مجله‌ی غدد درون‌ریز و متابولیسم ایران
دانشگاه علوم پزشکی و خدمات بهداشتی ـ درمانی شهید بهشتی
دوره‌ی پانزدهم، شماره‌ی ۱، صفحه‌های ۴۰ ـ ۳۳ (اردیبهشت ۱۳۹۲)

مقاله‌ی پژوهشی

کاربرد رگرسیون چندک در تعیین عوامل مرتبط با ید دفعی ساکنین شهر تهران

الهه بیگانه[۱]، دکتر یداله محرابی[۲]، دکتر پروین میرمیران[۳]، دکتر علی اکبر خادم معبودی[۱]، پانته‌آ ناظری[۳]

۱) گروه آمار زیستی، دانشکده‌ی پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، ۲) گروه اپیدمیولوژی، دانشکده‌ی بهداشت، دانشگاه علوم پزشکی شهید بهشتی، ۳) مرکز تحقیقات پیشگیری و درمان چاقی، پژوهشکده‌ی علوم غدد درون‌ریز و متابولیسم، دانشگاه علوم پزشکی شهید بهشتی، **نشانی مکاتبه‌ی نویسنده‌ی مسئول:** تهران، بزرگراه شهید چمران، ولنجک، دانشگاه علوم پزشکی شهید بهشتی، دانشکده بهداشت، گروه اپیدمیولوژی، دکتر یداله محرابی؛ e-mail: ymehrabi@gmail.com

چکیده

مقدمه: رگرسیون چندک بدون داشتن محدودیت مفروضات رگرسیون معمولی، امکان دخالت متغیرهای مستقل در تمام قسمت‌های توزیع به ویژه در دنباله‌های ابتدایی و انتهایی را فراهم می‌نماید. هدف پژوهش حاضر، به‌کارگیری مدل رگرسیون چندک برای داده‌های ید دفعی ساکنین تهران و عوامل مرتبط با آن بود. مواد و روش‌ها: در پژوهش حاضر، از داده‌های مطالعه‌ی مقطعی که در آن ۶۳۹ فرد ۱۹ ساله و بالاتر با روش نمونه‌گیری تصادفی خوشه‌ای از سطح شهر تهران (۸۸-۱۳۸۷) انتخاب شده بود، استفاده گردید. به دلیل چولگی غلظت ید ادرار ۲۴ ساعته (UIC24) و به منظور بررسی نقاط فرین آن، از مدل رگرسیون چندک خطی استفاده شد. UIC24 بر حسب متغیرهای محتوای ید نمک مصرفی و میزان نمک دریافتی مدل‌سازی شد. با جایگزینی این متغیرها توسط میزان ید دریافتی، مدل دیگری برازش یافت و هر دو مدل بر اساس سن تعدیل شدند. ضرایب مدل، از روش برنامه‌ریزی خطی و با استفاده از الگوریتم سیمپلکس برآورد گردید. معنی‌داری متغیرها از روش بوت‌استرپ و نیکویی برازش مدل‌ها با شاخص آکائیک(AIC) بررسی شد. تجزیه و تحلیل آماری با نرم‌افزار R نسخه‌ی ۲ .۱۲. ۲ انجام گرفت. یافته‌ها: برازش مدل ۱ نشان داد در طول صدک‌های مورد بررسی UIC24، ضریب محتوای ید نمک مصرفی و میزان نمک دریافتی افزایش و ضریب سن کاهش یافت. مدل ۲ یافته‌های مشابهی را نشان داد اما در صدک‌های پایین‌تر از میانه برازش بهتری داشت (AIC کمتر). نتیجه‌گیری: برای مدل‌سازی تمام قسمت‌های توزیع ید دفعی و ارایه‌ی شکل رگرسیونی کامل، رگرسیون چندک مناسب‌تر از رگرسیون معمولی است.

واژگان کلیدی: رگرسیون چندک، ید دفعی ادرار، نقاط فرین، ید دریافتی
دریافت مقاله: ۹۱/۵/۱۵ـ دریافت اصلاحیه: ۹۱/۸/۱ـ پذیرش مقاله: ۹۱/۸/۲

مقدمه

مدل رگرسیون کمینه‌ی مربعات، ارتباط بین میانگین شرطی یک متغیر پاسخ برحسب یک یا چند متغیر توضیحی را بیان می‌کند و با وجود این‌که محبوب‌ترین تحلیل رگرسیونی است گاهی عملکرد ضعیفی دارد. هنگامی‌که توزیع خطا غیر نرمال است به ویژه در توزیع‌های با دنباله‌های بلند و نامتقارن، هم‌چنین در صورت وجود ناهمگنی واریانس، برآوردگرهای کمینه‌ی مربعات به داده‌های دورافتاده خیلی حساس بوده و این مورد منجر به برآوردهای اریب می‌گردد.[۱-۳]

رگرسیون چندک، تکنیکی است که می‌تواند بر محدودیت‌های یاد شده غلبه نماید.[۴،۵] این مدل که توسط کاونکر و باست در سال ۱۹۷۸ معرفی گردید به تدریج به روش جامعی برای تجزیه و تحلیل آماری مدل‌های خطی و غیر خطی متغیرهای پاسخ، در زمینه‌های مختلف تبدیل گردید.[۶،۷] به عنوان نمونه چندک‌های وزن معیار مناسبی برای تشخیص مشکلات تغذیه‌ای کودکان و بررسی وزن آن‌ها در طول زمان می‌باشد. برای بررسی عوامل موثر در کم بودن

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]

وزن هنگام تولد (وزن کمتر از ۲۵۰۰ گرم)، رگرسیون معمولی نمی‌تواند اندازه‌ی اثرات عوامل را در انتهای سمت چپ توزیع وزن نشان دهد. با استفاده از رگرسیون چندک و برآورد یک خانواده از توابع چندک شرطی، شکل‌های کامل‌تری از اثر متغیرهای توضیحی در تمام قسمت‌های توزیع به دست می‌آید.[۱] انگیزه‌ی اصلی به کارگیری رگرسیون چندک این است که با نگاهی دقیق و جامع در ارزیابی متغیر پاسخ، مدلی ارایه شود تا امکان دخالت متغیرهای مستقل، نه تنها در مرکز ثقل داده‌ها، بلکه در تمام قسمت‌های توزیع به ویژه در دنباله‌های ابتدایی و انتهایی فراهم گردد، بدون این‌که با محدودیت مفروضات رگرسیون معمولی، ناهم‌واریانسی و حضور تاثیرگذار داده‌های دورافتاده در برآورد ضرایب روبرو باشیم.[۱،۸]

درواقع، رگرسیون چندک تعمیم مفهوم یک چندک به یک چندک شرطی می‌باشد؛ وقتی یک یا چند متغیر توضیحی وجود دارد. در مقایسه با روش‌های رگرسیون کمینه مربعات که در آن برآورد توابع شرطی میانگین بر پایه‌ی مینیمم کردن مجموع مربعات باقی‌مانده‌ها انجام می‌گیرد، روش‌های رگرسیون چندک بر اساس مینیمم کردن نامتقارن قدرمطلق موزون باقی‌مانده‌ها و با هدف برآورد توابع شرطی میانه و دامنه‌ی وسیعی از سایر توابع شرطی چندک مطرح می‌گردد.[۲] اگر مدل رگرسیون خطی را به صورت $Y_i = X_i'\beta_\tau + \varepsilon_{\tau_i}$ در نظر گرفته و فرض کنیم $Q_\tau(\varepsilon_{\tau_i} \mid x_i) = 0$، آن‌گاه تابع چندک شرطی τ ام توزیع Y به شرط متغیرهای توضیحی X به صورت زیر تعریف می‌شود:

$$Q_\tau(Y_i \mid x_i) = x_i'\beta_\tau$$

برآورد پارامترهای مدل رگرسیون چندک به روش حداقل قدر مطلق انحرافات[i](LAD) انجام می‌گیرد،[۳] که در آن پارامتر رگرسیونی τ امین چندک توزیع، به‌صورت زیر به دست می‌آید:

$$\hat{\beta}(\tau) = \min_{\beta \in R^p}\Big[\sum_{i \in \{i: y_i \geq x_i'\beta\}} \tau \mid y_i - x_i'\beta \mid + \sum_{i \in \{i: y_i < x_i'\beta\}} (1-\tau) \mid y_i - x_i'\beta \mid\Big]$$

$$= \min_{\beta \in R^p} \sum_{i=1}^{n} \rho_\tau(y_i - x_i'\beta)$$

که $\rho_\tau(u)$ تابع مقادیر قدر مطلق شیب[ii] است و به صورت $\rho_\tau(u) = u(\tau - I(u<0)) \quad 0<\tau<1$ تعریف شده و در آن I تابع نشانگر است.[۶]

حل این رابطه و یافتن برآورد پارامترها از روش‌های برنامه‌ریزی خطی[iii] شامل الگوریتم‌های سیمپلکس، الگوریتم نقاط داخلی و الگوریتم هموارسازی متناهی میسر است.[۶] الگوریتم سیمپلکس برخلاف برخورداری از سرعت پایین در حالت تعداد مشاهدات زیاد، نسبت به دو الگوریتم دیگر استوارتر است و برای انواع مختلف داده‌ها، به‌ویژه داده‌هایی که تعداد زیادی نقاط پرت دارند، می‌تواند یک حل پیدا کند. الگوریتم نقاط داخلی، سرعت عمل بالایی برای تعداد زیاد مشاهدات و تعداد کم متغیرهای توضیحی دارد. الگوریتم هموارسازی متناهی ضمن داشتن تئوری ساده، از سرعت بالایی برای تعداد زیاد متغیرهای توضیحی برخوردار است.[۹،۱۰] هم‌چنین، استنباط آماری مورد استفاده در رگرسیون چندک در انتخاب الگوریتم محاسباتی نقش دارد.[۸]

ید ریزمغذی ضروری برای رشد، تکامل و بقای انسان است و کمبود آن می‌تواند طیف وسیعی از اختلال‌ها را در گروه‌های سنی مختلف ایجاد نماید.[۱۱،۱۲] کمبود ید و اختلال‌های ناشی از آن از دهه‌ی ۴۰ به عنوان یک معضل بهداشتی ـ تغذیه‌ای درایران به اثبات رسیده است. برنامه‌های کشوری با تصویب و اجرای قانون یددار نمودن نمک مصرفی در سطح کشور، از سال ۱۳۷۳ تاکنون به عنوان یک برنامه‌ی موفق ادامه داشته و به منظور اطمینان از کفایت دریافت ید افراد جامعه، میانه‌ی غلظت ید ادرار جمعیت در معرض خطر، هر ۵ سال یک بار مورد پایش قرار می‌گیرد.[۱۳،۱۴] در ایران و کشورهایی که نمک یددار مهم‌ترین و اساسی‌ترین منبع ید دریافتی‌است، کفایت دریافت ید رژیم غذایی تحت تاثیر مقدار نمک دریافتی و میزان ید نمک مصرفی می‌باشد،[۱۵،۱۶] و از طرفی، با توجه به این که متغیر غلظت ید ادرار از توزیع نرمال برخوردار نیست و دارای توزیعی با چولگی شدید می‌باشد،[۱۳،۱۴] در پژوهش حاضر با به کارگیری مدل رگرسیون چندک، داده‌های ید دفعی وعوامل مرتبط با آن در ساکنین شهر تهران که طی یک بررسی در سال ۸۸-۱۳۸۷ اندازه‌گیری شده، مورد تجزیه و تحلیل قرار گرفت.

i-Least absolute deviations

ii-Tilted absolute value function

iii-linear programming

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]

مواد و روش‌ها

داده‌های پژوهش حاضر برگرفته از طرح تحقیقاتی بررسی وضعیت تغذیه‌ای ید خانوارهای تهران در سال ۱۳۸۷-۱۳۸۸ بود که در آن خانوارها توسط مراکز بهداشتی کلان‌شهر تهران با روش نمونه‌گیری خوشه‌ای تصادفی انتخاب شدند. از آنجا که این بررسی در کل سطح شهر تهران انجام شد برای سهولت کار، نقشه‌ی شهر تهران به چهار قسمت شمال، جنوب، شرق و غرب تقسیم‌بندی و از هر قسمت، دو منطقه شهرداری به طور تصادفی انتخاب شد، و در کل تعداد ۳۸۳ خانوار در سطح شهر تهران انتخاب و بررسی گردیدند. از مادر خانواده خواسته شد خود و فرد بزرگسال دیگری (بیش از ۱۹ سال) از اعضا خانواده نمونه‌ی ادرار خود را طی ۲۴ ساعت در روز جمعه که به طور معمول اعضا خانواده منزل هستند، جمع‌آوری نمایند. در طرح تحقیقاتی یاد شده، غلظت ید ادرار ۲۴ ساعته، دریافت روزانه‌ی نمک، میزان ید نمک مصرفی به ترتیب با روش‌های هضم اسید، فلیم فتومتری (از راه اندازه‌گیری سدیم ادرار ۲۴ ساعته) و تیتراسیون اندازه‌گیری گردید. هم‌چنین، متغیرهای سن، جنس، میزان تحصیلات، محل سکونت و جمعیت مناطق مختلف تهران مورد بررسی قرار گرفت.[۱۷]

برای به کارگیری مدل رگرسیونی، ابتدا داده‌های هر منطقه را بر اساس جمعیت همان منطقه، وزن و مدل‌های رگرسیون چندک به صورت وزنی برازش داده شد. بررسی معنی‌داری حضور متغیرها در مدل، با استفاده از روش بوت‌استرپ، منجر به کنار گذاشتن متغیرهای جنس و میزان تحصیلات از مدل به دلیل عدم معنی‌داری شد. بنابراین، اولین مدل خطی رگرسیون چندک وزنی با حضور متغیرهای سن، میزان ید نمک مصرفی و دریافت روزانه نمک برازش یافت. سپس با توجه به این‌که متغیر میزان ید دریافتی، از حاصل‌ضرب دو متغیر میزان ید نمک مصرفی و دریافت روزانه نمک به دست می‌آید،[۱۷] مدل دیگری با این متغیر و متغیر سن برازش گردید. ضرایب مدل رگرسیون چندک از روش الگوریتم سیمپلکس باردوال[i] و رابرتس[ii] در صدک‌های ۵، ۱۰، ۱۵، ۲۵، ۵۰ ، ۷۵، ۸۵، ۹۰ و ۹۵ ام برآورد شد. برای مقایسه‌ی میزان برازش مدل‌ها از معیار اطلاع آکائیک[iii](AIC) استفاده گردید. برازش مدل و دیگر تحلیل‌ها به کمک نرم‌افزار R نسخه‌ی ۲. ۱۲. ۲ صورت گرفت و سطح معنی‌داری ۰/۰۵ در نظر گرفته شد.

یافته‌ها

۶۳۹ فرد شامل ۳۹۷ زن و ۲۴۲ مرد بررسی گردیدند که در محدوده‌ی سنی ۸۰-۱۹ سال قرار داشتند. افراد مورد بررسی، ۲۳٪ از جنوب، ۲۵/۵٪ از غرب، ۲۵٪ از شرق و ۲۶/۴٪ از شمال شهر تهران انتخاب شده بودند. فراوانی تحصیلات آزمودنی‌ها به ترتیب ۶/۳٪، ۳۷/۱٪، ۳۸/۶٪ و ۱۸/۰٪ بی‌سواد، راهنمایی، دبیرستان و دانشگاهی بود. میانگین ± انحراف معیار و میانه (دامنه‌ی بین چارکی) غلظت ید ادراری ۲۴ساعته‌ی شرکت‌کنندگان به ترتیب ۸۲/۹±۹۷/۶ و (۱۳۳-۳۵)۷۱/۰ میکروگرم در لیتر بود. براساس طبقه‌بندی WHO/ICCIDD/UNICEF[iv]، ۳۶/۳٪، ۲۶/۷٪، ۲۵/۹٪ و ۱۱/۲٪ شرکت‌کنندگان به ترتیب غلظت ید ادراری بیشتر از ۱۰۰، ۹۹-۵۰، ۴۹-۲۰ و کمتر از ۲۰ میکروگرم در لیتر داشتند. میانگین±انحراف معیار و میانه (دامنه‌ی بین چارکی) برای دریافت روزانه نمک به ترتیب ۳/۴±۸/۰ و (۹/۹-۵/۶)۷/۶ گرم و برای ید نمک‌های مصرفی خانوار به ترتیب ۱۴/۹±۱۹/۸ و (۳۱/۷-۳/۲)۲۱/۲ گاما بود. با توجه به مقدار نمک دریافتی و میزان ید نمک‌های مصرفی، میانگین±انحراف معیار و میانه‌ی میزان ید دریافتی ۱۴۹/۳±۱۶۱/۱ و ۱۳۲ میکروگرم در روز بود.

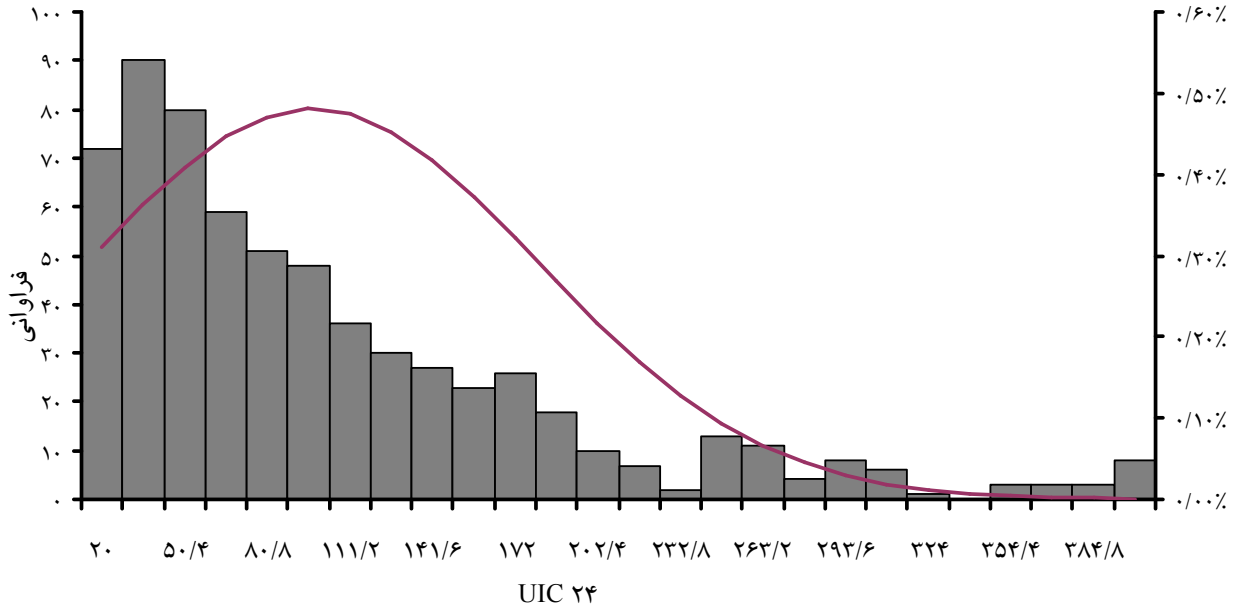
شکل ۱ نشان می‌دهد غلظت ید ادرار ۲۴ ساعته دارای چولگی به سمت راست و به میزان ۱/۵۴۹ (با خطای معیار ۰/۱) بود. تبدیل لگاریتمی متغیر نیز دارای چولگی بوده و با توزیع نرمال تفاوت معنی‌دار داشت.

یافته‌های مدل خطی رگرسیون چندک وزنی برازش داده شده با در نظر گرفتن متغیرهای توضیحی سن، میزان ید نمک مصرفی و دریافت روزانه نمک، در جدول ۱ و خلاصه‌ی یافته‌ها در نمودار ضرایب مشاهده می‌گردد (شکل ۲). در نمودار ضریب سن، با افزایش سن و به شرط ثابت فرض کردن بقیه‌ی متغیرها، غلظت ید ادرار ۲۴ ساعته کاهش یافت.

i-Barrodale
ii-Raberts
iii-Akaike information criterion (AIC)

iv-World health organization/International council for the control of iodine deficiency/The united nations children's fund

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]



نمودار ۱- هیستوگرام غلظت ید ادرار ۲۴ ساعته

تفسیر اندازه‌ی ضرایب در هر صدک، مشابه با رگرسیون کمینه مربعات است. به عنوان نمونه در مدل مربوط به چارک سوم به ازای یک واحد افزایش سن و به شرط ثابت فرض کردن بقیه‌ی متغیرها، غلظت ید ادرار ۲۴ ساعته به میزان ۱/۳۵ میکروگرم در لیتر کاهش خواهد داشت. ضریب سن در صدک ۵ ام مثبت و نزدیک به صفر بود، در حالی‌که در بقیه‌ی صدک‌های غلظت ید ادرار ۲۴ ساعته، ضریب سن منفی و با افزایش از نظر قدر مطلق همراه بود. در طول صدک‌های مورد بررسی، ضریب محتوای ید نمک مصرفی و میزان نمک دریافتی روند افزایشی را نشان داد که این روند، افزایش غلظت ید ادرار ۲۴ ساعته را در برداشت. نمودار ضریب میزان ید نمک مصرفی شیب خطی تندتر و انحراف استانداردهای کوچکتری را نسبت به نمودار میزان نمک دریافتی داشت.

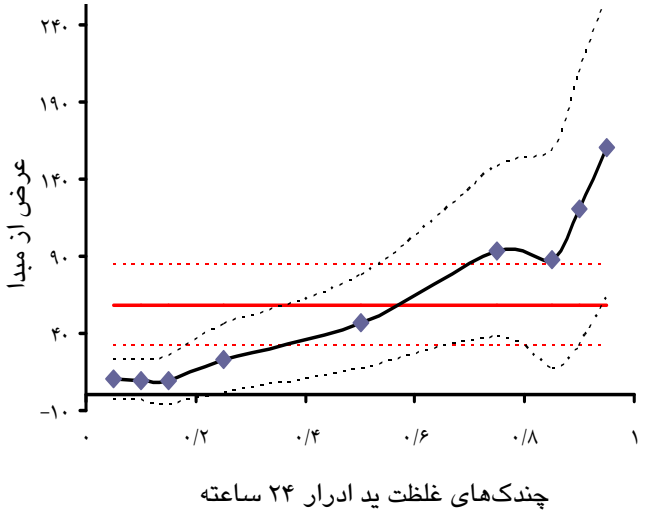
در مدل ۲ برازش داده شده، متغیرهای توضیحی سن و میزان ید دریافتی مورد بررسی قرار گرفتند (جدول ۲ و شکل۳).

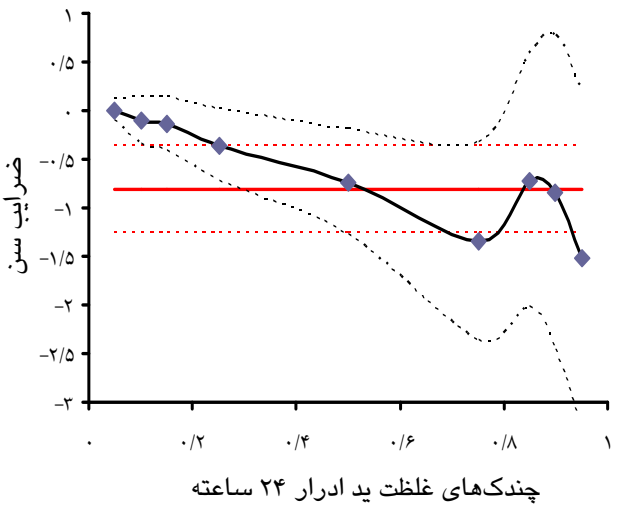
جدول ۱- یافته‌های برازش مدل رگرسیون چندک ید دفعی ادرار بر حسب متغیرهای سن، دریافت روزانه نمک و میزان ید آن (مدل ۱)

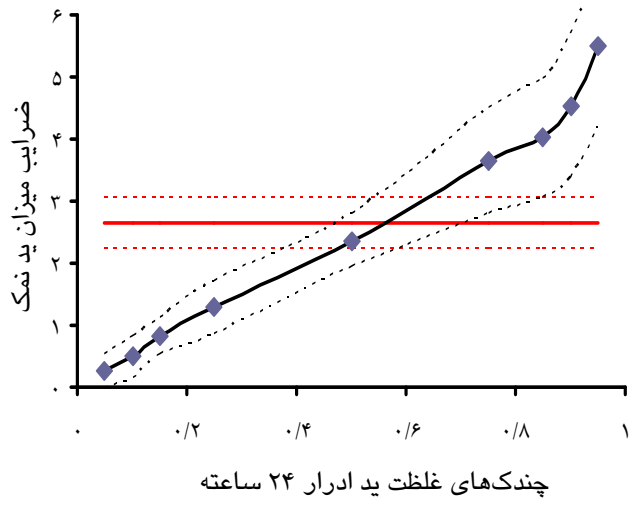
صدک	سن(سال)		میزان ید نمک مصرفی (گاما)		دریافت روزانه نمک(گرم/روز)	
	$\hat{\beta}$ (SE)	P*	$\hat{\beta}$ (SE)	P	$\hat{\beta}$ (SE)	P
۵	۰/۰۰۷(۰/۰۵۸)	۰/۹۰۷	۰/۲۵(۰/۱۳۹)	۰/۰۶۹	۰/۷۹(۰/۴۵۰)	۰/۰۷۸
۱۰	-۰/۱۰ (۰/۱۲۳)	۰/۴۳۰	۰/۴۹(۰/۱۷۰)	۰/۰۰۴	۱/۳۹(۰/۵۷۳)	۰/۰۱۵
۱۵	-۰/۱۴ (۰/۱۴۳)	۰/۳۱۹	۰/۸۱(۰/۱۴۸)	<۰/۰۰۱	۱/۵۷(۰/۶۱۴)	۰/۰۱۰
۲۵ (چارک اول)	-۰/۳۵ (۰/۱۸۹)	۰/۰۶۸	۱/۲۸(۰/۲۱۲)	<۰/۰۰۱	۱/۴۹(۰/۸۱۴)	۰/۰۶۹
۵۰ (میانه)	-۰/۷۴ (۰/۲۷۹)	۰/۰۰۸	۲/۳۶(۰/۲۱۸)	<۰/۰۰۱	۲/۶۵ (۰/۹۷۹)	۰/۰۰۷
۷۵ (چارک سوم)	-۱/۳۵ (۰/۵۲۰)	۰/۰۱۰	۳/۶۴(۰/۴۳۸)	<۰/۰۰۱	۳/۴۱ (۲/۰۶۷)	۰/۰۹۹
۸۵	-۰/۷۲ (۰/۶۶۶)	۰/۲۸۲	۴/۰۲(۰/۴۹۲)	<۰/۰۰۱	۴/۳۷ (۲/۱۴۸)	۰/۰۴۲
۹۰	-۰/۸۴ (۰/۸۲۵)	۰/۳۱۱	۴/۵۳(۰/۵۹۲)	<۰/۰۰۱	۳/۲۵ (۲/۹۰۰)	۰/۲۶۲
۹۵	-۱/۵۲ (۰/۸۷۶)	۰/۰۸۲	۵/۴۹(۰/۶۵۸)	<۰/۰۰۱	۳/۷۴ (۳/۸۱۰)	۰/۳۲۶
رگرسیون کمینه مربعات	-۰/۸۱ (۰/۲۲۹)	<۰/۰۰۱	۲/۶۵(۰/۲۱۱)	<۰/۰۰۱	۳/۰۷ (۰/۹۱۲)	<۰/۰۰۱

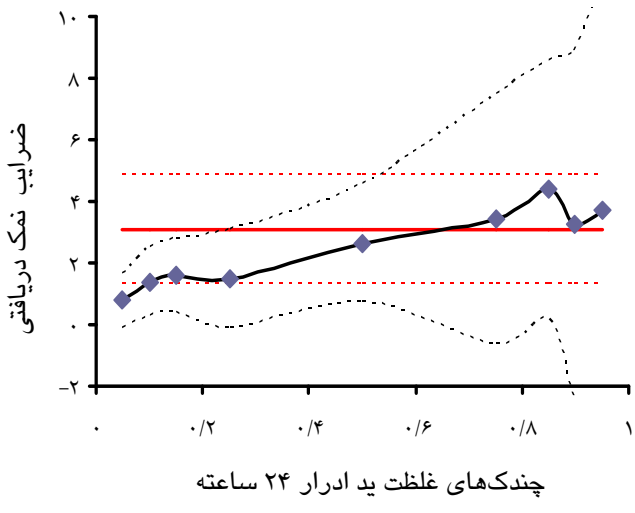
* مقدار P<۰/۰۵ از نظر آماری معنی‌دار است.

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]









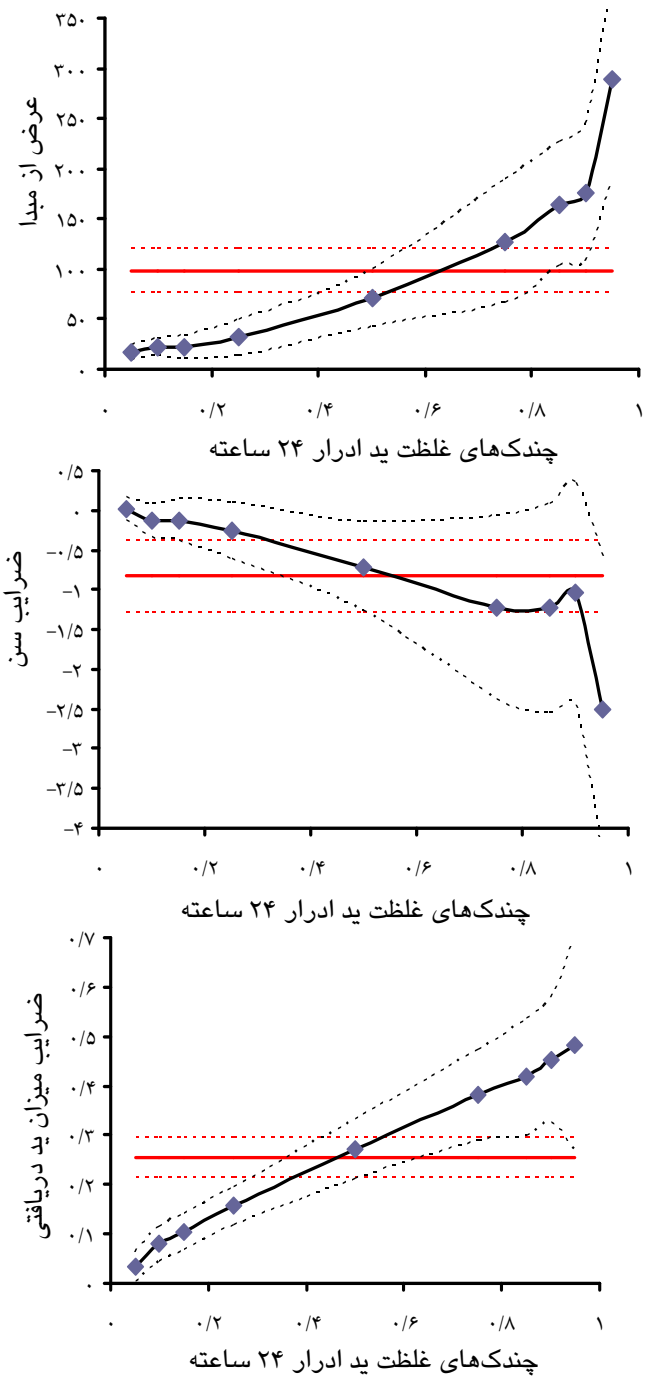
نمودار ۲- نمودارهای ضرایب مدل رگرسیون چندک (۱). خط پر رنگ اریب با نقاط روی آن، برآوردهای نقطه‌ای ضرایب رگرسیون چندک را نشان می‌دهد (به ازای هر چندک موردنظر، یک مدل رگرسیونی برازش داده شده است) و خطوط نقطه چین اطراف آن، فاصله اطمینـان ۹۵٪ است. خط ممتد افقی نشان‌دهنده برآورد کمینه مربعات اثر میانگین و خط چین‌های اطراف آن، فاصله اطمینان ۹۵٪ است.

جدول ۲- یافته‌های برازش مدل رگرسیون چندک ید دفعی ادرار بر حسب متغیرهای سن و میزان ید دریافتی (مدل ۲)

صدک	سن(سال)		میزان ید دریافتی(میکروگرم/روز)	
	$\hat{\beta}$ (SE)	P*	$\hat{\beta}$ (SE)	P
۵	-۰/۰۱(۰/۰۷۴)	۰/۸۹۴	۰/۰۳(۰/۰۱۵)	۰/۰۲۷
۱۰	-۰/۱۴(۰/۱۰۷)	۰/۱۹۰	۰/۰۸(۰/۰۱۸)	<۰/۰۰۱
۱۵	-۰/۱۲(۰/۱۳۴)	۰/۳۵۳	۰/۱۰(۰/۰۱۸)	<۰/۰۰۱
۲۵ (چارک اول)	-۰/۲۵(۰/۱۸۲)	۰/۱۷۰	۰/۱۶(۰/۰۱۹)	<۰/۰۰۱
۵۰ (میانه)	-۰/۷۱(۰/۲۸۷)	۰/۰۱۴	۰/۲۷(۰/۰۳۱)	<۰/۰۰۱
۷۵ (چارک سوم)	-۱/۲۲(۰/۵۹۰)	۰/۰۳۹	۰/۳۸(۰/۰۴۷)	<۰/۰۰۱
۸۵	-۱/۲۳(۰/۶۷۳)	۰/۰۶۷	۰/۴۲(۰/۰۶۱)	<۰/۰۰۱
۹۰	-۱/۰۴(۰/۷۱۴)	۰/۱۴۴	۰/۴۵(۰/۰۶۴)	<۰/۰۰۱
۹۵	-۲/۵۰(۰/۹۷۵)	۰/۰۱۰	۰/۴۸(۰/۱۰۹)	<۰/۰۰۱
رگرسیون کمینه مربعات	-۰/۸۳(۰/۲۳۱)	<۰/۰۰۱	۰/۲۵(۰/۰۲۰)	<۰/۰۰۱

* مقدار P<۰/۰۵ از نظر آماری معنی‌دار است.

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]



نمودار ۳- نمودارهای ضرایب مدل رگرسیون چندک (۲). خط پر رنگ اریب با نقاط روی آن، برآوردهای نقطه‌ای ضرایب رگرسیون چندک را نشان می‌دهد (به ازای هر چندک موردنظر، یک مدل رگرسیونی برازش داده شده است) و خطوط نقطه چین اطراف آن، فاصله اطمینان ۹۵٪ است. خط ممتد افقی نشان‌دهنده برآورد کمینه مربعات اثر میانگین و خط چین‌های اطراف آن، فاصله اطمینان ۹۵٪ است.

یافته‌های نمودار سن مشابه مدل ۱، ضرایب منفی با روند کاهشی را نشان داد. در مدل چارک سوم به ازای یک واحد افزایش سن و به شرط ثابت فرض کردن بقیه متغیرها، غلظت ید ادرار ۲۴ ساعته به میزان ۱/۲۲ میکروگرم در لیتر کاهش خواهد داشت. ضرایب میزان ید دریافتی مانند متغیرهای محتوای ید نمک مصرفی و میزان نمک دریافتی مدل ۱، در طول صدک‌های غلظت ید ادرار ۲۴ ساعته افزایش یافت. شیب خط و انحراف استاندارد متغیر میزان ید دریافتی نسبت به متغیرهای محتوای ید نمک و میزان نمک دریافتی کمتر بود.

معیار اطلاع آکائیک (AIC) مدل‌های برازش داده شده نشان داد مدل ۲ در صدک‌های پایین‌تر از میانه برازش بهتری نسبت به مدل ۱ داشت. در حالی‌که نیکویی برازش مدل ۱، در میانه و صدک‌های بالاتر مشاهده شد. در رگرسیون کمینه مربعات، برازش مدل ۱ بهتر از مدل ۲ بود. هم‌چنین، در هر دو مدل، نیکویی برازش صدک ۷۵ ام به طور تقریبی با نیکویی برازش رگرسیون کمینه مربعات یکسان بود (جدول ۳).

جدول ۳- معیار اطلاع آکائیک در مدل‌های برازش داده شده

صدک	مدل ۱ *	مدل ۲ †
۵	۶۹۲۷/۳	۶۹۱۳/۵
۱۰	۶۹۶۰/۳	۶۹۳۸/۴
۱۵	۶۹۸۱/۲	۶۹۶۰/۵
۲۵	۷۰۲۳/۲	۷۰۰۸/۳
۵۰ (میانه)	۷۲۰۴/۱	۷۲۱۰/۷
۷۵	۷۵۱۰/۷	۷۵۳۶/۵
۸۵	۷۷۲۱/۴	۷۷۵۹/۹
۹۰	۷۸۸۱/۹	۷۹۰۵/۷
۹۵	۸۱۰۱/۹	۸۱۵۷/۱
رگرسیون کمینه مربعات	۷۵۱۳/۶	۷۵۲۶/۷

* شامل سن، میزان ید نمک مصرفی و دریافت روزانه نمک، † شامل سن و میزان ید دریافتی

بحث

در بررسی حاضر، عوامل مرتبط با غلظت ید ادرار ۲۴ ساعته ـ به عنوان شاخص ارزیابی وضعیت تغذیه‌ای ید – با روش رگرسیون چندک مدل‌سازی شده است. با توجه به ماهیت چولگی متغیر غلظت ید ادرار ۲۴ ساعته، استفاده از رگرسیون کمینه مربعات برای بررسی عوامل مرتبط با این متغیر مناسب نمی‌باشد. در بیشتر بررسی‌ها، راهکار مورد استفاده برای این‌گونه متغیرها استفاده از رگرسیون لجستیک با تعیین یک نقطه برش است.[۱۸،۱۹] از طرفی، با توجه به ویژگی رگرسیون چندک در مدل‌سازی متغیرهای چوله، این شیوه روش مناسبی برای بررسی تمام قسمت‌های توزیع

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]

(بدون گروه‌بندی متغیر پاسخ پیوسته) و ارایه‌ی کامل شکل‌های رگرسیونی است.[۷،۸]

پژوهش حاضر نشان داد در هر دو مدل برازش داده شده، غلظت ید ادرار ۲۴ ساعته با افزایش سن، کاهش یافت. دلایل متعددی برای این امر وجود دارد که از جمله می‌توان به نمک دریافتی اشاره نمود. از آنجا که نمک دریافتی یکی از عوامل خطر فشار خون بالا و بیماری‌های قلبی ـ عروقی است، و با توجه به شیوع بالای این بیماری‌ها در کشور و افزایش توصیه‌های مبتنی بر کاهش نمک دریافتی به ویژه در سنین بالا، احتمال خطر کمبود ید در این سنین بیشتر می‌باشد.[۲۰]

با توجه به این‌که مهم‌ترین و اساسی‌ترین منبع ید دریافتی در ایران، نمک یددار می‌باشد، کفایت تغذیه‌ای ید (غلظت ید ادرار ۲۴ ساعته) بیشتر تحت تاثیر دو متغیر محتوای ید نمک مصرفی و میزان نمک دریافتی است. در بررسی حاضر مشخص شد در طول صدک‌های مختلف غلظت ید ادرار ۲۴ ساعته، شیب خط نمودار ضرایب محتوای ید نمک روند تندتری نسبت به میزان نمک دریافتی داشت که می‌توان آن را با میزان همبستگی متغیر وابسته با محتوای ید نمک مصرفی و میزان نمک دریافتی (به ترتیب $P<۰/۰۰۱$، $r=۰/۴۷$ و $P<۰/۰۰۱$ ، $r=۰/۱۲$) توجیه نمود.[۱۹]

مدلی که در آن متغیرهای سن و ید دریافتی مورد بررسی قرار گرفت (مدل ۲) نسبت به مدلی که شامل سن، محتوای ید نمک مصرفی و میزان نمک دریافتی بود (مدل۱)، برازش بهتری را در صدک‌های پایین‌تر از میانه نشان داد. با توجه به این‌که در بررسی حاضر، بیش از ۶۰٪ مقادیر غلظت ید ادرار ۲۴ ساعته در محدوده‌ی کمتر از ۱۰۰ و بقیه در محدوده‌ی ۴۰۰-۱۰۰ میکروگرم در لیتر قرار داشت، و از سوی دیگر، به دلیل اینکه غلظت ید ادرار ۲۴ ساعته بیش از آنکه تحت تاثیر هر یک از متغیرهای محتوای ید نمک مصرفی و میزان نمک دریافتی به طور مستقل باشد، متاثر از میزان ید دریافتی است که حاصل‌ضرب این متغیرهاست، مدل ۲ می‌تواند مبنای تصمیم‌گیری بهتری به منظور بررسی عوامل مرتبط با ید دفعی باشد. هم‌چنین، انحراف استانداردهای کمتر ضرایب متغیر میزان ید دریافتی، نسبت به دو متغیر دیگر در مدل ۱ نیز نشان داد میزان ید دریافتی اثر پیشگویی کننده‌ی بهتری برای غلظت ید ادرار ۲۴ ساعته دارد.

در طرح تحقیقاتی بررسی وضعیت تغذیه‌ای ید خانوارهای تهران در سال ۱۳۸۷-۱۳۸۸ یافته‌های به دست آمده نتیجه‌ی تحلیل رگرسیون لجستیک بود، که در آن متغیر طبقه‌بندی شده غلظت ید ادرار ۲۴ ساعته به عنوان متغیر وابسته (۱: کمتر از ۱۰۰، ۰: بیشتر یا مساوی ۱۰۰) و متغیرهای محل سکونت، جنس، سن، سطح تحصیلات، نمک دریافتی، سرانه‌ی مصرف نمک و میزان ید نمک مصرفی به عنوان متغیرهای مستقل وارد آنالیز شدند، این پژوهش نشان داد خطر غلظت ید ادراری کمتر از ۱۰۰ میکروگرم در لیتر در ساکنین جنوب، چارک‌های پایینی نمک دریافتی و چارک‌های پایینی میزان ید نمک مصرفی به ترتیب ۳/۰۹، ۱/۸۹ و ۸/۱۴ بود. ارتباط آماری معنی‌داری بین جنس، سن، سطح تحصیلات و غلظت ید ادراری کمتر از ۱۰۰ میکروگرم در لیتر وجود نداشت.

رگرسیون چندک علاوه بر نقاط قوت که به آن اشاره شد، دارای محدودیت‌هایی نیز می‌باشد که از آن جمله می‌توان به غیر واقعی بودن خطی فرض‌کردن چندک‌ها اشاره نمود که تعیین متغیرهای پیشگوی مناسب را با مشکل مواجه می‌سازد. محدودیت دیگر این است که در برآورد چندک‌های متعدد، احتمال تقاطع چندک‌ها وجود دارد، به ویژه زمانی‌که تعداد متغیرهای پیشگو زیاد و یا تعداد داده‌ها کم باشد.[۲۱] از محدودیت‌های داده‌های پژوهش حاضر این است که ید اندازه‌گیری شده تنها براساس نمک داده شده به خانوارهاست و دیگر منابع ید دریافتی درنظر گرفته نشدند.

پژوهش‌های مختلف بیان نموده است برای داده‌های دارای چولگی شدید، استفاده از رگرسیون چندک در مقایسه با رگرسیون معمولی و لجستیک یافته‌های بهتر و کامل‌تری را ارایه می‌دهد.[۱،۴،۷]

سپاسگزاری: نویسندگان برخود لازم می‌دانند از پژوهشکده‌ی علوم غدد درون‌ریز و متابولیسم دانشگاه علوم پزشکی شهید بهشتی که داده‌های این پژوهش را در اختیار قرار دادند و از سرکار خانم پروین سربخش که در برنامه نویسی R ما را یاری نمودند، صمیمانه تشکر وقدردانی نمایند. این مقاله برگرفته از پایان‌نامه‌ی کارشناسی ارشد آمار زیستی است.

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]

References

1. Koenker R. Quantile Regression. first ed. New York: Cambridge University Press 2005. p 1-25.
2. Koenker R, Bassett G. Regression Quantiles. Econometrica 1978; 46: 33-50.
3. Melly B. The Theory and Practice of Quantile Regression[dissertation]. University of St Gallen; 2001.
4. Eide E, Showalter M H. The effect of school quality on student performance: A quantile regression approach. Economics Letters 1998; 58: 345-50.
5. Buchinsky M. Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. The Journal of Human Resources 1998; 33: 88-126.
6. Koenker R, Hallock K F. Quantile Regression: An introduction. The Journal of Economic Perspectives 2001; 15: 143-56.
7. Keming Yu, Zudi Lu, Stander J. Quantile regression: applications and current research areas. Journal of the Royal Statistical Society: Series D (The Statistician) 2003; 52: 331-50.
8. Chen C, SAS Institute Inc, Cary N. An Introduction to Quantile Regression and the QUANTREG Procedure. SUGI 30/Statistics and Data Analysis 2005; 1-25.
9. Chen C, Wei Y. Computational Issues for Quantile Regression. The Indian Journal of Statistics 2005; 67: 399-417.
10. Koenker R, D'Orey V. Computing Regression Quantiles. Journal of the Royal Statistical Society. Series C (Applied Statistics) 1987; 36: 383-93.
11. Delange F. Iodine deficiency as a cause of brain damage. Postgrad Med J 2001; 77: 217-20.
12. Delange F. The role of iodine in brain development. Proc Nutr Soc 2000; 59: 75-9.
13. Azizi F, Mehran L, Sheikholeslam R, Ordookhani A, Naghavi M, Hedayati M, et al. Sustainability of a well-monitored salt iodization program in Iran: marked reduction in goiter prevalence and eventual normalization of urinary iodine concentrations without alteration in iodine content of salt. J Endocrinol Invest 2008; 31: 422-31.
14. Azizi F, Sheikholeslam R, Hedayati M, Mirmiran P, Malekafzali H, Kimiagar M, et al. Sustainable control of iodine deficiency in Iran: beneficial results of the implementation of the mandatory law on salt iodization. J Endocrinol Invest 2002; 25: 409-13.
15. Mustafa A, Muslimatun S, Untoro J, Lan MC, Kristianto Y. Determination of discretionary salt intake in an iodine deficient area of East Java-Indonesia using three different methods. Asia Pac J Clin Nutr 2006; 15: 362-7.
16. Melse-Boonstra A, Rozendaal M, Rexwinkel H, Gerichhausen MJ, van den Briel T, Bulux J, et al. Determination of discretionary salt intake in rural Guatemala and Benin to determine the iodine fortification of salt required to control iodine deficiency disorders: studies using lithium-labeled salt. Am J Clin Nutr 1998; 68: 636-41.
17. Nazeri P, Mirmiran P. Evaluation of iodine nutritional status in Tehranian households in 2008-09. [dissertation]. Tehran: Shahid Beheshti University of Medical Sciences and Health Services, Faculty of Nutrition Sciences and Food Technology/Department of Human Nutrition; 2009.
18. Seal JA, Doyle Z, Burgess J R, Taylor R, Cameron AR. Iodine status of Tasmanians following voluntary fortification of bread with iodine. Med J Aust 2007; 186: 69-71.
19. Nazeri P, Mirmiran P, Mehrabi Y, Hedayati M, Delshad H, Azizi F. Evaluation of iodine nutritional status in Tehran, Iran: iodine deficiency within iodine sufficeency. Thyroid 2010; 20: 1399-406.
20. Azizi F, Ghanbarian A, Momenan AA, Hadaegh F, Mirmiran P, Hedayati M, et al. Prevention of non-communicable disease in a population in nutrition transition: Tehran Lipid and Glucose Study phase II. Trials 2009; 10: 5.
21. Bjornar Bremnes J. Probabilistic Forecasts of Precipitation in Terms of Quantiles Using NWP Model Output. Monthly Weather Review 2004; 132: 338-47.

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]

Original Article

Application of Quantile Regression Model in Assessment of Urine Iodine Related Factors in Tehran Population

Biganeh E[1], Mehrabi Y[2], Mirmiran P[3], Maboudi A[1], Nazeri P[3]

[1]Department of Biostatistics, Faculty of Paramedical, & [2]Department of Epidemiology, Faculty of Public Health, & [3]Obesity Research Center, Research Institute for Endocrine Sciences, Shahid Beheshti University of Medical Sciences, Tehran, I.R. Iran

e-mail: ymehrabi@gmail.com

Received: 05/08/2012 **Accepted**: 23/10/2012

Abstract

Introduction: Quantile regression can be applied to model skewed variables, especially, when the objective is to model the tails of a response variable with highly skewed distribution. The aim of this study is to apply quantile regression to analyze urine iodine data and related factors in a Tehranian population. Materials and Methods: Data was collected in a cross-sectional study, in which 639 subjects, aged 19 years and over, were enrolled through randomized cluster sampling in Tehran between 2008-9. Due to the high skewness of 24hr urinary iodine concentrations (UIC24) and to evaluate its extreme points, two linear quantile regression models were fitted. In model I, UIC24 was regressed on iodine content of salt and daily salt intake. These variables were replaced by iodine intake in model II, both models were adjusted by age. Model coefficients were estimated using the linear programming method and simplex algorithm. Significancy of the variables were evaluated by the bootstrap method. The Akaike information criterion (AIC) was used to assess the fitting of the models. All analyses were performed using R software version 2.12.2. Results: Model I showed an increase in coefficients of iodine content of salt, daily salt intake, but a decrease in age coefficient in the length of the urinary iodine concentration percentiles. Model II showed similar results, but better fit (smaller AIC) in percentiles lower than median. Conclusion: Compared to ordinary regression, quantile regression models showed better fit, and a more complete picture and are recommended for modeling all parts of UIC24.

Keywords: Quantile regression, Urine iodine, Extreme points, Iodine intake

[Downloaded from ijem.sbmu.ac.ir on 2026-11-13]