآمد از دست رفته توسط بیمار و ارزش اقتصادی هزینه‌های نامحسوس مثل مرگ، درد، کیفیت زندگی و حتی ناراحتی می‌شود. در واقع هزینه‌های غیر مستقیم ضررهای ناشی از ناراحتی‌های بیمار است که با ضرب کردن متوسط درآمد روزانه‌ی افراد شاغل در تعداد روزهای غیبت از کار و یا روزهای ناراحتی محاسبه می‌شود (برای به‌دست آوردن متوسط درآمد روزانه‌ی هر فرد شاغل از اطلاعات موجود در بانک جهانی استفاده شده است). همان‌طور که ملاحظه می‌شود متغیر هزینه یک متغیر مثبت با مقادیر صفر زیاد است. هزینه‌های غیر صفر در واقع همان هزینه‌هایی هستند که فرد پرداخت آن را احساس کرده است. اما هزینه‌های صفر لزوماً صفر حقیقی نیستند و ممکن است فرد حداقل هزینه‌هایی را به‌صورت غیر مستقیم پرداخت کرده باشد اما به دلیل محسوس نبودن این هزینه‌ها آن را صفر در نظر گرفته است.

**تجزیه و تحلیل آماری داده‌ها:** در این تحقیق برای مدل بندی هزینه‌ی بیماری‌های عملکردی دستگاه گوارش، مدل دو بخشی انتخاب شد. به این صورت که در بخش اول مدل آمیخته‌ی دو بخشی، روی هزینه‌های دو حالتی (صفر و غیر صفر) رگرسیون لوجستیک و در بخش دوم روی هزینه‌های غیر صفر، رگرسیون خطی چند متغیره برازش داده شد. در بخش اول مدل برای بررسی

روده تحریک پذیر، درد شکم، نفخ، سوزش سر دل، تهوع و استفراغ، اشکال در بلع، سوء هاضمه، کاهش وزن و درد رکتوم)، برآورد دقیقی از میانگین هزینه‌های بیماری‌های دستگاه گوارش ارابه و متغیرهای تأثیر گذار بر آن شناسایی شوند. داده‌های موجود از هزینه‌های بیماری‌های دستگاه گوارش به دلیل حجم بالای صفر در هزینه‌ها و همچنین هزینه‌های سنگین استفاده از مدل‌های پارامتری معمولی را غیر ممکن می‌سازند، بنابراین به نظر می‌رسد استفاده از مدل دو بخشی پارامتری مشکل مدل‌های معمولی پارامتری حل کند.

### روش کار

**داده‌های مربوط به بیماری‌های عملکردی دستگاه گوارش:** اطلاعات مربوط به ۲۹۲۹ بیمار مبتلا به بیماری‌های عملکردی که طی سال‌های ۸۵-۸۶ دستگاه گوارش به مرکز تحقیقات بیماری‌های گوارش و کبد دانشگاه علوم پزشکی شهید بهشتی تهران مراجعه کرده بودند، به صورت گذشته‌نگر جمع‌آوری گردید. اطلاعات افراد به صورت پرسشنامه و مصاحبه حضوری از افراد ثبت گردید. پس از ویرایش اطلاعات و حذف داده‌های ناقص، تعداد ۱۹۰۷ نفر در مطالعه باقی ماندند. این طرح در کمیته‌ی اخلاق مرکز تحقیقات بیماری‌های گوارش و کبد بیمارستان طالقانی به تصویب رسیده است. برای همه‌ی بیماران و بر اساس اطلاعات موجود، متغیرهایی که از نظر محقق می‌توانند بر هزینه تأثیر گذار باشند نظیر سن، جنسیت (زن و مرد)، تحصیلات (فوق لیسانس و بالاتر، لیسانس، دیپلم، پایین تر از دیپلم و ابتدایی)، وضعیت تأهل (مجرد، مطلقه، بیوه، متأهل)، میزان توانایی انجام کار در طول دوره‌ی بیماری (هرگز، کم، متوسط، زیاد و خیلی زیاد)، تعداد بیماری‌هایی که فرد به آن‌ها مبتلا است (رفلکس معده به مری، یبوست، اسهال، سندرم روده تحریک پذیر و ...)، تعداد روزهای غیبت از کار و کاهش کارایی (تعداد روزهایی که خود فرد اعلام کرده است که به دلیل بیماری غیبت از کار یا کاهش کارایی داشته است)، تعداد دفعات مراجعه به پزشک عمومی و متخصص، تعداد

جدول ۱. توصیف متغیرهای کیفی مورد مطالعه

متغیر	طبقه	فراوانی	درصد
جنسیت	زن	۱۰۹۳	۵۷/۳
	مرد	۸۱۴	۴۲/۷
تحصیلات	فوق لیسانس و بالاتر	۳۵	۱/۸
	لیسانس	۲۵۹	۱۳/۶
	دیپلم	۵۲۸	۲۷/۷
	پایین‌تر از دیپلم	۶۰۵	۳۱/۷
وضعیت تأهل	ابتدایی	۴۸۰	۲۵/۲
	مجرد	۱۹۸	۱۰/۴
	متأهل	۱۵۵۷	۸۱/۶
	بیوه	۱۲۲	۶/۴
سرپرست خانوار	مطلقه	۳۰	۱/۶
	بله	۷۰۰	۳۶/۷
	خیر	۱۲۰۷	۶۳/۳
بیمه	نداشته‌اند	۵۲۵	۲۷/۵
	بیمه دولتی	۱۳۶۳	۷۱/۵
میزان توانایی انجام کار	بیمه تکمیلی	۱۹	۱/۰
	هرگز	۱۱۰۸	۵۸/۰
	کم	۳۸۴	۲۰/۱
	متوسط	۳۰۷	۱۶/۱
	زیاد	۶۱	۳/۲
مصرف دارو	خیلی زیاد	۴۸	۲/۵
	ندارد	۸۳۶	۴۳/۸
بستری در بیمارستان	دارد	۱۰۷۱	۵۶/۲
	نداشته‌اند	۱۸۳۷	۹۶/۳
	داشته‌اند	۷۰	۳/۷

جدول ۲. توصیف آماره‌های متغیرهای کمی مورد مطالعه

متغیر	میانگین	انحراف معیار	حداقل	حداکثر
تعداد بیماری‌هایی که فرد به آن‌ها مبتلا است	۲/۵۲	۱/۵۷	۰	۱۰
تعداد روزهای غیبت از کار و کاهش کارایی	۱/۲۸	۶/۶۷	۰	۱۰۵
دفعات مراجعه به پزشک عمومی و متخصص	۱/۶۵	۲/۲۴	۰	۲۰

در هم، برآورد نهایی از هزینه‌ی هر فرد آرایه شد. همچنین در بخش اول و دوم مدل، متغیرهای تأثیرگذار بر هزینه به دست آورده شد. برای تجزیه و تحلیل داده‌ها از SPSS ویراست ۱۶ استفاده شد. در تمامی مراحل سطح معنی‌داری ۰/۰۵ در نظر گرفته شده است.

مدل پارامتری دو بخشی با جدا کردن رفتار مدل به دو قسمت، مدل لوجیت برای احتمال هزینه‌ی غیر صفر در قسمت اول:

$$\begin{aligned} \text{logit}\{P(Y_i|X_i)\} \\ &= X_i\beta_1 \\ &+ \eta_{1i} \quad \eta_{1i} \sim N(0, 1) \quad i \\ &= 1, \dots, N \end{aligned}$$

رابطه‌ی بین داشتن هزینه و متغیرهای مستقل سن، جنسیت، تحصیلات، وضعیت تأهل، میزان توانایی انجام کار، دفعات مراجعه به پزشک عمومی و متخصص، تعداد آزمون‌های تشخیصی و بیمه از رگرسیون لوجستیک استفاده شد و احتمال داشتن هزینه‌ی غیر صفر برای هر فرد محاسبه گردید. سپس در بخش دوم مدل برای بررسی رابطه‌ی بین هزینه‌های غیر صفر و متغیرهای مستقل میزان توانایی انجام کار، تعداد روزهای غیبت از کار یا روزهای کاهش کارایی و بستری در بیمارستان از رگرسیون خطی چندگانه استفاده و میانگین هزینه‌ی مورد انتظار برای هر فرد ثبت شد. با ضرب نتایج به دست آمده از دو بخش مدل

در  $E(Y_i|X_i)$ ، متوسط هزینه‌ی پزشکی مورد انتظار برای هر فرد را برآورد می‌کند (۲).

### یافته‌ها

پس از ویرایش اطلاعات، نمونه‌ای به حجم ۱۹۰۷ نفر در مطالعه باقی ماند. سن افرادی که در مطالعه حضور داشتند بین ۲۰ تا ۸۰ سال با میانگین ۴۵/۸۵ و انحراف معیار ۱۴/۸۸ بود. همچنین حداقل و حداکثر وزن این افراد به ترتیب برابر با ۴۳ و ۱۵۰ کیلوگرم با میانگین ۷۲/۴۷ و انحراف معیار ۱۱/۷۴ کیلوگرم بود. توصیف سایر متغیرهای کیفی و کمی مورد مطالعه در جدول ۱ آورده شده است.

در جداول ۱ و ۲ توصیفی از متغیرهای تأثیر

که در آن  $X_i$  متغیر کمکی هستند و مدل خطی بر پایه‌ی مقیاس لگاریتم برای سطوح هزینه‌ی غیر صفر در قسمت دوم:

$$\log(Y_i|Y_i > 0) = X_i\beta_2 + \eta_{2i} \quad \eta_{2i} \sim N[0, \sigma^2]$$

برآوردی از هزینه‌ی آرایه می‌دهد. در مدل دو بخشی هزینه‌ی پزشکی مورد انتظار برای یک بیمار به صورت زیر است:

$$E(Y_i|X_i) = \pi_i \cdot E(Y_i|Y_i > 0, X_i) = \pi_i \cdot \exp[X_i\beta_2 + \sigma^2/2]$$

که در آن:

$$\pi_i = pr(Y_i > 0) = \{1 + \exp(-X_i\beta_1)\}^{-1}$$

با برآورد  $\beta_1$ ،  $\beta_2$  و  $\sigma^2$  و جایگزین کردن آن

جدول ۳. نتایج رگرسیون لوجستیک چندگانه برای بررسی رابطه بین داشتن هزینه و متغیرهای مستقل

متغیر	طبقه	برآورد	انحراف معیار	p-value	OR
سن	-	-۰/۰۰۶	۰/۰۰۵	۰/۲۲۵	۰/۹۹۴
جنسیت	زن	۰/۰۸۸	۰/۱۴۰	۰/۵۲۷	۱/۶۴۲
	مرد				
تحصیلات	فوق لیسانس و بالاتر	۰/۲۵۵	۰/۵۸۷	۰/۶۶۴	۱/۲۹۰
	لیسانس	-۱/۰۳۳	۰/۲۴۴	<۰/۰۰۱	۰/۳۵۶
	دیپلم	-۰/۹۵۸	۰/۲۰۸	<۰/۰۰۱	۰/۳۸۴
	پایین‌تر از دیپلم	-۰/۵۹۶	۰/۲۰۲	۰/۰۰۳	۰/۵۵۱
وضعیت تأهل	ابتدایی				
	مجرد	۱/۹۲۷	۰/۴۸۸	<۰/۰۰۱	۶/۸۷۰
	مطلقه	۱/۸۲۵	۰/۴۵۱	<۰/۰۰۱	۶/۲۰۳
	بیوه	۱/۸۰۹	۰/۵۳۷	<۰/۰۰۱	۶/۱۰۹
میزان توانایی انجام کار	متأهل				
	هرگز	-۰/۷۴۲	۰/۶۳۵	۰/۲۴۲	۰/۴۷۶
	کم	-۰/۲۸۵	۰/۶۴۸	۰/۶۶۰	۰/۷۵۲
	متوسط	-۰/۵۶۳	۰/۶۵۲	۰/۳۸۳	۰/۵۶۷
	زیاد	-۰/۲۳۱	۰/۷۴۵	۰/۷۵۶	۰/۷۹۴
	خیلی زیاد				
	تعداد بیماری‌هایی که فرد به آن‌ها مبتلا است -		۰/۱۰۰	۰/۰۴۲	۰/۰۱۷
تعداد روزهای غیبت از کار و کاهش کارایی -		۰/۱۷۹	۰/۰۴۶	<۰/۰۰۱	۱/۱۹۶
دفعات مراجعه به پزشک عمومی و متخصص -		۰/۶۱۳	۰/۰۶۱	<۰/۰۰۱	۱/۸۴۵
بیمه	نداشته‌اند	۱/۹۲۰	۰/۵۸۲	۰/۰۰۱	۶/۸۱۹
	بیمه دولتی	۱/۳۵۶	۰/۵۸۸	۰/۰۲۱	۳/۸۷۹
	بیمه تکمیلی				

جدول ۴. نتایج رگرسیون خطی چندگانه بین میزان هزینه‌های غیر صفر و متغیرهای مستقل

متغیر	برآورد	انحراف معیار	p-value
ثابت	۳۱/۲۰۶	۶/۹۰۲	<۰/۰۰۱
دفعات مراجعه به پزشک عمومی و پزشک متخصص	۱۴/۳۰۹	۲/۲۳۶	<۰/۰۰۱
تعداد روزهای غیبت از کار و کاهش کارایی	۱۲/۴۳۷	۰/۶۸۹	<۰/۰۰۱
بستری شدن در بیمارستان	۴۵۳/۶۷۶	۲۴/۸۲۷	<۰/۰۰۱

جدول ۵. نتایج برآورد هزینه‌ی بیماری‌های دستگاه گوارش

متغیر	میانگین	انحراف معیار	حداقل	حداکثر	R <sup>2</sup>
هزینه‌ی واقعی	۷۸/۳۵	۲۲۲/۳۶	۰	۵۱۸۳/۸۱	-
هزینه‌ی برآورد شده با استفاده از مدل دو بخشی	۷۵/۹۳	۱۲۲/۲۹	۲/۸۹	۱۳۹۴/۳۲	۰/۳۳

پزشک [عمومی یا متخصص]، شانس اینکه فرد هزینه‌ی غیر صفر داشته باشد، ۸۴ درصد افزایش می‌یابد (جدول ۳). با برآزش این مدل، احتمال داشتن هزینه‌ی غیر صفر برای هر فرد به صورت جداگانه محاسبه شد تا در مراحل بعد از آن استفاده شود.

با برآزش رگرسیون خطی میانگین هزینه‌ی بیماری برای هر فرد بدون در نظر گرفتن اینکه فرد هزینه‌ی صفر داشته است به دست آمد. با تلفیق نتایج به دست آمده از دو بخش مدل، میانگین هزینه نهایی برای هر فرد به دست آمد.

### بحث و نتیجه‌گیری

در مطالعه حاضر وضعیت تأهل، وزن افراد بیمار، تعداد بیماری‌های دستگاه گوارش که فرد به آن مبتلا است، تعداد روزهای غیبت از کار یا کاهش کارایی، تعداد دفعات مراجعه بیمار به پزشک عمومی و متخصص و نوع بهره‌مندی از بیمه در داشتن هزینه‌های غیر صفر تأثیر گذار بوده است. از طرفی در قسمت دوم مدل، تعداد روزهای غیبت از کار یا کاهش کارایی داشته است و تعداد دفعات مراجعه بیمار به پزشک عمومی و متخصص و نوع درمان (سرپایی یا بستری در بیمارستان) روی هزینه به طور معنی داری تأثیر گذار بوده‌اند. نتایج بخش دوم مدل نشان می‌دهد، در صورت ثابت بودن سایر شرایط، با هر بار بستری شدن در بیمارستان، بیمار متحمل هزینه‌ای حدود ۴۵۴ دلار بین المللی (۴۰۰-۵۰۰ هزار تومان) می‌شود. هر روز غیبت از کار یا کاهش کارایی، هزینه‌ای در

گذار بر هزینه ارایه شده است. در قسمت پایانی به شرح کامل متغیر اصلی یعنی متغیر هزینه پرداخته شده است: هزینه‌ی هر فرد با استفاده از شاخص‌های اقتصادی و بر اساس دلار بین المللی ثبت شده است. در این مطالعه ۲۱ درصد افراد هزینه‌ی صفر دلار و حدود یک درصد هزینه‌ی بالغ بر ۱۰۰۰ دلار داشته‌اند. میانگین هزینه در نمونه‌ی مورد مطالعه ۷۸/۳۵ دلار با انحراف معیار ۲۲۲/۳۷ دلار بوده است. با برآزش رگرسیون لجستیک مدل نهایی به صورت زیر به دست آمد:

$$\begin{aligned} \text{logit (داشتن هزینه)} &= \beta_1 (\text{تحصیلات}) \\ &+ \beta_2 (\text{وضعیت تاهل}) \\ &+ \beta_3 (\text{تعداد بیماری‌ها}) \\ &+ \beta_4 (\text{تعداد روزها}) \\ &+ \beta_5 (\text{دفعات مراجعه}) \\ &+ \beta_6 (\text{بیمه}) \end{aligned}$$

ولی علاوه بر این متغیرها، علاقه مندی به کنترل اثر متغیرهای سن، جنسیت و میزان توانایی انجام کار باعث شد این متغیرها در مدل باقی بمانند. نتایج نشان می‌دهد شانس این که زنان هزینه‌ی غیر صفر داشته باشند، ۶۴ درصد بیشتر از مردان بوده است. همچنین شانس اینکه افراد بدون بیمه، هزینه‌ی غیر صفر داشته باشند، ۶/۸ برابر افرادی است که بیمه‌ی تکمیلی دارند. به علاوه شانس این که افراد دارای بیمه‌ی دولتی هزینه‌ی غیر صفر داشته باشند، ۳/۹ برابر افرادی است که بیمه‌ی تکمیلی دارند. با هر بار افزایش مراجعه فرد به

بیماری‌های سوء هاضمه و رفلاکس معده به مری را ۲۳۱/۷ دلار به ازای هر فرد در یک سال گزارش کرده است. همچنین هزینه‌ی کل برای بیماری سوء هاضمه به تنهایی ۱۲۰/۲ دلار و برای بیماری رفلاکس معده به مری ۱۱۱/۴ دلار گزارش شد. به علاوه در آن تحقیق هزینه‌ی مستقیم بیماری‌های سوء هاضمه و رفلاکس معده به مری به ترتیب ۱۰۸ و ۹۸ دلار برای هر فرد در یک سال به دست آمد (۹). در پژوهشی دیگر با استفاده از روش‌های اقتصادی برای هزینه‌ی بیماری‌های سوء هاضمه و رفلاکس معده به مری محدوده بین ۱۷۲ تا ۱۷۶ دلار ارائه شد. در این تحقیق سعی شده است که با یک روش آماری، برآوردی از میانگین هزینه بیماری‌های عملکردی دستگاه گوارش ارائه شود (۱۸).

همان‌طور که گفته شد یکی از ویژگی‌های برجسته‌ی هزینه‌های مراقبت بهداشتی، داده‌های پراکنده با بخش عمده‌ای از مقادیر صفر است. هنگام مدل‌سازی این داده‌ها یکی از عواملی که باید در نظر گرفته شود درصد بالایی از مقادیر صفر در مجموعه داده‌ی کم حجم ولی با اطلاعات مهم است. مدل دو بخشی با انعطاف پذیری زیاد بهترین مدل را برای هزینه‌های با مقادیر صفر زیاد و چوله برآزش می‌دهد و برآوردهایی معقول و مناسب به دست می‌آورد (۱۰). این مدل در مطالعات مختلف مورد بررسی قرار گرفته است. دوان در سال ۱۹۸۳ مدل‌های یک بخشی، دو بخشی و چهار بخشی را برای مدل بندی داده‌های هزینه‌های پزشکی در مطالعات متقاطع بررسی کرد و نشان داد مدل‌های چهار بخشی اشکالات موجود در رفتار مدل‌های دو بخشی و یک بخشی را تصحیح می‌کنند (۱۰). اولسن و شفر مدل رگرسیون دو بخشی را برای داده‌های شبه پیوسته با اضافه کردن مؤلفه‌ی تصادفی به قسمت اول و دوم مدل در داده‌های طولی توسعه دادند (۱۹). تحقیقی برآورد کننده‌های متفاوت [حداقل مربعات و مدل‌های خطی تعمیم یافته] و مدل‌های یک بخشی و دو بخشی را از نظر کارایی برای برآورد هزینه‌های پزشکی مقایسه شد و عنوان شد که دقت مدل‌های دو بخشی در صورتی که بیش از ده

حدود ۱۲ دلار بین‌الملل (۱۰-۱۵ هزار تومان) به بیمار تحمیل می‌کند. همچنین بیمار در هر مراجعه به پزشک عمومی یا متخصص در حدود ۱۴ دلار (۱۵-۲۰ هزار تومان) هزینه می‌کند. با توجه به نتایج مدل دو بخشی میانگین هزینه‌ی بیماری‌های دستگاه گوارش ۷۵/۹۳ دلار با انحراف معیار ۱۲۲/۲۹ دلار برآورد شد.

هزینه‌های بهداشت از عوامل مهم تشکیل دهنده رشد اقتصادی‌اند که در کنار نیروی کار و سرمایه‌ی فیزیکی می‌توانند باعث توسعه و پیشرفت اقتصادی کشور شوند. در کشور ایران مخارج بهداشتی اهمیت فراوانی دارند زیرا جمعیت ایران جوان است و حفظ سلامت و تندرستی این نیروی کار همراه با گسترش آموزش عمومی و عالی بهتر می‌تواند به رشد و توسعه اقتصادی کمک کند (۱۳). در ایران مطالعات زیادی در زمینه‌ی هزینه‌ی اقتصادی بیماری‌های مختلف صورت گرفته است ولی تقریباً تمام این مطالعات از روش‌های اقتصادی متعارف برای تعیین هزینه‌های مستقیم و غیر مستقیم یک بیماری خاص بهره برده‌اند و در تمامی این تحقیقات از هیچ مدل آماری برای به دست آوردن میانگین هزینه استفاده نشده است. طی سال‌های متمادی در ایران هزینه‌ی بیماری‌های مختلف از جمله دیابت، نروپاتی دیابتی، سرطان ریه و شکستگی مفصل ران با استفاده از روش‌های اقتصادی مورد بررسی قرار گرفته‌اند (۱۴-۱۷). در زمینه‌ی هزینه‌های اقتصادی بیماری‌های دستگاه گوارش روش‌های هزینه‌ی مستقیم را برای دو گروه با مشاوره و بدون مشاوره در سندرم روده تحریک پذیر به ترتیب برابر ۹۲/۰۴ و ۱/۰۴ دلار، اختلالات ناشناخته‌ی روده‌ای برابر ۱۰۰/۹۴ و ۰/۳۹ دلار، یبوست برابر ۵۷/۲۳ و ۱/۰۴ دلار و نفخ شکم برابر ۷۱/۳۵ و ۰/۶۳ دلار و هزینه‌ی غیر مستقیم را برای دو گروه با مشاوره و بدون مشاوره در سندرم روده تحریک پذیر به ترتیب برابر ۱۱/۸۵ و ۶۶۹/۰۹ دلار، اختلالات ناشناخته‌ی روده‌ای برابر ۷۰۵/۸۵ و ۲۶۷/۴۷ دلار، یبوست برابر ۵۸۷/۴۸ و ۹۷/۴۹ دلار و نفخ شکم برابر ۱۴۷/۸۸ و ۳۸/۶۰ گزارش کرده است (۸۵).

برازش این مدل بر داده‌های موجود به طور خلاصه ارائه شد.

از آن جا که در این پژوهش از اطلاعاتی استفاده شده بود که قبلاً به هدف دیگری جمع آوری شده بودند، بنابراین بسیاری از اطلاعات که می‌توانست در این زمینه راهگشا باشد، موجود نبود. مواردی همچون طول دوره بیماری، نحوه اثرگذاری بیمه، میزان درآمد و محل سکونت از این قبیل اطلاعات بودند.

با توجه به نتایج مدل دو بخشی میانگین هزینه‌ی بیماری‌های دستگاه گوارش ۷۵/۹۳ دلار با انحراف معیار ۱۲۲/۲۹ دلار برآورد شد. میزان حداقل هزینه ۲/۸۹ و حداکثر هزینه ۱۳۹۴/۳۲ دلار به دست آمد. در حالی که میانگین هزینه‌ی واقعی در نمونه‌ی مورد نظر ۷۸/۳۵ دلار با انحراف معیار ۲۲۲/۳۶ دلار بوده است. حداقل و حداکثر هزینه در نمونه‌ی ما صفر و ۵۱۸۳/۸۱ دلار بود. با در نظر گرفتن میانگین واقعی و مقدار مشابه برآورد شده در این روش، دیده می‌شود که این مقادیر بسیار نزدیک به هم هستند. استفاده از مدل دو بخشی باعث کاهش انحراف معیار هزینه‌ها شده است، به علاوه حداقل هزینه‌ی برآورد شده از این روش صفر نیست و ایده‌ی حداکثر ساختن هزینه‌های صفر تحقق یافته است. مقدار  $R^2$  (شاخص نیکویی برازش) مدل دو بخشی برابر ۰/۳۳ است، در نتیجه مدل با توجه به حجم نمونه برازش نسبتاً مناسبی به داده‌ها دارد.

با استفاده از مدل دو بخشی میانگین هزینه‌های بیماری‌های دستگاه گوارش بسیار نزدیک به مقدار واقعی برآورد شد. به علاوه این مدل در حداکثر ساختن هزینه‌های صفر و کاهش انحراف معیار بسیار خوب عمل کرده است. با توجه به این که داده‌های پزشکی چوله و دارای مقادیر صفر زیاد و همچنین مقادیر بزرگ هستند، مدل دو بخشی برآوردهایی مناسب ارائه می‌دهد.

### تقدیر و تشکر

این مقاله بخشی از پایان نامه محدثه شجاعی شاهرخ آبادی در مقطع کارشناسی ارشد رشته آمار زیستی به راهنمایی آقای دکتر انوشیروان کاظم

درصد هزینه‌های افراد نمونه صفر باشد، بیشتر از حالتی است که نسبت صفرها کمتر باشد. هزینه‌ی ۲۱ درصد از اعضای نمونه‌ی موجود در این پایان نامه صفر بود، در نتیجه مدل دو بخشی استفاده شده در این پایان نامه از دقت کافی برخوردار بود. همچنین نشان داده شده است که اگر هیچ کدام از مدل‌های خطی تعمیم یافته با تابع پیوندهای متفاوت برازش خوبی نسبت به داده‌ها نداشته باشند، استفاده از مدل‌های دو بخشی برای بالا بردن انعطاف برآوردگرها مناسب است (۲۰). دیب و برای برآورد هزینه‌های بهداشتی از تحلیل بیزی مدل دو بخشی با متغیرهای درون‌زا استفاده کرد و نشان داد برای متغیرهای درون‌زا این مدل کارآمدتر است (۲۱). علاوه بر این ژو و لیانگ در همان سال مدل رگرسیونی نیمه پارامتری تک شاخصی دو بخشی را برای کم کردن فرضیات روش‌های پارامتری رگرسیونی که برای تحلیل داده‌های چوله به کار می‌روند، ارائه کردند و برآورد کننده‌های سازگار و مجانباً نرمال به دست آوردند (۱). همچنین در پژوهش دیگری هزینه‌های پزشکی ماهانه‌ی ۱۴۵۵ بیمار مبتلا به سکته قلبی را به صورت طولی در نظر گرفته شد و مدل دو بخشی و مدل بقا برای این داده‌ها ترکیب و مدل انعطاف پذیرتری به دست آمد (۲۲). در همین راستا محقق‌ی مدل چهار بخشی با اثرات تصادفی را معرفی نمود و کاربرد آن را در هزینه‌های پزشکی همبسته (اندازه‌های تکراری) نشان داد، همچنین پس از آن، مدل اثرات تصادفی دو بخشی انعطاف پذیر را برای هزینه‌های وابسته برازش داد. مدل ارائه شده در مقاله وی از دو مدل خطی تعمیم یافته‌ی آمیخته تشکیل شده است که توسط اثرات تصادفی وابسته، به هم ربط داده شده‌اند. این مدل علاوه بر کارایی از دقت بالاتری نیز برخوردار است. بر اساس تمامی این مقالات، برآوردهای به دست آمده از مدل‌های دو بخشی از انعطاف پذیری بالایی برخوردار هستند، کارا، سازگار و مجانباً نرمال بوده و دقت بالایی نیز دارند (۲۳و۲). در این مطالعه با استفاده از ویژگی‌های این مدل، میانگین هزینه‌های بیماری دستگاه گوارش برآورد شده است. همچنین نتایج حاصل از

drinking. Health, Econometrics and Data Group. 2006; 6 [12]: 1751-1782.

12. Mullahy J. Much ado about two: reconsidering retransformation and the two-part model in health econometrics. Journal of Health Economics. 1998; 17: 247-281.

13. Hadian M, Shojaei S, Rajabzadeh D. The impact of health expenditure on the economic growth in Iran [1980-2004]. Journal of health administration. 2006; 9[24]:39-44. (Persian).

14. Amini M, Khadivi R, Haghghi S. Cost of type 2 diabetes in Isfahan- Iran in 1998. Iranian Journal of Endocrinology and Metabolism. 2002; 3[2]: 97-104. (Persian).

15. Amini M, Khadivi R. Direct medical cost of diabetic nephropathy in non-insulin dependent diabetic patients in Isfahan- Iran in 1998. Iranian Journal of Endocrinology and Metabolism. 2002; 1[3]: 215-223. (Persian).

16. Zamani B, Ebrahimi S A, Ahmadvand A, Mousavi Gh. The frequency of osteoporosis in HIP fracture following minor trauma and the resulting mortality rate and direct treatment costs in patients over 45 years old in Kashan Naghavi hospital during 2005- 2007. Journal of Kerman University of Medical Sciences. 2010; 17[2]:137-144. (Persian).

17. Akbarzadehbaghban A, Esmaili M, Kimiafar Kh. Medical information management and assessment of direct costs of treatment of Lung cancer. Health Information Management. 2008; 5[2]: 168. (Persian).

18. Lashkajani RM, Roshandel D, Shafae S. A cost analysis of gastroesophageal reflux disease and dyspepsia in Iran. Dig Liver Dis. 2008; 40[6]: 412-417.

19. Olsen MK, Schafer JL. A two-part random effects model for semicontinuous longitudinal data. Journal of the American Statistical Association. 2001; 96: 730-745.

20. Buntin M, Zaslavsky A. Too much ado about two part models and transformation? Comparing methods of modeling medicare expenditures. Journal of Health Economics. 2004; 23: 525-542.

21. Deb P, Munkin M K, Trivedi P K. Bayesian analysis of the two-part model with endogeneity: application to health care expenditure. Journal of Applied Econometrics. 2006; 21: 1081-1099.

22. Liu A, Kronmal R, Zhou X H. Semiparametric two part models with proportionality constraints: analysis of the multi-ethnic study of atherosclerosis [MESA]. University of Washington Biostatistics Working Paper Series. 2008; working paper 341.

23. Liu L, Conoway M R, Knaus W A. A random effects four-part model, with application to correlated medical costs. Computational Statistics and Data Analysis. 2008; 52: 4458-4473.

نژاد و مشاوره آقای دکتر فرید زایری در سال ۱۳۹۰ و کد ۲۰۵۹۰۳۱ می‌باشد که با حمایت مالی دانشکده علوم پزشکی دانشگاه تربیت مدرس اجرا شده است. نگارندگان از کلیه همکاران مرکز تحقیقات بیماری‌های گوارش و کبد دانشگاه علوم پزشکی شهید بهشتی تشکر و قدردانی می‌نمایند.

## منابع

1. Zhou X H, Liang H. Semiparametric single-index two-part regression models. Computational Statistics and Data Analysis. 2006; 50: 1378-1390.

2. Liu L, Strawderman R L, Cowen M E, Shih Y T. A flexible two-part random effects model for correlated medical costs. Journal of Health Economics. 2010; 29: 110-123.

3. Ghadir MR, Ghanouni AR. A review of the treatment of irritable bowel syndrome. Journal of Ghom University of Medical Sciences. 2010; 2[4]: 59-66. (Persian).

4. Roshandel D, Rezailashkajani M, Shafae S, Zali MR. A cost analysis of gastro-oesophageal reflux disease and dyspepsia in Iran. Digestive and Liver Disease. 2008; 40: 412-417.

5. Nyrop K A, Palsson O S, Levy R L, Korff M V, Feld A D, Turner M J, et al. Costs of health care for irritable bowel syndrome, chronic constipation, functional diarrhea and functional abdominal pain. Alimentary Pharmacology & Therapeutics. 2007; 26: 237-248.

6. Schwenkglens M, Marbet U A, Szucs TD. Epidemiology and costs of gastroesophageal reflux disease in Switzerland: a population-based study. Original Article. 2004; 49: 51-61.

7. Hoveyda N, Heneghan C, Mahtani K R, Perera R, Roberts N, Glasziou P. A systematic review and meta-analysis: probiotics in the treatment of irritable bowel syndrome. BMC Gastroenterology. 2009; 9: 15-26.

8. Roshandel D, Rezailashkajani M, Shafae S, Zali M R. A cost analysis of functional bowel disorders in Iran. Int J Colorectal Dis. 2007; 22: 791-799.

9. Moghimi Dehkordi B, Vahedi M, Khoshkrood Mansoori B, Kasaeian A, Safae A, Habibi M, et al. Economic burden of gastro-oesophageal reflux disease and dyspepsia: A community-based study. Arab Journal of Gastroenterology. 2011; 3[5]: 37-41.

10. Doan N, Manning W G, Morris C N, Newhouse J P. A comparison of alternative models for the demand for medical care. Journal of Business and Economic Statistics. 1983; 1: 115-126.

11. Madden D. Sample selection versus two-part models revisited: the case of female smoking and

## Estimating the economic costs of functional gastrointestinal tract disease using regression two-part parameters model

**Mohadeseh Shojaei**, MSc of Biostatistics, Biostatistics Department, Faculty of Medical Sciences, Tarbiat Modarres University, Tehran, Iran. [m\\_shojai@modares.ac.ir](mailto:m_shojai@modares.ac.ir)

\***Anoshirvan Kazemnejad**, PhD. Professor of Biostatistics, Biostatistics Department, Faculty of Medical Sciences, Tarbiat Modarres University, Tehran, Iran. (\*Corresponding author) [kazem\\_an@modares.ac.ir](mailto:kazem_an@modares.ac.ir)

**Farid Zayeri**, PhD. Assistant Professor of Biostatistics, Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [fzayeri@yahoo.com](mailto:fzayeri@yahoo.com)

**Mohsen Vahedi**, MSc of Biostatistics, Research Center for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [mohsenvahedi540@gmail.com](mailto:mohsenvahedi540@gmail.com)

### Abstract

**Background:** Accurate estimation of medical costs is one of the health care policymakers' goals. Regarding to the feature of health data and the complexity of their analysis conventional models are not suitable for them. To model these data, all these characteristics should be taken into consideration. This study has estimated the average cost of gastrointestinal tract diseases and some factors influencing this cost using two part regression model.

**Methods:** During 2006 to 2007 information of total 1907 gastrointestinal tract patients were collected in the research center for gastroenterology and liver disease of Shahid Beheshti University of medical science of Tehran as a retrospective cross-sectional study. For the purpose of modeling gastrointestinal tract disease cost, the two-part model was employed. In the first part, a logistic regression was fitted to the dichotomous events of having zero or positive expense and in the second part a multiple linear regression was fitted to positive expense and the effect of demographic variables, work ability, number of times visiting a doctor, the number of diagnostic tests, insurance, number of days absent from work or reduced efficiency and hospitalization on expense were assessed.

**Results:** The average costs of gastrointestinal tract diseases and their standard deviations in parametric method were yielded as  $\$75.93 \pm 122.29$ . Minimum and maximum of costs were respectively estimated  $\$2.89$  and  $\$1394.32$ . While the actual average cost was  $\$78.35 \pm 222.36$ , minimum and maximum were respectively zero and  $\$5183.81$ . The results obtained from the two- part model revealed that "some demographical variables", "number of days absent from work or reduced efficiency" and "number of times visiting a doctor" have influenced having expense and the variables "number of days absent from work or reduced efficiency", "number of times visiting a doctor" and " hospitalization" " have influenced positive expenses.

**Discussion:** The estimated cost made by this method is close to the true value. According to the lower standard deviation of estimations made by model, compared to the actual values, estimations from this model have high precision. In addition, this model has performed well to maximize zero costs.

**Keywords:** Two-part model, Economic costs of disease, Two-part parametric regression, Gastrointestinal tract disease.