

## خود-ارزیابی سلامت در نوجوانان: کاربرد روش جانهی چند متغیره رگرسیون سلسله مراتبی

سید محسن حسینی - محبوبه مقامی

گروه آمار زیستی و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی اصفهان

**چکیده:** هدف این مطالعه کاربرد روش جانهی رگرسیون سلسله مراتبی است. در این روش، جانهی به وسیله برازش دنباله ای از مدل های رگرسیونی و به دست آوردن مقادیر از توزیع های پیشگویانه صورت می گیرد. نوع مدل های رگرسیونی که استفاده می شود، با توجه به نوع متغیرهایی که جانهی می شوند می تواند خطی، لجستیک، پواسن، لجیت تعمیم یافته یا ترکیبی از آنها باشد. اطلاعات مربوط به خود-ارزیابی سلامت که شامل داده های گمشده می باشد با استفاده از این روش، جانهی و سپس با استفاده از رگرسیون چندگانه عوامل مؤثر آن بررسی می شوند. و سپس بعد از کنار گذاشتن داده های گمشده با رگرسیون چندگانه برازش داده می شود و با نتایج حاصل از روش جانهی مقایسه می شود. در این مطالعه از اطلاعات طرح نظام مراقبت رفتارهای مرتبط با سلامت و عوامل خطرزای بیماری ها در دانش آموزان استفاده شده است. نوع این مطالعه مقطعی و حجم نمونه برابر با ۳۸۴۳ است.

**واژه های کلیدی:** داده های گمشده، جانهی چندگانه، جانهی چند متغیره رگرسیون سلسله مراتبی، خود-ارزیابی سلامت، نوجوانان

### ۱ مقدمه

آماردان ها و تحلیل گران داده ها اغلب با مسئله ی تحلیل داده های گمشده در مطالعات و پرسش نامه ها مواجه می شوند. ( فینچ ۲۰۱۰ ) محققان آمار یا علوم اجتماعی اثر این قبیل داده های گمشده را روی تحلیل های آماری در مطالعات بررسی کرده اند. ( دی لیو، هاکس و همکاران ۲۰۰۳ ) این مطالعات نشان می دهد که داده های گمشده می توانند اثرات زیان آوری روی نتایج تحلیل های آماری داشته باشد، که منجر به برآورد پارامترهای اریب و خطاهای استاندارد متورم شود. به علاوه وقتی که حجم داده های گمشده زیاد باشد، توان آزمون های آماری کاهش می یابد. ( دی لیو، هاکس و همکاران ۲۰۰۳ )

با توجه به مسائل زیادی که با وجود داده های گمشده به وجود می آید، چندین روش برای جانهی داده های گمشده پیشنهاد شده است. (شفه و گراهام ۲۰۰۲)

در بین این روش ها جانشینی میانگین است که متوسط مقادیر نمونه برای مشاهدات گمشده از یک متغیر خاص جایگذاری می شود. جانهی رگرسیون که مشاهدات گمشده با استفاده از پیشگویی که از تحلیل رگرسیون چندگانه به دست می آید جانهی می شوند و روش بی درنگ<sup>۱</sup> که مقادیر گمشده با مقدار مشاهده شده ی مطابقت داده شده بر اساس متغیرهایی که گمشده ندارند، جایگذاری می شود. هریک از این روش های جانهی منفرد در تولید پارامترهای معلوم جامعه و خطاهای استاندارد نامناسب می باشند. (شفه و گراهام ۲۰۰۲) با توجه به مشکلاتی که این روش ها دارند، محققان در جستجوی روشهایی بوده اند که برای جایگذاری مشاهدات گمشده مناسب تر باشند. جانهی چندگانه<sup>۲</sup> یکی از این روش هاست. در واقع تعداد زیادی از محققان پیشنهاد کرده اند که برای داده هایی که دارای توزیع نرمال پیوسته هستند این روش بهینه می باشد. (شفه و اولسن ۱۹۹۸) این روش که توسط روین پیشنهاد شده است، (روین ۱۹۹۶) شامل جانهی چندین مجموعه ی ممکن برای مقادیر گمشده می باشد که در نتیجه ی آن چندین مجموعه داده ی کامل به دست می آید. هر مجموعه داده به صورت جداگانه به وسیله ی یک مدل رگرسیونی خاص تحلیل می شود. سپس نتایج به دست آمده (برآوردهای نقطه ای و ماتریس های کواریانس) با استفاده از فرمولی که در روین آمده است (a) ۱۹۸۷ ترکیب می شود. (راگونتان، لیکواسکی و همکاران. ۲۰۰۱)

در مطالعات مجموعه داده ها اغلب شامل تعداد زیادی از متغیرها که شکل های توزیعی مختلفی دارند، هستند. برای مثال این قبیل مجموعه داده ها صدها متغیر دارند که پیوسته، شمارشی، دوتایی یا چندتایی و یا حتی نیمه پیوسته یا متغیرهایی با محدودیت هایی هستند. به علاوه توزیع متغیرهای پیوسته به تنهایی ممکن است نرمال، لگ نرمال یا توزیع های دیگر باشد.

به علاوه داده های پیمایشی اغلب دو خصیصه ی دیگر هم دارند که فرآیند مدل بندی را پیچیده تر می کند. اول اینکه، وجود محدودیت های خاص ضروری است. برای مثال متغیر "سن آغاز سیگار کشیدن" تنها برای کسانی است که سابقه ی سیگار کشیدن دارند. بنابراین باید فرآیند جانهی این متغیر، برای سیگاری ها محدود شود.

دوم اینکه محدوده های منطقی یا توافقی خاصی برای مقادیر گمشده ای که باید در فرآیند جانهی شوند، وجود دارد. این قبیل روابط متقابل بین متغیرها تشخیص

<sup>۱</sup> Hot Deck Method

<sup>۲</sup> Multiple imputation

مدل را پیچیده می کند. برای مثال متغیر "تعداد سالهای سیگار کشیدن" که محدود به سیگاری ها می شود باید کمتر از مقدار سن باشد. (راگونتان، لپکواسکی و همکاران. ۲۰۰۱)

هدف این مطالعه کاربرد یک روش جانهای چندگانه است که می توان برای یک مجموعه داده با ساختار پیچیده به کار برد. استراتژی آن به این صورت است که جانهای از طریق دنباله ی رگرسیون های چندگانه انجام می شود. نوع مدل رگرسیون بستگی به نوع متغیرهایی دارد که جانهای می شوند. فرض می شود که متغیرها در مجموعه داده ها از پنج نوع باشند، (۱) پیوسته، (۲) دوتایی، (۳) رسته ای (بیشتر از دو رسته)، (۴) شمارشی، (۵) آمیخته (یک متغیر پیوسته که قسمت عمده ی آن صفر است). همینطور فرض می کنیم که جمعیت تحت مطالعه نامحدود، نمونه یک نمونه تصادفی ساده و مکانیسم گمشدگی داده قابل چشم پوشی است (راگونتان، لپکواسکی و همکاران. ۲۰۰۱).

برای استفاده از این روش اطلاعات مربوط به طرح کشوری نظام مراقبت رفتارهای مرتبط با سلامت و عوامل خطرزای بیماری ها در دانش آموزان به کار برده شده است. هدف این مطالعه بررسی عوامل مرتبط با خود-ارزیابی سلامت در دانش آموزان مقطع راهنمایی و دبیرستان می باشد. خود-ارزیابی سلامت<sup>۳</sup> یک اندازه گیری معتبر از وضعیت سلامت است که به طور وسیعی در مطالعات مربوط به اپیدمیولوژی سلامت و سلامت اجتماعی جامعه استفاده می شود (لاندبرگ و ماندرباکا ۱۹۹۶؛ کاکس، اوین و همکاران ۲۰۰۹).

در اپیدمیولوژی اجتماعی خود-ارزیابی سلامت به عنوان یک ابزار اندازه گیری برای پیش بینی ناتوانی ها، ناخوشی ها و مرگ و میر می باشد. هم چنین وقتی که با ارزیابی های عینی سلامت همراه شده باشد، پایین بودن خود-ارزیابی سلامت پیش بینی کننده ی افزایش مرگ و میر می باشد. (ایدلر و بنیامینی ۱۹۹۷) اگرچه بعضی شکایات در رابطه با این اندازه گیری فردی و ذهنی وجود دارد (لنگ و دلپیر ۲۰۰۹). به نظر می رسد که درک از سلامتی فردی ابعاد زیستی، اجتماعی-اقتصادی و روانشناسی را با هم ترکیب می کند و یک اندازه ی خام و ساده را می دهد که فرای روشهای ارزیابی عینی می باشد (یولها ۲۰۰۹). در مقایسه با سلامتی از دیدگاه پزشکی، که اساسا مربوط به دلایل و علائم امراض می باشد، ارزیابی از خود شامل اجزای فیزیکی و احساسی است که کاملا همانند زندگی است. به علاوه بیان احساس بیماری غالبا مربوط به مسائل اجتماعی و روانی است تا داشتن درد یا ناراحتی های فیزیکی (بلنک و دیدرسن ۱۹۹۶).

<sup>۳</sup> Self-rated health

در این مطالعه برای بررسی عوامل مرتبط با خود-ارزیابی سلامت، پس از جانمایی داده های گمشده با استفاده از روش جانمایی چند متغیره رگرسیون سلسله مراتبی، از رگرسیون چندگانه استفاده شده است.

## ۲ روش جانمایی چند متغیره رگرسیون سلسله مراتبی

برای یک نمونه با اندازه  $n$ ، یک ماتریس پیش بینی با بعد  $n \times p$  شامل همه ی متغیرهایی است که هیچ داده ی گمشده ای ندارند.  $X$  شامل متغیرهای پیوسته، رسته ای دوتایی، شمارشی یا آمیخته می باشد.  $Y_1, Y_2, \dots, Y_k$  بیان کننده ی  $k$  متغیر هستند که مقادیر آنها گمشده است که بدون از دست دادن کلیت آنها، براساس تعداد گمشدگی از کمترین به بزرگترین مرتب شده اند.

برای یک مدل وابسته به جانمایی چگالی شرطی توأم  $Y_1, Y_2, \dots, Y_k$  به شرط  $X$  به صورت زیر تفکیک می شود:

$$f(Y_1, Y_2, \dots, Y_k | X, \theta_1, \theta_2, \dots, \theta_k) = \\ f_1(Y_1 | X, \theta_1) f_2(Y_2 | X, Y_1, \theta_2) \dots \\ f_k(Y_k | X, Y_1, Y_2, \dots, Y_{k-1}, \theta_k)$$

به طوریکه  $f_j, j = 1, 2, \dots, k$  تابع چگالی شرطی است و  $\theta_j$  برداری از پارامترها در توزیع شرطی می باشد. هر یک از چگالی های شرطی را از طریق یک مدل رگرسیونی مناسب با پارامترهای نامعلوم  $\theta_j$  مدل بندی می کنیم و توزیع های پیشگویانه ی مقادیر گمشده را با داشتن مقادیر مشاهده شده دریافت می کنیم. فرض می کنیم که توزیع پیشین برای پارامترهای  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$  برابر است با  $\pi(\theta) \propto 1$  به هر حال روش را می توان برای توزیع های پیشین مناسب اصلاح کرد.

هر رگرسیون شرطی بر اساس یکی از مدل های زیر می باشد:

(۱) یک مدل رگرسیون خطی نرمال بر اساس مقیاس های مناسب (برای مثال تبدیل توانی باکس-کاکس را می توان برای نرمال کردن به دست آورد)، اگر  $Y_j$  پیوسته باشد.

(۲) یک مدل لجستیک، اگر  $Y_j$  دوتایی باشد.

(۳) یک مدل رگرسیون لجیت تعمیم یافته، اگر  $Y_j$  رسته ای باشد.

(۴) یک مدل لگاریتم خطی پواسن، اگر  $Y_j$  متغیر شمارشی باشد.

۵) یک مدل دو مرحله‌ای به طوریکه حالت صفر- غیر صفر با استفاده از رگرسیون لجستیک که شرطی روی حالت غیر صفر می‌شود، جانهی شده است و مدل رگرسیون خطی نرمال برای جانهی حالت غیر صفر استفاده شده است، وقتی که  $Y_j$  یک متغیر آمیخته می‌باشد.

هر جانهی شامل  $c$  دوره می‌باشد. دوره یک با رگرسیون کردن متغیر با کمترین مقدار گمشدگی،  $Y_1$  روی  $X$ ، مقادیر گمشده را تحت یک مدل رگرسیونی مناسب جانهی می‌کند. سپس  $X$  با اضافه شدن  $Y_1$  جدید می‌شود و به متغیر  $Y_2$  که در کم بودن مقدار گمشدگی در جایگاه بعدی قرار دارد، حرکت داده می‌شود. فرایند جانهی با جدید کردن  $X$  به عنوان متغیرهای پیش بینی تا وقتی که همه ی متغیرها جانهی شوند، تکرار می‌شود. یعنی،  $Y_1$  روی  $Y_2, U = X$ ، روی  $(X, Y_1) U = (Y_1)$  دارای مقادیر جانهی می‌باشد،  $Y_2$  روی  $(X, Y_1, Y_2) U = (X, Y_1, Y_2)$  و  $Y_2$  دارای مقادیر جانهی می‌باشند) و به همین ترتیب الی آخر برازش می‌شوند. سپس فرایند جانهی در دوره ی دوم تا  $c$  تکرار می‌شود، با این اصلاح که مجموعه ی پیش بینی کننده ها شامل همه ی متغیرهای  $Y$  می‌باشد، به جز متغیری که به عنوان متغیر وابسته است. بنابراین  $Y_1$  روی  $X$  و  $Y_2, Y_3, \dots, Y_k$  روی  $X$  و  $Y_1, Y_2, \dots, Y_k$  و به همین ترتیب الی آخر رگرسیون می‌شود. تکرار دوره ها تا تعدادی از پیش تعیین شده یا تا وقتی که مقادیر جانهی شده پایا شوند، ادامه می‌یابد. (راگوتان، لپکواسکی و همکاران ۲۰۰۱)

### ۳ روش جمع آوری اطلاعات

اطلاعات مورد نیاز در این مطالعه از طرح نظام مراقبت رفتارهای مرتبط با سلامت و عوامل خطرزای بیماری‌ها در دانش آموزان می‌باشد.

این مطالعه ی مقطعی در سطح ملی توسط وزارت بهداشت درمان و آموزش پزشکی و با همکاری وزارت آموزش و پرورش و حمایت سازمان جهانی بهداشت انجام شد. این مطالعه یک مطالعه ی چند مرکزی است که در بین دانش آموزان در سنین ۶-۱۸ سال که در مناطق شهری و روستایی در شهرهای مرکزی استان های کشور زندگی می‌کنند، انجام می‌شود.

جامعه ی آماری کلیه دانش آموزان دو مقطع راهنمایی و دبیرستان و واحدهای آماری، مدارس دو مقطع راهنمایی و دبیرستان مناطق شهری و روستایی شهرستان‌های مرکزی استان های کشور می‌باشد. به دلیل اینکه دانش آموزان مقطع راهنمایی و دبیرستان در سن نوجوانی هستند و بهتر می‌توانند سلامتی خود را ارزیابی کنند، در این مطالعه

فقط از اطلاعات مربوط به مقطع راهنمایی و دبیرستان استفاده می شود.

با در نظر گرفتن حداقل ۲٪ تغییر در شاخص های اصلی مورد مطالعه و با در نظر گرفتن خطای آلفای ۵٪ و خطای بتای ۲۰٪ و با توجه به وجود شش گروه سنی و جنسی، حجم نمونه ۴۹۵۰ دانش آموز محاسبه گردید که بایست به صورت تصادفی از جامعه انتخاب شوند. همچنین با عنایت به پیش بینی میزان ۸۰٪ پاسخ دانش آموزان به سوالات میزان ۲۰٪ به حجم نمونه اضافه می شود که بدین ترتیب حجم نمونه نهایی حدود ۶۰۰۰ نفر خواهد بود، که با حذف مقطع ابتدایی حدود ۴۰۰۰ نفر می شود. معیار ورود دانش آموزان به بررسی رضایت از سوی ایشان و نیز رضایت والدین / ولی قانونی دانش آموز است.

پرسشنامه ها به زبان فارسی است که بر اساس پرسشنامه های استفاده شده در پروژه های روش گام به گام سازمان بهداشت جهانی برای بیماری های غیر واگیردار و بررسی سلامت وابسته به مدرسه در دانش آموزان طراحی شده است. (کلیشادی، اردلان و همکاران ۲۰۰۷؛ کلیشادی، اردلان و همکاران ۲۰۰۸؛ مطلق، کلیشادی و همکاران ۲۰۰۹)

#### ۴ روش جهانی داده های گمشده و تحلیل اطلاعات

با توجه به اینکه بعضی از سؤالات توسط دانش آموزان پاسخ داده نشده است، بنابراین ابتدا گمشده های مربوط به سؤالات پرسشنامه جهانی می شوند و سپس برای به دست آوردن متغیرها، سؤالات با هم جمع می شوند.

از آنجایی که پرسشنامه ساختار پیچیده ای دارد، یعنی شامل متغیرهای پیوسته، گسسته و شمارشی می باشد و همچنین در بعضی از قسمت ها باید محدودیت ها و حدودی را در نظر گرفت، بنابراین از روش (SRMI) برای جهانی استفاده می شود.

برای جهانی از ابزار IMPUTE مربوط به بسته ی نرم افزاری IVEware که وابسته به نرم افزار SAS می باشد استفاده شده است.

تولید ۵-۱۰ مجموعه داده ی کامل کافی می باشد (شفه ۱۹۹۷). در این مطالعه ۵ مجموعه تولید شده است.

متغیر خود-ارزیابی سلامت به عنوان متغیر وابسته و متغیرهای مشخصات خانوادگی، فعالیت بدنی، فعالیت های اوقات فراغت، تغذیه و مصرف سیگار متغیرهای مستقل هستند. از تحلیل رگرسیون چندگانه برای تعیین عوامل مرتبط با خود-ارزیابی سلامت

در هریک از مجموعه داده ها به طور جداگانه استفاده می شود. نتایج به دست آمده از ۵ مجموعه داده با استفاده از PROC MIANALYZE در نرم افزار SAS ترکیب می شود. در مرحله ی دیگر ابتدا داده های گمشده کنار گذاشته می شوند و با استفاده از رگرسیون چندگانه عوامل مرتبط با خود- ارزیابی سلامت بررسی می شوند. سپس با نتایج حاصل از رگرسیون جانهی داده های گمشده مقایسه می شوند.

## ۵ یافته ها

حجم نمونه ی مورد استفاده در این مطالعه برابر با ۳۸۴۳ دانش آموز می باشد. در این مطالعه کمترین مقدار خود-ارزیابی سلامت برابر با ۱۶ و بیشترین آن ۶۲ می باشد. جدول ۲ لیست متغیرها همراه با میانگین و انحراف معیار و درصد گمشدگی هر متغیر را نشان می دهد.

جدول ۱:

نام متغیر	تعداد مشاهدات	تعداد گمشده ها	درصد گمشده ها	میانگین	انحراف معیار
مشخصات خانوادگی	۳۱۱۱	۷۳۲	۱۹/۰٪	۲۲/۷۰	۳/۶۸
فعالیت بدنی	۳۲۲۹	۶۱۴	۱۶/۰٪	۸/۱۹	۲/۶۳
فعالیت های اوقات فراغت	۳۱۶۹	۶۷۴	۱۷/۵٪	۳۶/۰۱	۵/۷۵
خود - ارزیابی سلامت	۲۷۱۰	۱۱۳۳	۲۹/۵٪	۴۷/۳۳	۸/۴۲۳
مصرف دخانیات	۲۶۸۳	۱۱۶۰	۳۰/۲٪	۴/۱۶	۵/۹۴
تغذیه	۲۶۸۹	۱۱۵۴	۳۰/۰٪	۴۱/۰۵	۵/۵۷

جدول ۲ مربوط به اطلاعات واریانس جانهی چندگانه می باشد. در این جدول واریانس درون- جانهی (میانگین حسابی از واریانس های  $m = 5$  مجموعه داده) و میان-جانهی (بیان کننده ی تغییرپذیری برآورد پارامترها در بین هریک از  $m = 5$  مجموعه داده) و واریانس کل (مجموع این دو واریانس) را برای ترکیب نتایج مجموعه داده های کامل نشان می دهد. هم چنین درجه آزادی واریانس کل، افزایش نسبی در واریانس ناشی از مقادیر گمشده و کسر اطلاعات گمشده که نشان دهنده ی نسبتی از واریانس کل که ناشی از داده های گمشده می باشد، نشان داده شده است.

جدول ۲: اطلاعات واریانس جانهی چندگانه

پارامتر	میان جانهی درون جانهی		کل	درجه آزادی افزایش کسر نسبی اطلاعات واریانس گمشده	
	۱	۲		۱	۲
عرض از مبدأ	۰/۱۴	۱/۷۰	۱/۸۷	۴۷۱/۶۷	۰/۱۰
مشخصات خانوادگی	۰/۰۰۰۱	۰/۰۰۱	۰/۰۰۱	۴۳۲/۵۷	۰/۱۰
تغذیه	۰/۰۰	۰/۰۰	۰/۰۰	۳۰۲/۴۷	۰/۱۳
فعالیت بدن	۰/۰۰	۰/۰۰۲	۰/۰۰۲	۲۰۹/۴۵	۰/۱۶
اوقات فراغت	۰/۰۰	۰/۰۰	۰/۰۰	۲۰۷/۳۶	۰/۱۶
مصرف دخانیات	۰/۰۰	۰/۰۰	۰/۰۰	۳۴/۷۴۵	۰/۵۱

جدول ۳ برآوردهای ترکیب شده، خطاهای استاندارد برای هر ضریب رگرسیونی، فاصله اطمینان ۹۵٪، آماره  $t$  همراه با  $p$ -value و کمترین و بیشترین برآوردهای پارامترها از مجموعه داده های جانهی شده را نشان می دهد.

همانطور که جدول (۳) نشان می دهد، متغیر مشخصات خانوادگی معنی دار نشده است ( $P - value > 0.05$ ). متغیرهای تغذیه، مصرف دخانیات، فعالیت بدنی و اوقات فراغت روی خود-ارزیابی سلامت تأثیر گذار می باشند.

جدول (۴) مربوط به برآوردهای ناشی از رگرسیون چندگانه بعد از کنار گذاشتن داده های گمشده می باشد. این جدول نشان می دهد که متغیرهای تغذیه، مصرف دخانیات، فعالیت بدنی و مشخصات خانوادگی تأثیر گذار می باشد و متغیر اوقات فراغت اثر گذار نمی باشد. که در مورد دو متغیر مشخصات خانوادگی و اوقات فراغت با نتایج حاصل از رگرسیون جانهی متفاوت می باشد.



جدول ۳: برآورد پارامترهای جان‌هی چندگانه

پارامتر	برآورد	خطای استاندارد	حدود ۹۵٪	کمینه	بیشینه	آماره t	p-value
عرض از مبدأ	۲۶/۳۳	۱/۳۷	۲۳/۶۳	۲۹/۰۲	۲۵/۷۶	۱۹/۲۳	۰/۰۰ >
مشخصات خانوادگی	۰/۰۵۸	۰/۰۴	-۰/۰۱	۰/۱۳	۰/۰۵	۱/۶۳	۰/۱۰
تغذیه	۰/۲۲	۰/۰۲	۰/۱۷	۰/۲۷	۰/۲۱	۸/۸۶	۰/۰۰ >
فعالیت بدن	۰/۸۳	۰/۰۵	۰/۷۲	۰/۹۴	۰/۸۰	۱۵/۲۲	۰/۰۰ >
اوقات فراغت	۰/۱۰	۰/۰۲	۰/۰۵	۰/۱۵	۰/۰۹	۴/۳۲	۰/۰۰ >
مصرف دخانیات	-۰/۳۳	۰/۰۳	-۰/۳۸	-۰/۲۷	-۰/۳۵	-۱۲/۲۹	۰/۰۰ >

## بحث و نتیجه‌گیری

استفاده از روش SRMI برای جان‌هی داده‌های گمشده وقتی که تشخیص توزیع توأم تمامی متغیرها شامل داده‌های گمشده مشکل است، بسیار مفید می‌باشد. ارزیابی سلامتی که خلاصه‌ای از درک ذهنی از سلامتی می‌باشد نسبتاً دارای نتایج یکسانی در مطالعات مختلف در مورد نوجوانان است (بردمن ۲۰۰۶). در این مطالعه متغیر مصرف دخانیات روی خود-ارزیابی سلامت اثر معکوس دارد که با نتایج مطالعاتی که در کانادا و ایالت متحده صورت گرفته است همخوانی دارد (وید و وینجلیز ۱۹۹۹؛ وینجلیز، وید و همکاران ۲۰۰۲). متغیرهای فعالیت بدنی و نحوه‌ی گذراندن اوقات فراغت تأثیر مثبتی روی خود ارزیابی سلامت دارند که با مطالعه‌ی سوئد که در همین رابطه صورت گرفته است، مشابه است (الیندر، ساندبلوم و همکاران ۲۰۱۱).

پس از کنار گذاشتن داده‌های گمشده متغیر اوقات فراغت با استفاده از رگرسیون چندگانه روی خود-ارزیابی سلامت تأثیر گذار نیست. مطالعه‌ی طولی در دانمارک نشان می‌دهد که کیفیت پایین تغذیه با خود-ارزیابی سلامت در سطح پایین رابطه‌ی مستقیم دارد که با نتیجه‌ی مطالعه‌ی حاضر یکسان است (اسلر، هیتمن و همکاران ۲۰۰۱).

در مورد متغیر مشخصات خانوادگی مطالعاتی که در نروژ و ایالت متحده انجام شده است، این متغیر روی خود-ارزیابی سلامت اثر مستقیمی دارد. (وید، پوالین و

جدول ۴: برآورد پارامترهای رگرسیون چندگانه بعد از کنار گذاشتن داده های گمشده

پارامتر	برآورد	خطای آماره $t$ استاندارد	$p - value$
عرض از مبدأ	۳۰/۷۵	۲/۲۴	۰/۰۰ >
مشخصات خانوادگی	۰/۱۳	۰/۰۶	۰/۰۲
تغذیه	۰/۲۵	۰/۰۴	۰/۰۰ >
فعالیت بدن	۰/۴۶	۰/۰۸	۰/۰۰ >
اوقات فراغت	۰/۰۲	۰/۰۴	۰/۵۶
مصرف دخانیات	-۰/۳۳	۰/۰۴	۰/۰۰ >

همکاران ۲۰۰۰؛ بریدابلیک، ملند و همکاران (۲۰۰۸) ولی رابطه بین این متغیر و ارزیابی از سلامت به صورت های مختلف مطالعه شده است و نتایج متناقضی به دست آمده است. (وینجلیز، وید و همکاران ۲۰۰۲) در مطالعه ی حاضر این متغیر با استفاده از رگرسیون جهانی معنی دار نشده است، ولی با استفاده از رگرسیون چندگانه پس از کنار گذاشتن داده های گمشده معنی دار شده است.

## مراجع

- BLANK, N. & F. DIDERICHSEN 1996. The prediction of different experiences of longterm illness: a longitudinal approach in Sweden, *Journal of epidemiology and community health*, **50**, 156-161.
- BOARDMAN, J. D. 2006. Self-rated health among US adolescents. *Journal of Adolescent Health*, **38**, 401-408.
- BREIDABLIK, H. J., MELAND, E. & LYDERSEN, S. 2008. Self-rated health in adolescence: A multifactorial composite. *Scandinavian journal of public health*, **36**, 12-20.
- COX, B., OYEN, H. V., CAMBOIS, E., JAGGER, C., ROY, S., ROBINE, J. M. & ROMIEU, I 2009. The reliability of the Minimum European Health

- Module. *International journal of public health*, **54**, 55-60.
- DE LEEUW, E. D., HOX, J. J. & HUISMAN, M. 2003. Prevention and treatment of item nonresponse. *Journal of Official Statistics*, **19**, 153-176.
- ELINDER, L. S., SUNDBLOM, E. & ROSENDAHL, K.I 2011. Low Physical Activity Is a Predictor of Thinness and Low Self-rated Health: Gender Differences in a Swedish Cohort. *Journal of Adolescent Health*, **48**, 481-486.
- FINCH, W. H. 2010. Imputation Methods for Missing Categorical Questionnaire Data: A Comparison of Approaches. *Journal of Data Science*, **8**, 361-378.
- IDLER, E. L. & BENYAMINI, Y. 1997. Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, 21-37.
- Jylhä, M. 2009. What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social Science & Medicine*, **69**, 307-316.
- KELISHADI, R., ARDALAN, G., GHEIRATMAND, R., GOUYA, M. M., RAZAGHI, E. M., DELAVARI, A., MAJDZADEH, R., HESHMAT, R., MOTAGHIAN, M. & BAREKATI, H. 2007. Association of physical activity and dietary behaviours in relation to the body mass index in a national sample of Iranian children and adolescents: CASPIAN Study. *Bulletin of the World Health Organization*, **85**, 19-26.
- KELISHADI, R., ARDALAN, G., GHEIRATMAND, R., MAJDZADEH, R., HOSSEINI, M., GOUYA, M., RAZAGHI, E., DELAVARI, A., MOTAGHIAN, M. & BAREKATI, H. (2008). Thinness, overweight and obesity in a national sample of Iranian children and adolescents: CASPIAN Study. *Child: Care, Health and Development*, **34**, 44-54.
- LANG, T. & DELPIERRE, C. 2009. 'How are you?': what do you mean? *The European Journal of Public Health*, **19**, 353-353.
- LUNDBERG, O. & MANDERBACKA, K. 1996. Assessing reliability of a measure of self-rated health. *Scandinavian journal of public health*, **24**, 218-224.

- MOTLAGH, M., KELISHADI, R., ARDALAN, G., GHEIRATMAND, R., MAJZADEH, R. & HEIDARZADEH, A. 2009. Rationale, methods and first results of the Iranian national programme for prevention of chronic diseases from childhood : CASPIAN Study. *East. Mediterr. Health J*, **15**, 302-314.
- OSLER, M., HEITMANN, B., HØIDRUP, S , JØRGENSEN, L. M. & SCHROLL, M. 2001. Food intake pattern, self-rated health and mortality in Danish men and women. A progressiveobservational study. *J Epidemiol Community Health*, **55**, 399-403.
- RAGHUNATHAN, T. E., LEPKOWSKI, J. M., VAN HOEWYK, J. & SOLENBERGER, P. 2001. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey methodology*, **27**, 85-95.
- RUBIN, D. B. 1987. *Multiple imputation for nonresponse in surveys*. New York, USA: Willey Online Library.
- RUBIN, D. B. 1996. Multiple imputation after 18+ years. *Journal of the American Statistical Association*, 473-489.
- SCHAFFER, J. L. 1997. *Analysis of incomplete multivariate data*, Chapman & Hall/CRC.
- SCHAFFER, J. L. & GRAHAM, J. W. 2002. Missing data: Our view of the state of the art. *Psychological methods*, **7**, 147.
- SCHAFFER, J. L. & OLSEN, M. K. 1998. Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research*, **33**, 545-571.
- VINGILIS, E. R., WADE, T. J. & SEELEY, J. S. 2002. Predictors of adolescent self-rated health: analysis of the National Population Health Survey. *Canadian Journal of Public Health*, **93**, 193-197.
- WADE, T. & VINGILIS, E. 1999 . The development of self-rated health during adolescence: an exploration of inter-and intra-cohort effects. *Canadian journal of public health. Revue canadienne de sante publique*, **90(2)**, 90.
- WADE, T. J., D. J. PEVALIN & VINGILIS, E. 2000 . Revisiting student self-rated physical health. *Journal of adolescence*, **23**, 785-791.