

## مقایسه تعداد فرزندان زنده به دنیا آمده زنان شهری و روستایی مهاجر به تهران

مهسا سعادت<sup>۱</sup>

۱. استادیار، مؤسسه مطالعات و مدیریت جامع و تخصصی جمعیت کشور، آمار زیستی، تهران، ایران

### چکیده

**زمینه و هدف:** مهاجرت به عنوان یکی از پدیده‌های جمعیتی در هر شکل و با هر انگیزه یا پیامدی، تأثیرات اقتصادی-اجتماعی و فرهنگی گوناگونی را در سطوح محلی، منطقه‌ای، ملی و بین‌المللی به دنبال دارد؛ از دیگر سو باروری نیز در مطالعات علوم بهداشتی و جمعیتی جایگاه ویژه‌ای دارد و پژوهشگران این حوزه‌ها همواره به بررسی تغییرات و روند این مقوله از ابعاد مختلف پرداخته‌اند. هدف از مطالعه حاضر، مدلسازی میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر شهری و روستایی به شهرهای استان تهران با استفاده از رگرسیون درختی می‌باشد.

**روش بررسی:** داده‌ها بر اساس ۲ درصد از داده‌های خام سرشماری سال ۱۳۹۰ به دست آمد و با استفاده از مدل رگرسیون درختی تحلیل شد. مدل‌های درختی، روش‌های آماری ناپارامتری هستند که پیش‌فرض‌های پیچیده و بعضاً غیرقابل تحقق روش‌های پارامتری متداول را نداشته و از دقت مناسبی برای مدلسازی نیز برخوردار هستند. این روش‌ها تفسیر ساده نتایج را به دنبال دارند و از این رو مورد توجه بسیاری از پژوهشگران در حوزه‌های مختلف اعم از علوم اجتماعی قرار گرفته‌اند. **یافته‌ها:** سن، سطح تحصیلی، وضعیت شغلی، دلیل مهاجرت و وضعیت استفاده از اینترنت زنان مهاجر شهری و سن زنان مهاجر روستایی، به عنوان متغیرهای تأثیرگذار در پیش‌بینی میانگین تعداد فرزندان زنده به دنیا آمده به دست آمد. **نتیجه‌گیری:** نتایج مستخرج از رگرسیون درختی حاکی از آن است که زنان مهاجر شهری در گروه‌های سنی بالاتر، با سطح تحصیلی پایین‌تر، غیرشاغل و کسانی که از اینترنت استفاده نمی‌کنند، میانگین تعداد فرزندان زنده بیشتری به دنیا آورده‌اند. **کلمات کلیدی:** زنان، مهاجرت، تهران، فرزندان زنده به دنیا آمده، رگرسیون درختی.

نویسنده مسئول: مهسا سعادت

آدرس: ایران، تهران، مؤسسه مطالعات و مدیریت جامع و تخصصی جمعیت کشور

ایمیل: mahsa.saadati@gmail.com, mahsa.saadati@psri.ac.ir



## مقدمه

اجتماعی، اقتصادی و سیاسی بستگی دارد. در دیدگاه رفتاری (Behavioral) هنجارهای فرهنگی و ایدئولوژیکی بر شانس زنان برای مهاجرت تأثیر می‌گذارند. ارتباط میان اعضای خانواده بر حسب قدرت و اقتدار، بر روی مهاجرت زنان تأثیرگذار است و در دیدگاه استراتژی خانواری (Household strategy) اغلب زنان مجبور به مهاجرت می‌شوند (۱۲).

با این‌که عمده‌ترین رفتار جمعیتی که در مورد زنان مورد بررسی قرار می‌گیرد، رفتارهای باروری و فرزندآوری آنان می‌باشد، با ارتقاء پایگاه اجتماعی- اقتصادی زنان سایر رفتارهای جمعیتی آنان از جمله مهاجرت نیز اهمیت می‌یابد. بررسی ارتباط میان وضعیت مهاجرت، رفتارهای باروری و نتایج حاصل از آن، همواره مورد توجه پژوهشگران جمعیت‌شناسی و علوم بهداشتی است. رفتار باروری، بر اساس محرک‌ها، فرایند تصمیم‌گیری و نگرش‌های مرتبط می‌تواند به عنوان رفتاری اجتماعی که در یک محیط اجتماعی حادث می‌شود، در نظر گرفته شود. روابط اجتماعی بین اعضای یک جامعه می‌تواند ارزش‌ها و هنجارهای مرتبط با رفتار باروری را سازمان دهد، رفتاری که تنها در به دنیا آوردن فرزند خلاصه نمی‌شود، بلکه تمامی جنبه‌های فرزندزایی نظیر ازدواج، علاقه به داشتن فرزند و نگرش‌ها نسبت به محرک‌ها برای تنظیم باروری را شامل می‌شود. به دنیا آوردن فرزند می‌تواند به عنوان نتیجه یک فرآیند بزرگتر که تمامی این موارد را در برمی‌گیرد، تلقی شود. مهاجرت، مهاجران را در یک محیط جدید اجتماعی، اقتصادی و فرهنگی مستقر می‌سازد. اگر این دو محیط با هم مشابه نباشند، که اغلب نیستند، مهاجران در مقصد، رفتاری متفاوت از آنچه در مبدأ داشتند، در پیش می‌گیرند. بنابراین، از آن‌جا که مهاجرت محل زندگی مهاجران و محیط اقتصادی و اجتماعی، آنان را تغییر می‌دهد، بررسی رابطه بین مهاجرت و باروری حائز اهمیت است. پژوهش‌های مربوط به مهاجرت و باروری، رابطه بین وضعیت مهاجرت که به

مهاجرت یکی از رخدادهای مهمی است که زندگی بسیاری از افراد را تحت تأثیر قرار می‌دهد و بر روی جنبه‌های مختلف رفتارهای فردی و اجتماعی آن‌ها تأثیر می‌گذارد؛ در نتیجه نیازمند توجه ویژه است (۱). علاوه بر باروری و مرگ و میر، مهاجرت یکی از پدیده‌های جمعیتی است که با توجه به پیامدهای اقتصادی و اجتماعی که دارد می‌توان از جهات مختلفی به مطالعه و بررسی آن پرداخت. این مؤلفه جمعیتی یکی از عوامل مهم تغییردهنده تعداد و ترکیب سنی جمعیت است که به دو صورت در مبدأ و مقصد مهاجرت ظاهر می‌شود. در جمعیت‌شناسی، مهاجرت به شکل عام عبارت از تغییر محل اقامت فرد از یک نقطه به نقطه دیگر می‌باشد که این نقطه را معمولاً محدوده یک شهر یا یک روستا در نظر می‌گیرند (۲). در میان عوامل مؤثر بر مهاجرت، جنس و سن نقش مهمی دارند؛ افراد اغلب در سنین ویژه‌ای (بین ۲۰-۴۰ سالگی) تصمیم به مهاجرت می‌گیرند و در کل، مردان بیشتر از زنان مهاجرت می‌کنند و زنان اغلب در پیروی از همسرانشان اقدام به مهاجرت می‌نمایند (۷-۳). با این حال سهم زنان در مهاجرت طی دهه‌های اخیر افزایش یافته و در نتیجه جهان با مهاجرت زنانه‌تر مواجه شده که نتایج اقتصادی، اجتماعی و جمعیتی متفاوتی دارد (۸). در ایران نیز با توجه به تغییر موقعیت زنان طی دهه‌های اخیر مانند افزایش سواد و سطح تحصیلی، فراهم آمدن فرصت‌های شغلی و کاهش باروری کل کشور تا زیر سطح جانشینی (۹-۱۰)، سهم زنان در مهاجرت داخلی افزایش یافته است. نسبت زنان مهاجر از ۰/۳۸ در سال ۱۳۵۵ به ۰/۴۹ در سال ۱۳۹۰ افزایش یافته است (۱۱). در مطالعه‌های نیز مهاجرت زنان از دیدگاه‌های مختلف نظری تبیین شده است؛ دیدگاه نئوکلاسیک (Neoclassic)، مهاجرت روستا به شهر زنان را به فرصت‌های شغلی محدود در روستا مرتبط می‌داند که انگیزه مشترک زنان و مردان برای مهاجرت است. در دیدگاه ساختاری (Structural)، مهاجرت به تغییرات

۴ دلیل شغلی، تحصیلی، پیروی از خانوار و سایر) و وضعیت استفاده از اینترنت (به صورت استفاده می‌کند و استفاده نمی‌کند)، برای مدلسازی متغیر تعداد فرزندان زنده به دنیا آمده (منظور تعداد کل فرزندان زنده‌ای است که یک زن حداقل یکبار ازدواج کرده تا زمان سرشماری به دنیا آورده) در نظر گرفته شده‌اند.

### رگرسیون درختی

پیشرفت فناوری امکان جمع‌آوری مشاهدات و متغیرهای گوناگون را برای موضوعی خاص فراهم آورده است. در این شرایط، با توجه به حجم بالای داده‌ها دیگر نمی‌توان از روش‌های کلاسیک برای مدلسازی آماری استفاده نمود و روش‌هایی مانند درخت تصمیم که امکان محاسبات سریع و حصول نتایج دقیق را فراهم می‌کنند، اهمیت و جایگاه ویژه‌ای می‌یابند. درخت تصمیم در بسیاری از زمینه‌ها از جمله مباحث علوم پزشکی، جمعیت‌شناسی، جامعه‌شناسی، روان‌شناسی، مدیریت، هواشناسی و غیره به منظور تشخیص مدل، پیش‌بینی و تصمیم‌گیری کاربرد دارد (۲۰-۱۸). درخت‌های تصمیم برای تحلیل اکتشافی داده‌ها و ایجاد طبقه‌بندی ساده و قابل تفسیر، همچنین تعریف ضوابط پیش‌بینی با کمترین پیش‌فرض لازم مفید هستند.

درخت‌های تصمیم که به‌طور عمومی در دو دهه اخیر عمومیت یافتند، برحسب متغیر پاسخ مورد مطالعه به سه دسته مدل‌های طبقه‌بندی درختی ( Classification tree) (اگر متغیر پاسخ رسته‌ای باشد)، مدل‌های رگرسیون درختی (اگر متغیر پاسخ پیوسته باشد) و مدل‌های بقای درختی (Survival tree) (اگر متغیر پاسخ زمان بقا باشد) تقسیم می‌شوند (۲۱). با توجه به این‌که در این مقاله متغیر مورد بررسی متغیری کمی می‌باشد، برای تحلیل آن از مدل رگرسیون درختی با استفاده از الگوریتم CHAID به دلیل امتیاز این الگوریتم نسبت به الگوریتم کارت که درختی با افرازهای چند

تمامی جنبه‌های زندگی فرد مهاجر توجه می‌کند و باروری را مورد مطالعه قرار می‌دهند. رفتار باروری و مهاجرت توسط پژوهشگران بسیاری مورد بحث و بررسی قرار گرفته است (۱۷-۱۳) که در بخش بحث و نتیجه‌گیری به آن‌ها پرداخته خواهد شد.

با توجه به اهمیت باروری و عوامل مؤثر بر آن و بحث تغییرات جمعیتی بر پایه مهاجرت زنان که تا کنون بر اساس مدل‌های پیشرفته آماری در ایران مطالعه نشده است، هدف اصلی این مقاله مدلسازی میانگین تعداد فرزندان زنده به دنیا آمده زنان شهری و روستایی مهاجر به شهرهای استان تهران با استفاده از مدل رگرسیون درختی (Tree Regression) و مقایسه عوامل مؤثر بر این دو مدل با برازش مدل‌های آماری مجزا با استفاده از ۲ درصد داده‌های خام سرشماری سال ۱۳۹۰ است. بدین منظور در بخش روش بررسی به‌صورت مجزا به تشریح داده‌ها و روش رگرسیون درختی پرداخته می‌شود، در بخش یافته‌ها نتایج حاصل از برازش مدل درختی برای دو جمعیت مهاجران زن شهری و روستایی و در بخش بحث و نتیجه‌گیری نکات مهم مستخرج از این مطالعه ارائه خواهد شد.

### روش بررسی

در سرشماری‌های بسیاری از کشورها، پاسخ به سؤال «تعداد فرزندان که تاکنون به دنیا آورده‌اید» داده‌های مربوط به تعداد فرزندان زنده به دنیا آمده هر زن را تولید می‌کند که جمعیت‌شناسان اغلب از این داده‌ها برای مدلسازی آماری باروری استفاده می‌کنند. بر اساس ۲ درصد داده‌های خام سرشماری ۱۳۹۰، متغیرهای پیش‌بین سن (با ۷ طبقه زیر ۲۰ سال، ۲۹-۲۱، ۳۹-۳۰، ۴۹-۴۰، ۵۹-۵۰، ۶۹-۶۰ و ۷۰ ساله و بیشتر)، سطح تحصیلی (در ۶ طبقه بی‌سواد، زیردیپلم، دیپلم، فوق دیپلم و لیسانس، فوق لیسانس و دکتری و بالاتر)، وضعیت شغلی (در ۴ طبقه بیکار، کارکن مستقل، کارکنان مزد و حقوق بگیر و کارکن فامیلی بدون مزد)، دلیل مهاجرت (با

برای ساختن مدل رگرسیون درختی معادله (۲) باید مینیمم گردد. مقدار مینیمم این عبارت زمانی رخ می‌دهد که مقدار پیش‌بینی شده  $y(t)$  در هر گره برابر میانگین  $y_n$ ، برای کلیه حالت‌های  $(x_n, y_n)$  موجود در گره باشد.

$$(۳) \quad \bar{y}(t) = \frac{1}{N(t)} \sum_{x_n \in t} y_n$$

که در آن  $N(t)$  تعداد کل حالت‌های موجود در گره  $t$  است.

در رگرسیون درختی،  $S^*$  بهترین افزازکننده است که  $R(t)$  را کاهش می‌دهد. در واقع برای هر افزازکننده که گره  $t$  را به  $t_R$  (گره راست) و  $t_L$  (گره چپ) افزایش می‌کند، معادله (۴) را می‌توان نوشت:

$$(۴) \quad \Delta R(s, t) = R(t) - R(t_L) - R(t_R)$$

و بهترین افزاز، افزازی است که در معادله زیر صدق کند:

$$(۵) \quad \Delta R(s^*, t) = \max_{s \in S} \Delta R(s, t)$$

بنابراین رگرسیون درختی به وسیله افزازهای مکرر که منجر به بیشترین کاهش در  $R(t)$  می‌گردند، ساخته می‌شود. در این مدل، می‌توان گفت که بهترین افزازکننده روی یک متغیر در گره  $t$ ، افزازی است که به صورت موفق مقادیر بزرگ متغیر پاسخ را از مقادیر کوچک آن تفکیک نماید. لذا به‌طور کلی در این مدل رگرسیونی  $y(t_L)$  یا  $y(t_R)$  از  $y(t)$  کمتر و دیگری بزرگتر است (۲۳).

#### یافته‌ها

براساس ۲ درصد داده‌های خام سرشماری ۱۳۹۰ در مجموع ۳۳۴۲ زن حداقل یکبار ازدواج کرده از سایر استان‌ها به شهرهای استان تهران مهاجرت نمودند که ۳۰۶۱ نفر آنان (۹۱/۶ درصد) از مناطق شهری و مابقی (۸/۴ درصد) از مناطق روستایی به شهرهای استان تهران وارد شدند. بیشترین مهاجرت در سن ۲۱-۲۹ سالگی (۴۶/۹ درصد) و کمترین مهاجرت در سن ۷۰ سالگی و بیشتر (۱/۵ درصد) رخ داده است. در هر دو گروه مهاجران

شاخه‌ای به منظور تفسیر و نمایش کارآتر ایجاد می‌نماید، استفاده شد (۲۲).

ساختار رگرسیون درختی، فضای داده را به وسیله افزازهایی برای دستیابی به گره‌های پایانی تقسیم می‌نماید که در هر گره پایانی  $t$ ، مقدار پیش‌بینی شده  $y(t)$  قرار می‌گیرد. در کلیه مدل‌های رگرسیونی، داده‌ها شامل زوج مرتب‌های  $(x, y)$  هستند که  $x$  فضای اندازه‌گیری و متغیر مستقل نامیده می‌شود و  $y$  متغیر پاسخ با مقادیر حقیقی است. تشکیل یک مدل رگرسیونی با دو هدف پیش‌بینی متغیر پاسخ مطابق بردار اندازه متغیر مستقل و تشخیص رابطه میان متغیرهای اندازه‌گیری (مستقل) و متغیر پاسخ انجام می‌گیرد.

برای ساخت مدل رگرسیون درختی باید مجموعه داده‌های آموزشی دارای  $N$  حالت به‌صورت  $(x_n, y_n), \dots, (x_1, y_1)$  باشند. در رگرسیون درختی سه جزء اساسی تعیین روشی برای انتخاب یک افزاز در هر گره میانی، انتخاب ضابطه‌ای برای تعیین گره‌های پایانی و تخصیص مقادیر  $y(t)$  به هر گره پایانی  $t$  (مقادیر پیش‌بینی شده متغیر پاسخ) برای برازش درخت است (۲۳).

سؤال مطرح در مورد مدل رگرسیون درختی «چگونگی بررسی دقت پیش‌بینی مدل حاصل» است. برای پاسخ به این سؤال یک مجموعه داده آزمون بزرگ همانند مجموعه داده‌های آموزشی به حجم  $N_2$  در نظر گرفته می‌شود و دقت مدل رگرسیون درختی همانند رگرسیون معمولی به وسیله میانگین خطا و با استفاده از معادله (۱) به‌صورت زیر تعریف می‌شود:

$$(۱) \quad \frac{1}{N_2} \sum_{n=1}^{N_2} |y_n - d(x_n)|$$

که در این معادله  $d(x_n)$  پیش‌بینی‌کننده  $y_n$  (بریمن، ۱۹۸۸) و مقادیر به دست آمده همان قدرمطلق کمترین انحرافات است. معمولاً به‌طور کلاسیک در رگرسیون معمولی از روش میانگین مربعات خطا برای محاسبه برآوردها استفاده می‌شود که با استفاده از معادله زیر محاسبه می‌شود:

$$(۲) \quad R(d) = \frac{1}{N_2} \sum_{n=1}^{N_2} (y_n - d(x_n))^2$$

شده است. میزان خطای مدل با استفاده از روش اعتبارسنجی متقابل ۱۰ مرتبه‌ای ( 10-fold cross validation) ۰/۸۸۳ با خطای استاندارد ۰/۰۴۱ برآورد شد؛ در نتیجه مدل رگرسیون درختی با دقتی نزدیک به ۶۰ درصد در پیش‌بینی میانگین تعداد فرزندان زنده به‌دنیا آمده زنان مهاجر شهری کارا است.

شهری و روستایی نیز بیشترین درصد مهاجرت در گروه سنی ۲۱-۲۹ سال اتفاق افتاده است. کمترین درصد مهاجران روستایی را گروه سنی ۶۰-۶۹ سال (۱/۴ درصد) تشکیل داده و در سایر گروه‌های سنی نیز تفاوت‌هایی میان گروه‌های سنی در دو مبدأ مهاجرت وجود دارد.

بیشترین مهاجران زن دارای تحصیلات زیردیپلم (۴۰/۲ درصد) می‌باشند؛ زنان با تحصیلات دیپلم، فوق دیپلم و لیسانس، بی‌سواد و فوق لیسانس و بالاتر در رتبه‌های بعدی مهاجرت به تهران قرار دارند. این الگو کاملاً منطبق بر الگوی سطح تحصیلات زنان مهاجر شهری است که به شهرهای استان تهران مهاجرت نموده‌اند. رتبه‌بندی زنان مهاجر روستایی از نظر سطح تحصیلی به ترتیب دارای تحصیلات زیردیپلم، دیپلم، بی‌سواد و فوق دیپلم و لیسانس می‌باشد. دلایل پیروی از خانوار، سایر، شغلی و تحصیلی به ترتیب بیشترین دلیل زنان برای مهاجرت به تهران بوده است. همچنین حدود ۹۲ درصد از مهاجران بیکار بوده‌اند، این درصد در میان مهاجران شهری تقریباً برابر ۹۱ درصد و برای مهاجران روستایی تقریباً برابر ۹۷ درصد است. تنها ۳/۲ درصد از زنان مهاجر روستایی و ۱۷/۴ درصد زنان مهاجر شهری از اینترنت طی یکسال قبل از سرشماری استفاده نموده‌اند.

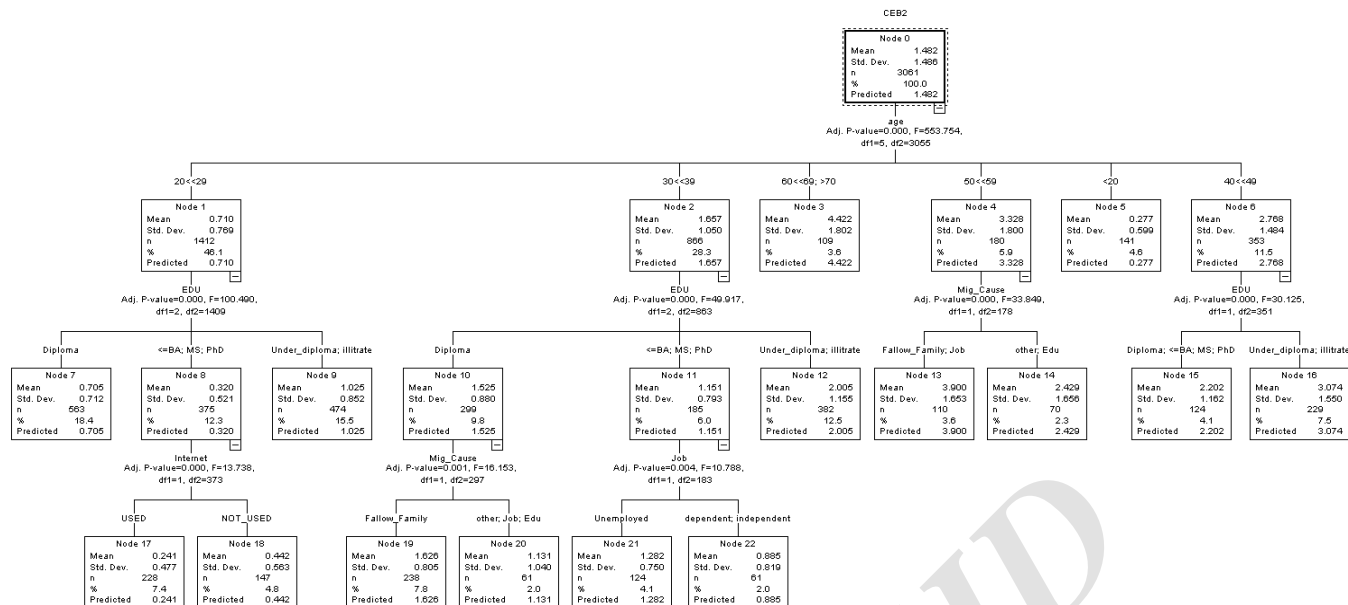
جدول (۱)، آماره‌های میانگین، انحراف معیار و نما را برای متغیر پاسخ تعداد فرزندان زنده به دنیا آمده زنان مهاجر به شهرهای استان تهران نشان می‌دهد. میانگین تعداد فرزندان زنده به‌دنیا آمده مهاجران زن شهری بیشتر از مهاجران زن روستایی است.

جدول ۱. آماره‌های توصیفی تعداد فرزندان زنده به دنیا

آمده

| تعداد فرزندان زنده به دنیا آمده |              |         | مبدأ مهاجرت        |
|---------------------------------|--------------|---------|--------------------|
| نما                             | انحراف معیار | میانگین |                    |
| ۱                               | ۱/۶۶۹        | ۱/۴۱۶   | مهاجران زن روستایی |
| ۱                               | ۱/۴۸۶        | ۱/۴۸۲   | مهاجران زن شهری    |

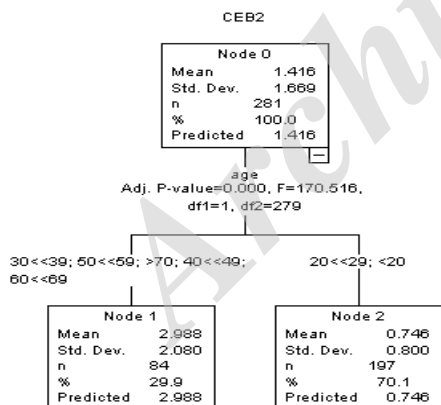
رگرسیون درختی زنان مهاجر شهری به شهرهای استان تهران که شامل ۲۳ گره و ۱۵ برگ و با استفاده از الگوریتم CHAID ساخته شده در شکل (۱) نشان داده



شکل ۱. درخت رگرسیونی میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر شهری به شهرهای استان تهران

- زنان مهاجر شهری ۵۹-۵۰ ساله‌ای که به دلیل «شغلی و پیروی از خانوار» به شهر تهران مهاجرت کرده‌اند، به صورت میانگین ۳/۹۰۰ و زنانی که به سایر دلایل مهاجرت کرده‌اند، ۲/۴۲۹ فرزند زنده به دنیا آورده‌اند.
- زنان مهاجر شهری ۶۰ ساله و بیشتر بدون تأثیر سایر متغیرهای پیش‌بین به صورت میانگین ۴/۴۲۲ فرزند زنده به دنیا آورده‌اند.

شکل (۲) رگرسیون درختی زنان مهاجر روستایی به شهرهای استان تهران را که شامل ۳ گروه و ۲ برگ است و با استفاده از الگوریتم CHAID ساخته شده نشان می‌دهد. همان‌گونه که ملاحظه می‌شود از میان متغیرهای پیش‌بین تنها سن زنان در مدل وارد شده است؛ میزان خطای مدل با استفاده از روش اعتبارسنجی متقابل ۱۰ مرتبه‌ای ۱/۲۵۳ با خطای استاندارد ۰/۱۸۹ برآورد شد؛ در نتیجه مدل رگرسیون درختی با دقتی حدود ۵۵ درصد در پیش‌بینی میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر روستایی به تهران درست عمل می‌نماید.



شکل ۲. درخت رگرسیونی میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر روستایی به شهرهای استان تهران

با توجه به درخت رگرسیونی در این شکل، زنان مهاجر روستایی با سنین کمتر یا مساوی ۲۹ سال به صورت میانگین ۰/۷۴۶ فرزند و زنان با سنین بالاتر از

همان‌گونه که در این شکل ملاحظه می‌شود، سن زنان در ریشه درخت قرار گرفته است و می‌توان گفت که در پیش‌بینی میانگین تعداد فرزندان زنده به دنیا آمده زنان بیشترین نقش را ایفا می‌کند.

نتایج زیر از درخت رگرسیونی شکل (۱) قابل استخراج است:

- زنان مهاجر شهری در گروه سنی زیر ۲۰ سال، بدون تأثیر هیچ متغیر دیگری به صورت میانگین ۰/۲۷۷ فرزند زنده به دنیا آورده‌اند.
- زنان مهاجر شهری در گروه سنی ۲۹-۲۰ سال که بی‌سواد یا زیردیپلم هستند به صورت میانگین ۱/۰۲۵ فرزند، زنانی که دارای تحصیلات دیپلم هستند به صورت میانگین ۰/۷۰۵ فرزند و زنان با تحصیلات دانشگاهی برحسب این که از اینترنت استفاده می‌کنند یا خیر به ترتیب به صورت میانگین ۰/۲۴۱ و ۰/۴۴۲ فرزند زنده به دنیا آورده‌اند.
- زنان مهاجر شهری در گروه سنی ۳۹-۳۰ سال که بی‌سواد یا زیردیپلم هستند، به صورت میانگین ۲/۰۰۵ فرزند، زنان دارای تحصیلات دیپلم که به دلیل پیروی از خانوار مهاجرت کرده‌اند به صورت میانگین ۱/۶۲۶ فرزند و زنان با سایر دلایل مهاجرت به صورت میانگین ۱/۱۳۱ فرزند زنده به دنیا آورده‌اند. زنان با تحصیلات دانشگاهی در این گروه سنی برحسب این که وضعیت شغلی آنان «بیکار» یا «کارکن مستقل یا مزد و حقوق بگیر» بوده است به صورت میانگین به ترتیب ۱/۲۸۲ و ۰/۸۸۵ فرزند زنده به دنیا آورده‌اند.

- زنان مهاجر شهری بی‌سواد و دارای تحصیلات زیر دیپلم گروه سنی ۴۹-۴۰ سال به صورت میانگین ۳/۰۷۴ فرزند زنده به دنیا آورده‌اند. میانگین تعداد فرزندان زنده به دنیا آمده زنان با تحصیلات دیپلم و بالاتر در این گروه سنی ۲/۲۰۲ فرزند می‌باشد.

نرخ‌های استفاده از وسایل پیشگیری از بارداری بیشتری نسبت به غیرمهاجران دارند. این نتیجه بعد از کنترل بر روی متغیرهای سن، طول مدت ازدواج و سطح تحصیلی به دست آورده شد (۲۶).

با توجه به اهمیت بررسی مباحث مربوط به باروری و مهاجرت از یکسو و ضرورت به کارگیری مدل‌های آماری که ضمن داشتن دقت کافی از سادگی در تفسیر نتایج نیز برخوردار باشند، در این مطالعه از مدل رگرسیون درختی در بررسی میانگین تعداد فرزندان زنده به دنیا آمده زنان حداقل یکبار ازدواج کرده شهری و روستایی مهاجر به شهرهای استان تهران استفاده شد.

از مزایای مدل درختی می‌توان به موارد زیر اشاره کرد (۲۷):

- روش‌های مبتنی بر مدل‌های درختی برخلاف روش‌های کلاسیک نیازمند فرض‌های کمتری هستند و شامل طیف وسیعی از داده‌ها می‌باشند، به همین دلیل در دو دهه اخیر این روش‌ها در برابر مدل‌های کلاسیک تحلیل ممیزی خطی ( Linear Discriminant analysis) و رگرسیون خطی مقبولیت عام یافتند. به خصوص که این مدل‌ها برای حجم بالای داده‌ها به خوبی برازش یافته و مشکلات ناشی از داده‌های گمشده (Missing data) که همواره در مدل‌های کلاسیک مطرح است، در این مدل‌ها وجود ندارد.
- انعطاف‌پذیری مدل‌های درختی نسبت به مدل‌های کلاسیک یکی از دلایل مقبولیت آن‌ها است. در واقع برای هر نوع از داده‌ها با تعدیل ضوابط افزای می‌توان یکی از مدل‌های درختی شامل طبقه‌بندی درختی برای داده‌های اسمی، رگرسیون درختی برای داده‌های پیوسته و درخت بقا برای داده‌های بقا با داده‌های سانسور شده را برازش کرد.
- بیان نتایج حاصل از مدل براساس ضوابط افزای منطقی به صورت یک درخت تصمیم و کشف ساختارهای پیچیده داده‌ها که منجر به تفسیر ساده و قابل درکی از

۳۰ سال به صورت میانگین ۲/۹۸۸ فرزند زنده به دنیا آورده‌اند.

### بحث و نتیجه گیری

با مقایسه دو درخت رگرسیونی در شکل‌های (۱) و (۲)، می‌توان نتیجه گرفت که تنها عامل مؤثر روی میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر روستایی سن آنان است، در حالی که برای زنان مهاجر شهری علاوه بر سن متغیرهای سطح تحصیلی، وضعیت شغلی، دلیل مهاجرت و وضعیت استفاده از اینترنت نیز اثرگذارند.

میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر در مطالعات بسیاری در جهان و ایران بررسی شده است؛ Bacal، رفتار باروری زنان مهاجر و غیرمهاجر فیلیپینی را با استفاده از مدل رگرسیون خطی با یکدیگر مقایسه کرد و نشان داد که در سنین اولیه باروری تفاوت معنی‌داری میان میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر و غیرمهاجر وجود ندارد. زنان تحصیل کرده و کسانی که از روش‌های پیشگیری از بارداری استفاده می‌کردند، میانگین تعداد فرزندان زنده کمتری داشتند (۲۴). Bhatia در مطالعه زنان مهاجر به شهر و روستا در هند نشان دادند که رابطه‌ای منفی میان باروری روستاییان و باروری مهاجران روستا به شهر وجود دارد. سطح تحصیلی، سن پایین ازدواج، وضعیت شغلی زنان و درآمد سرانه عوامل مؤثر در تبیین این رابطه منفی بودند (۲۵).

کازمی‌پور بیان نمود که تفاوت معنی‌داری میان ویژگی‌ها و شاخص‌های باروری مهاجران شهر و روستا وجود دارد. مهاجران شهری تحصیل‌کرده‌تر هستند و نرخ باروری کمتری نسبت به مهاجران روستایی دارند. با این حال مهاجران روستایی و شهری نسبت به غیرمهاجران این مناطق نرخ باروری کمتری دارند. برخلاف انتظار بر اساس این مطالعه، ۴ گروه مهاجر داخلی نرخ‌های باروری (که بر اساس تعداد حاملگی‌ها، تعداد فرزندان زنده به دنیا آمده و در حال حاضر زنده محاسبه شده‌اند) کمتر و



گروه‌های سنی ۵۹-۵۰ سال و ۳۹-۳۰ سال دارای تحصیلات دیپلم می‌باشد. میانگین تعداد فرزندان زنده به دنیا آمده زنان در گروه سنی ۳۹-۳۰ سال دارای تحصیلات دیپلم که دلیل مهاجرت آن‌ها پیروی از خانوار بوده، نسبت به سایر زنان مهاجر بیشتر بوده است (۱/۶۲۶ در مقابل ۱/۱۳۱ فرزند).

میانگین تعداد فرزندان زنده به دنیا آمده زنان در گروه سنی ۵۹-۵۰ سال که دلایل مهاجرت آن‌ها پیروی از خانوار و شغلی است نسبت به سایر زنان مهاجر بیشتر بوده است (۳/۹۰۰ در مقابل ۲/۴۲۹ فرزند).

متغیر دیگری که شاید کمتر از سایر عوامل مؤثر بر تعداد فرزندان زنده به دنیا آمده در مطالعات جمعیتی به آن پرداخته شده است، استفاده از اینترنت می‌باشد. الگوی درختی مهاجران زن شهری نشان داد که این متغیر بر روی میانگین تعداد فرزندان زنده به دنیا آمده زنان گروه سنی ۲۹-۲۰ سال دارای تحصیلات دانشگاهی مؤثر است؛ به نحوی که میانگین تعداد فرزندان زنده به دنیا آمده در میان زنانی که از اینترنت استفاده می‌کنند کمتر از زنانی است که از اینترنت استفاده نمی‌کنند (۰/۴۷۷ در مقابل ۰/۵۶۳ فرزند).

#### تقدیر و تشکر

این مقاله مستخرج از طرح کاوش داده‌های جمعیتی با استفاده از درخت تصمیم (ابلاغ شماره ۲۰/۱۵۲۸۳ مورخ ۹۳/۱۱/۵) می‌باشد که با حمایت مالی مؤسسه مطالعات و مدیریت جامع و تخصصی جمعیت کشور در سال ۱۳۹۳ انجام شده است.

آن‌ها می‌شود، یکی از نکات جالب توجه برای پژوهشگران در زمینه‌های مختلف علمی است.

نتایج زیر از برآزش رگرسیون درختی میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر شهری و روستایی به شهرهای استان تهران در این مطالعه (شکل‌های ۱ و ۲) به دست آمد:

- سن زنان مهمترین عامل مؤثر در میان مهاجران شهری و تنها عامل مؤثر در میان مهاجران روستایی برای تعیین میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر استان تهران می‌باشد. با افزایش سن زنان مهاجر، میانگین تعداد فرزندان زنده به دنیا آمده آنان نیز افزایش می‌یابد. نتایج مشابه در مطالعه‌ای که توسط رازقی انجام گرفته نیز به دست آمده است (۲۸).

- سطح تحصیلی زنان مهاجر شهری به استان تهران نقش مهمی در میان کلیه گروه‌های سنی (۲۹-۲۰، ۳۹-۳۰ و ۴۹-۴۰ ساله) دارد؛ افراد با تحصیلات پایین‌تر میانگین تعداد فرزندان زنده به دنیا آمده بیشتری دارند. اثر سطح تحصیلی بر روی میانگین تعداد فرزندان زنده به دنیا آمده در مطالعات بسیاری در جهان و ایران بررسی شده است (۳۱-۲۹).

- وضعیت شغلی تنها در میان قشر تحصیل کرده (زنان مهاجر شهری با تحصیلات دانشگاهی) گروه سنی ۳۹-۳۰ سال (متولدین دهه ۱۳۶۰) اثرگذار بوده به قسمی که زنان شاغل میانگین تعداد فرزندان زنده به دنیا آمده (۰/۸۸۵ فرزند) کمتری نسبت به زنان غیرشاغل (۱/۲۸۲ فرزند) دارا هستند. بررسی تأثیر وضعیت شغلی بر روی تعداد فرزندان زنده به دنیا آمده که توسط مهربار در سال ۲۰۰۳ انجام شد، نشان داد که استان‌های با میزان بالای اشتغال زنان و میزان استفاده از وسایل پیشگیری از بارداری بیشتر، باروری پایین‌تری را تجربه کرده‌اند (۳۰).

- دلیل مهاجرت نیز از عوامل مؤثر بر میانگین تعداد فرزندان زنده به دنیا آمده زنان مهاجر شهری

## References

1. Weeks, J. R. Population: An Introduction to Concepts and Issues. Belmont ,California: Wadsworth Group. 2002.
2. Momen Tuefeh, M The impact of immigration and the provinces of origin on urbanization and relative density of population census in two decades (2005-2010 census). 2012, <http://www.amar.org.ir/Portals/7/hamayesh-poster>. (Persian)
3. Chant, S. and Radcliffe S. A. Migration and Development: The Importance of Gender, In S. Chant (Ed.), Gender and Migration in Developing Countries, London :Belhaven Press, pp. 1992:1-29.
4. Fincher, R. Commentary, gender relations and the geography of migration. Environment and Planning. 1993; 25(12), pp. 1703-1715.
5. DeJong, G. Expectations, gender and norms in migration decision-making. Population Studies. 2000; 54(3), pp. 307-319.
6. Boyle, P. Population geography: Transnational women on the move. Progress in Human Geography.2002; 26(4), pp. 531-543.
7. Heering, L., van Der, R., and Wissen, L. V. The role of family networks and migration culture in the continuation of Moroccan emigration: A gender perspective. Journal of Ethnic and Migration Studies. 2004; 30(2), pp. 323-337.
8. Castles, S. and Miller, M. J. The Age of Migration, London: Macmillan.1993.
9. Abbasi-Shavazi, M. J., McDonald, P. and Hosseini-Chavoshi, M. The Fertility Transition in Iran: Revolution and Reproduction. New York: Springer.2009.
10. Statistical Center of Iran, Population indicators of census 2010. Tehran, Iran 2012.
11. Mahmoudian H., Majid Koosheshi, M. Torkashvand, and Rashvand M A Comparative Study of Social, Cultural, and Economic Characteristics of Provinces with Special Focus on Demographic Indices, Tehran: Iran Civil Registration. 2012.
12. Chant, S. Gender and Migration in Developing Countries. London and New York: Belhaven Press. 1992.
13. Sharma, H. L. A study of relationship between migration and fertility. Demography India, 1992 Jan-Jun; 21(1): 51-7.
14. Lee, B. S. and Pol, L. G. The influence of rural-urban migration on migrants' fertility in Korea, Mexico and Cameroon. Population Research and Policy Review, 1993; 12(1): 3-26.
15. Liu, G. Migrant-nonmigrant differentials in level and timing of fertility in Anhui, China. Ann Arbor, Michigan, University Microfilms International. xxii 392 p. Doctoral dissertation, Brown University.1993.
16. Goldstein, S., Goldstein, A. and Limanonda B. Migration and fertility-related attitudes and behavior in urban Thailand. In: Chang Y, Kwon TK, Donaldson PJ, ed. Society in transition: with special reference to Korea. Brown Univ. Population Studies Training Center. 1982. PSTC Reprint series; no. 82-06.
17. Manner, C. K. A model of rural-urban migration and fertility. Journal of Developing Areas; 2003;37(1): 55-71.
18. Loris, N., Lumini, A. and Manna, C. A Data Mining Approach for Predicting the Pregnancy Rate in Human Assisted Reproduction, Advanced Computational Intelligence Paradigms in Healthcare, 2011;Volume 326, pp 97-111.
19. Aurangabad, M. Comparative Analysis of Classification Techniques on Soil Data to Predict Fertility Rate for Aurangabad District INDIA, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 2014; Volume 3, Issue 2.
20. Ramezankhani, A., Pournik, O., Shahrabi, J., Khalili, D., Azizi, F. and Hadaegh, F. Applying decision tree for identification of a low risk population for type 2 diabetes. Tehran

Lipid and Glucose Study, *Diabetes Res Clin Pract*, 2014; 105(3):391-8.

21. Fan G. Regression and Survival Tree Analysis Using TARGET [dissertation], Doctor of philosophy, Alabama, University of Alabama. 2004.

22. Witten, I. and Frank, E. Data Mining Practical Machine Learning Tools and Techniques, Second Edition: ELSEVIER. 2005.

23. Breiman, L., Friedman, Jerome, Olsen, Richard, Stone, Charler. Classification and Regression Trees, Chapman & Hall. 1988.

24. Bacal, R. A. Migration and fertility in the Philippines: Hendershot's selectivity model revisited. *Philippine Population Journal*, 1988; Jan-Dec; 4(1-4): 53-67.

25. Bhatia J. C. and Sabagh, G. Migration and fertility in India. *Demography India*. 1980; Jan-Dec; 9(1-2): 54-74.

26. Kazemipour, Sh, Haghshenas N .M. Fertility and Contraceptive use among migrant and non-migrant women in Iran. IUSSP, XXV International Population Conference Tours, France, 2005:18-23 July 2005.

27. Noon A .M. and Banerjee, M. Computational Methods in Biomedical Research. In Chow SC, Jones B, Liu P J and Peace K (ed), New York, Chapman& Hall, 2008, pp.77-101.

28. Razeghi H. Marriage and fertility behavior at least once married women, 15-49 years old in 2013, Smnan-Iran. Research report; 2013. (Persian)

29. Kannan, S K. and Nagarajan, V. Factor and Multiple Regression Analysis for Human Fertility in Kanyakumari District. *Anthropologist*, 2008; 10(3): 211-214.

30. Mehryar, A. H. Demographic and Health Survey of Iran, 2000: A Summary of Main Findings. Tehran: Regional Center for Population Studies and Research, Working paper, 2003.

31. Shiri, T. and Bidarian, S. Economic Factors Affecting the Fertility of women 15-49 years old population working in school, Tehran. *Journal of social sciences*, 2009, Vol.3, No.3.

## Comparing the number of children ever born (CEB) of rural and urban migrant women to Tehran by regression tree model

Mahsa Saadati

Assistant Professor, National Population Studies & Comprehensive Management Institute, Biostatistics, Tehran, Iran

### Abstract

**Background & Objective:** Migration, in any forms and by any motivations or outcomes, as a demographic phenomenon, has various cultural and socio-economic effects on local, regional, national and international levels. On the other hand, fertility plays an important role in health and population studies and researchers have examined its changes and trends in various aspects. The aim of this research was modeling the mean number of children ever born (CEB) for women who have left their cities or villages and migrated to Tehran city using regression tree model.

**Methods:** Data was obtained from 2% of raw data from the census of 2011 and analyzed by regression tree model. Tree models are nonparametric statistical techniques which do not need complicated and unreachable assumptions of traditional parametric ones and have a considerable accuracy of modeling. These models are associated with simple interpretation of results. Therefore, they have been used by researches in many fields such as social sciences.

**Results:** Age, educational level, job status, cause of migration, internet use for urban migrant women and age for rural migrant women were assumed as influential covariates in predicting the mean number of CEB.

**Conclusion:** Regression tree findings revealed that urban migrants who were in higher age groups, lower educational levels, unemployed and have not used internet have had more mean number of CEBs.

**Key words:** Women, Migration, Tehran, Children Ever Born, Regression tree



**Corresponding Author:** Mahsa Saadati

**Address:** Comprehensive Management Institute, Tehran, Iran.

**E-mail:** mahsa.saadati@gmail.com