

## آشنایی با روش بسط شبکه‌ای: شیوه‌ای نوین برای برآورد غیر مستقیم جمعیت گروه‌های پنهان

مصطفی شکوهی<sup>۱</sup>، الهام محبی<sup>۲</sup>، اعظم رستگاری<sup>۳</sup>، سعیده حاجی مقصودی<sup>۴</sup>، علی اکبر حق دوست<sup>۵</sup>، محمدرضا بانسی<sup>۶</sup>

<sup>۱</sup> کارشناسی ارشد اپیدمیولوژی، پژوهشگر، مرکز منطقه‌ای آموزش نظام مراقبت اج‌آی‌وی/ایدز، پژوهشکده‌ی آینده‌پژوهی در سلامت، دانشگاه علوم پزشکی کرمان، کرمان، ایران

<sup>۲</sup> دکترای عمومی دامپزشکی، پژوهشگر، مرکز تحقیقات مدل‌سازی در سلامت، پژوهشکده‌ی آینده‌پژوهی در سلامت، دانشگاه علوم پزشکی کرمان، کرمان، ایران

<sup>۳</sup> کارشناسی ارشد آمار زیستی، پژوهشگر، مرکز تحقیقات مدیریت ارایه خدمات سلامت، پژوهشکده‌ی آینده‌پژوهی در سلامت، دانشگاه علوم پزشکی کرمان، کرمان، ایران

<sup>۴</sup> کارشناسی ارشد آمار زیستی، پژوهشگر، کمیته‌ی تحقیقات دانشجویی، دانشگاه علوم پزشکی کرمان، کرمان، ایران

<sup>۵</sup> دکترای اپیدمیولوژی، استاد، مرکز تحقیقات مدل‌سازی در سلامت، و مرکز منطقه‌ای آموزش نظام مراقبت اج‌آی‌وی/ایدز، پژوهشکده‌ی آینده‌پژوهی در سلامت، دانشگاه علوم

پزشکی کرمان، کرمان، ایران

<sup>۶</sup> دکترای آمار زیستی، دانشیار، مرکز تحقیقات عوامل اجتماعی مؤثر بر سلامت، پژوهشکده‌ی آینده‌پژوهی در سلامت، دانشگاه علوم پزشکی کرمان، کرمان، ایران

نویسنده رابط: علی اکبر حق دوست، نشانی: کرمان، چهارراه سمیه، خیابان جهاد، خیابان ابن سینا، معاونت تحقیقات و فناوری دانشگاه علوم پزشکی کرمان، مرکز تحقیقات مدل‌سازی در

سلامت، تلفن: ۰۳۴۱۲۲۶۳۷۲۵، نامبر: ۰۳۴۱۲۲۶۳۷۲۵، پست الکترونیک [ahaghdooost@kmu.ac.ir](mailto:ahaghdooost@kmu.ac.ir)

تاریخ دریافت: ۹۱/۱۲/۱۵؛ پذیرش: ۹۲/۶/۲

**مقدمه و اهداف:** یکی از مهم‌ترین چالش‌ها در حوزه‌ی بهداشت عمومی و نظام مراقبت بیماری‌های، برآورد تعداد افراد بیمار مبتلا به بیماری‌های نادر یا برآورد برخی زیرگروه‌های جمعیتی خاص مانند تعداد استفاده‌کنندگان مواد تزریقی (IDUs) (Injected Drug Users) می‌باشد. این در حالی است که برآورد جمعیت‌های پنهان یا دور از دسترس کاری، بسیار دشوار بوده و با مشکلات فراوانی همراه است. در سال‌های اخیر، روش‌های جدیدی برای برآورد جمعیت‌های پنهان یا دور از دسترس ارایه شده است. یکی از روش‌های نوین برای برآورد چنین جمعیت‌هایی، روش بسط شبکه‌ای (NSUM) (Network Scale-up Method) می‌باشد. این روش شامل دو بخش برآورد شبکه‌ی اجتماعی افراد جامعه، و تخمین تعداد افراد زیر جمعیت‌های پنهان یا دور از دسترس با استفاده از شبکه‌ی اجتماعی است. در این مقاله، سعی شده است با زبانی ساده، در ابتدا به بحث مختصری درباره روش‌های غیر مستقیم برآورد اندازه جمعیت‌ها، تاریخچه‌ی روش بسط شبکه‌ای، و مفهوم شبکه‌ی اجتماعی پرداخته شده، و در ادامه به برآورد گروه‌های پنهان با استفاده از این روش و نکته‌های کاربردی آن در مورد چنین جمعیت‌هایی پرداخته خواهد شد.

**واژگان کلیدی:** بسط شبکه‌ای، شبکه‌ی اجتماعی، روش‌شناسی، جمعیت‌های دور از دسترس

### مقدمه

پیش‌شرط‌های خود را دارند و بنابراین باید در انتخاب آن‌ها دقت زیادی شود. روش بسط شبکه‌ای (NSUM)<sup>۱</sup>، شیوه‌ای بدیع و مؤثر است که به سبب عدم تماس مستقیم با اعضای جامعه هدف و جمعیت‌های پنهان، به معنی واقعی روش غیر مستقیم محسوب می‌شود.

روش‌های مختلفی برای برآورد اندازه جمعیت‌ها با یک بیماری یا یک مشکل خاص بهداشتی وجود دارد. به‌طور کلی، تعیین اندازه جمعیت دو روش مستقیم، و غیر مستقیم صورت می‌گیرد، هرچند انتخاب روش برآورد اندازه‌ی جمعیت تا حد زیادی به سطح دسترسی به گروه مورد مطالعه بستگی دارد. روش‌های مستقیم

آیا تا کنون به این فکر افتاده‌اید، که چه تعداد مصرف‌کننده مواد در جامعه وجود دارد؟ آیا می‌توانید به راحتی تعداد افرادی که تماس جنسی خارج از چارچوب ازدواج دارند؛ را برآورد کنید؟ چگونه می‌توان تعداد افرادی مبتلا به یک بیماری که از شیوع پایینی در جامعه برخوردار هستند مانند صرع را برآورد کرد؟

اگر دقت کنید این گروه‌ها مخفی یا دور از دسترس هستند و یا به علت نادر بودن، برای تعیین شیوع آن‌ها نیاز به حجم نمونه بسیار سنگینی است، که به راحتی نمی‌توان از عهده‌ی هزینه‌های آن برآمد. در این گونه مواقع چه باید کرد؟ آیا راهی غیر مستقیم برای برآورد این جمعیت‌ها وجود دارد؟ پاسخ مثبت است، بلکه روش‌هایی برای برآورد چنین جمعیت‌هایی به شیوه غیر مستقیم وجود دارد. البته هر یک از این روش‌ها محدودیت‌ها، و

<sup>۱</sup> Network Scale-up Method

در روش مدل ضریبی نیز دو منبع مستقل اطلاعات یعنی عدد مبنا<sup>۱۱</sup> و ضریب<sup>۱۱</sup> مورد نیاز است. در این روش، احتمال این که نمونه‌های مورد مطالعه به صورت تصادفی گرفته شوند، بسیار ضعیف است. مناسب بودن این روش به میزان زیادی به کیفیت داده‌های موجود بستگی دارد. از فرضیه‌های اصلی روش مدل ضریبی، مستقل بودن دو منبع داده و وجود تعریف مشابه از جمعیت است (۱۲-۱۰، ۸). جدول شماره ۱، خلاصه‌ای از معایب و مزایای استفاده از این روش‌ها را نشان داده است.

با توجه به توضیحات فوق، و برای آشنایی بیشتر با روش بسط شبکه‌ای در این مقاله، ابتدا مبنای فکر، روش و تاریخچه، سپس مفهوم شبکه اجتماعی و روش‌های برآورد آن توضیح داده خواهد شد. در گام بعدی، مفهوم اصلی روش بسط شبکه‌ای بر پایه‌ی فرمول‌های ریاضی ارائه خواهد شد. هم‌چون تمام روش‌های برآوردی، این روش هم دارای برخی پیش‌فرض‌ها، نکات مثبت (قدرت) و محدودیت‌ها است، که باید به هنگام استفاده از این روش مدنظر قرار گیرند.

### روش بسط شبکه‌ای و تاریخچه‌ی آن

هم‌چنان که در بالا ذکر شد، سایر روش‌ها از مشکلات خاصی برخوردار بودند، که برآورد اندازه‌ی جمعیت را دستخوش خطا می‌کردند. بنابراین روش بسط شبکه‌ای به عنوان یک روش برآورد اندازه جمعیت‌های پنهان می‌تواند مورد استفاده قرار گیرد. این روش مزیت‌هایی دارد که شامل (۱۴-۱۳):

(۱) در این روش نیازی به تماس مستقیم با اعضای زیر جمعیت‌های مورد مطالعه نیست، به عبارت دیگر در این روش نیازی نیست که مستقیماً با افراد بیمار یا دارا بودن آن مشکل بهداشتی در ارتباط باشید، بلکه انتخاب یک نمونه‌ی تصادفی از جامعه کافی است (مهم‌ترین مزیت).

(۲) برآورد بسیاری از گروه‌ها یا بیماری‌ها با یک تحقیق منفرد امکان‌پذیر می‌باشد.

(۳) نتایج به دست آمده از روش در زمان‌های متفاوت، قابل مقایسه هستند.

این روش توسط انسان‌شناسان<sup>۱۲</sup>، متخصصان شبکه‌ی اجتماعی، و ریاضی دانان ارائه شد و برای برآورد اندازه‌ی جمعیت‌هایی که

شامل سرشماری<sup>۱</sup>، شمارش<sup>۲</sup>، و ارجاع<sup>۳</sup> می‌باشند. این روش‌ها تنها در صورت در دسترس بودن، و تماس مستقیم با جمعیت مورد بررسی قابل استفاده و اجرا هستند. در کنار برآوردهای دقیق و قابل اطمینانی که دارند؛ اجرای این روش‌ها به علت پرهزینه بودن، زمان‌بر بودن، نیاز به حجم نمونه‌ی بالا، درگیر کردن بسیاری از نیروهای انسانی، و بسیاری از مشکلات اجرایی دچار مشکل خواهد شد. علاوه بر این، عدم دسترسی یا سخت بودن دسترسی به جمعیت مورد نظر منجر به کم‌شماری<sup>۴</sup> اندازه جمعیت گروه‌های پنهان می‌شود. به عبارت دیگر، برای گروه‌ها یا بیماری‌هایی که در سطح جامعه پنهان<sup>۵</sup> هستند، یا دسترسی به آن‌ها سخت است<sup>۶</sup> امکان‌پذیر نیست. هم‌چنین برای بیماری‌ها یا مشکلاتی که از شیوع کم یا نادری برخوردار هستند، به سختی می‌توان به برآوردهای به دست آمده از این روش‌ها اعتماد کرد (۳-۱).

بر عکس روش‌های مستقیم برآورد جمعیت، روش‌های غیر مستقیم مانند صید-باز-صید<sup>۷</sup>، مدل ضریبی<sup>۸</sup>، و استفاده از روش بسط شبکه‌ای<sup>۹</sup> از چنین مشکلات اجرایی برخوردار نیستند (۴-۵). استفاده از روش‌های غیر مستقیم برای برآورد زیرگروه‌های جمعیتی پنهان، و یا بیماری یا مشکلات بهداشتی که شیوع کمی در جامعه دارند، از اهمیت ویژه‌ای برخوردار هستند. در روش صید-باز-صید، افراد متعلق به گروه‌های پنهان باید در ۲-۳ نوبت شمرده شوند، و یا بر اساس چند مرکز ثبت اطلاعات- یعنی در ۲-۳ مرحله‌ی زمانی یا در ۲-۳ سازمان مستقل- شناسایی شوند. هر چند این روش نسبتاً ساده است، اما متکی بر فرضیه‌هایی است که به سختی برقرار می‌شوند. در این روش، جمعیت باید بسته باشد، نمونه‌ها باید از یکدیگر مستقل باشند، و تمامی افراد جامعه باید شانس مساوی برای انتخاب شدن داشته باشند. علاوه بر این، برای رسیدن به برآورد معنی دار تعداد موارد مشابه یا همسان در ۲-۳ سازمان باید بیش از ۷ نمونه، مورد یا رکورد باشند. برای مثال، برای برآورد مرگ‌ومیر مادران می‌توان از داده‌های ثبت شده در وزارت بهداشت، درمان و آموزش پزشکی و هم‌چنین سازمان ثبت احوال کشور استفاده کرد (۹-۶، ۳).

<sup>۱</sup>.Census

<sup>۲</sup>.Enumeration

<sup>۳</sup>.Nomination

<sup>۴</sup>.Underestimation

<sup>۵</sup>.Hidden population

<sup>۶</sup>.Hard to reach

<sup>۷</sup>.Capture-recapture

<sup>۸</sup>.Multiplier

<sup>۹</sup>.Network scale-up

<sup>۱۰</sup>.Benchmark

<sup>۱۱</sup>.Multiplier

<sup>۱۲</sup>.Anthropologists

برآورد اندازه‌ی جمعیت مبتلایان به برخی از بیماری‌های عصبی شامل آلزایمر، پارکینسون، مالتیپل اسکلروزیس، و صرع در جمعیت شهر کرمان در سال ۱۳۹۱ (۲۱).

برآورد اندازه‌ی جمعیت مبتلایان به دیابت، افراد دچار عقب‌ماندگی ذهنی، معلولیت جسمی، قطع دست و پا، کم‌بینایی، نابینایی، کم‌شنوایی، و ناشنوایی (۲۲)، برآورد اندازه‌ی جمعیت مبتلایان سرطان پستان، سرطان رحم و تخمدان، سرطان پروستات، سرطان مثانه در جمعیت شهر کرمان در سال ۱۳۹۱ (۲۳).

تعیین بروز سقط جنین با و بدون اندیکاسیون پزشکی در کشور (۲۴).

تعیین اندازه‌ی جمعیت مصرف‌کنندگان مواد مخدر شامل تریاک، الکل، شیرهی تریاک، آمفتامین‌ها، استازی، کریستال، هروئین، کراک، حشیش، ماری جوانا، و هم‌چنین مصرف‌کنندگان تزریقی مواد (۲۵).

### منطق روش بسط شبکه‌ای

منطق این روش بسیار ساده است. فکر کنید ۵ درصد افرادی که شما می‌شناسید؛ مبتلا به دیابت باشند، و شما یک فرد معمولی جامعه باشید، می‌توان فرض کرد، که احتمالاً شیوع بیماری دیابت در جامعه ۵ درصد است. البته طبیعی است که با سؤال از یک نفر نمی‌توان چنین حرفی در مورد جامعه زد، پس باید یک نمونه با حجم نمونه قابل قبول و تصادفی از جامعه بیرون انتخاب کرد، و شیوع بیماری و یا پدیده‌های مورد نظر را در آن‌ها سنجید، و سپس آن را به جامعه تعمیم داد. احتمالاً این منطق ساده را می‌پذیرید، اما اشکالات و محدودیت‌هایی نیز به ذهن‌تان می‌رسد. ابتدا باید بررسی کرد، که چگونه می‌توان شیوع پدیده را در شبکه‌ی اجتماعی افراد تعیین کرد. اصلاً شبکه‌ی اجتماعی افراد چقدر است، و چگونه می‌توان آن را تعیین نمود؟ تعریف شبکه‌ی اجتماعی چیست؟ بعد چگونه مطمئن شویم که تعداد افراد بیمار یا مرتبط به پدیده مورد سنجش در شبکه‌ی افراد پاسخ‌گو درست برآورد می‌شود؟

شمارش آن‌ها سخت است، مورد استفاده قرار گرفت. این روش برای اولین بار در سال ۱۹۸۶ برای برآورد جمعیت گروه‌هایی که به آسانی در دسترس نبودند، به ویژه افرادی که در زلزله‌ی شهر مکزیک مفقود شدند، مورد استفاده قرار گرفت (۱۳). در سال‌های بعد نیز در چندین مطالعه که در آمریکا انجام شد. از این روش برای برآورده جمعیت گروه‌های اچ‌آی‌وی مثبت، مصرف‌کنندگان هروئین، زنانی که مورد تجاوز قرار گرفتند، و افراد خیابانی و بی‌خانمان استفاده شد (۱۷-۱۳). در این مطالعه‌ها، از روش‌های مصاحبه‌ی تلفنی و مصاحبه‌ی حضوری- با پرسش‌نامه‌های استاندارد- استفاده گردید. این محققان اعلام کرده‌اند که، برآورد زیرگروه‌هایی از جمعیت که دسترسی با آن سخت است، یا در درون جمعیت عادی پنهان هستند، و یا حتی، به سبب برخی از انگ‌های اجتماعی- مانند گروه‌های افراد معتاد به مواد مخدر، افراد تزریق کننده مواد مخدر<sup>۱</sup>، زنان روسپی<sup>۲</sup>، مردان هم‌جنس باز<sup>۳</sup>، و غیره- در دسترس نیستند؛ با روش‌های مستقیم خیلی سخت و دشوار می‌باشد. در چنین حالاتی به‌کارگیری روش‌های غیر مستقیم برای برآورد گروه‌های مورد مطالعه با چنین ویژگی‌هایی مفید می‌باشد. چگونگی برآورد اندازه جمعیت گروه‌ها، بیماری‌ها یا مشکلات بهداشتی و پیش‌فرض‌های مورد نیاز در این روش در زیر ارایه شده است (۱۰). این روش در مطالعه‌هایی برای برآورد گروه‌های پرخطر برای اچ‌آی‌وی/ایدز در کشورهایمانند ایالات متحده آمریکا، اوکراین، مولداوی، تایلند، رواندا، ژاپن، قزاقستان، برزیل، و ایران مورد استفاده قرار گرفته است.

### کاربردهای روش بسط شبکه‌ای در ایران

در سال‌های اخیر از روش بسط شبکه‌ای در برآورد برخی بیماری‌ها، اختلالات و تخمین برخی جمعیت‌های پرخطر در ایران استفاده شده است. از جمله:

تعیین اندازه‌ی شبکه اجتماعی فعال مردان شهر کرمان (۱۸)، و اندازه‌ی شبکه اجتماعی مردم ایران (۱۹).

تعیین اندازه‌ی جمعیت گروه‌های پرخطر برای اچ‌آی‌وی/ایدز شامل مصرف‌کنندگان مواد مخدر، الکل، قرص‌های اکستازی، (۲۰).

### مفهوم شبکه‌ی اجتماعی

<sup>۱</sup> Injected drug users (IDUs)

<sup>۲</sup> Female sex workers (FSW)

<sup>۳</sup> Men who have sex with men (MSM)

سؤال را بشکنیم و بپرسیم که مثلاً چند همسایه با این تعریف می‌شناسید، چند نفر در بین همکلاسی‌های قدیمی و ... می‌شناسید. شاید با جمع زدن این افراد بتوان یک برآوردی از اندازه‌ی شبکه داشت.

تعریف شناختن: محققان باید دقیقاً مشخص نمایند که منظور از شناختن یک فرد چیست. برخی مطالعه‌ها برای برآورد تعداد افراد با اچ‌آی‌وی مثبت، شناختن را به این صورت تعریف کرده‌اند که «افرادی که از طرف فرد مصاحبه شونده معرفی می‌شوند، باید حداقل در یک یا دو سال گذشته، یک‌بار با وی تماس تلفنی، حضوری و یا ایمیلی داشته و او را به اسم و چهره می‌شناسد، و او نیز متقابلاً این افراد را می‌شناسد و هر زمان که لازم باشد باید به راحتی بتوانند با آن‌ها تماس برقرار نمایند» (۱۵،۲۰). این تعریف ممکن است بر اساس ماهیت بیماری متفاوت باشد. یکی از پیش‌فرض‌های کلیدی و بسیار مهم این روش این است که «تعداد افرادی که نمونه‌های مورد مطالعه می‌شناسند، باید نمونه‌ی گویایی<sup>۴</sup> از جمعیت عمومی باشند، که نمونه‌های مورد مطالعه به آن تعلق دارند».

در روش مستقیم، ما باید تعداد افرادی را که هر فرد می‌شناسد بتوانیم بشماریم، اما آیا شما خودتان می‌توانید به این سؤال پاسخ دهید؟ لحظه‌ای تأمل کنید و به این سؤال جواب دهید. آیا می‌توانید بگویید در کل چند نفر را با تعریف بیان شده بالا به عنوان شبکه‌ی فعال خود می‌شناسید؟ طبیعی است که بسیار دشوار باشد، که انتظار داشته باشیم با یک سؤال مستقیم به نتیجه‌ای مطلوب برسیم. روش مستقیم، اما شکسته شده به زیرگروه‌ها قطعاً بهتر از شمارش افراد به صورت کلی است، اما باز اطمینان از صحت داده‌ها از یک طرف و احتمال تکراری بودن یک نفر در چند زیر گروه و شمارش چندباره وی احتمالاً اشکال جدی در برآورد ایجاد خواهد کرد. با توجه به این معذوریت‌ها و محدودیت‌ها باید به فکر روش‌های غیر مستقیم در برآورد اندازه‌ی شبکه بود (۲۷). در یک مطالعه ذکر شده است که این نوع سؤال پرسیدن در مورد اندازه‌ی شبکه اجتماعی، حدود ۲۰ درصد از افراد درون شبکه از دست می‌روند، به عبارتی دیگر ۲۰ درصد از اندازه‌ی واقعی شبکه کم‌شماری خواهد شد.

**روش غیر مستقیم:** فرض کنید در یک جامعه ۱۰۰۰۰۰ نفری که آن را با T نشان می‌دهیم، و بر اساس آمار ثبت احوال آن جامعه، ۱۰۰۰ نفر با اسم سعید در آن جامعه وجود داشته باشد، که آن را با حرف e نشان می‌دهیم. از فرد مورد مطالعه خود می‌پرسیم شما چند نفر با اسم سعید می‌شناسید. فرض کنید جواب ایشان ۵ نفر است

**شبکه اجتماعی:** هم‌چنان که در بالا اشاره شد، برای استفاده از روش بسط شبکه‌ای و برآورد پدیده یا بیماری‌های مورد نظر با استفاده از آن، باید مفهوم شبکه اجتماعی و چگونگی تعیین آن را دانست. به طور کلی یک فرد می‌تواند سه شبکه داشته باشد (۱۳-۱۴،۲۶):

۱- شبکه‌ی کلی<sup>۱</sup>: فکر کنید از شما پرسیده می‌شود، که به صورت کلی چند نفر را می‌شناسد، یعنی کل افرادی که از ابتدای عمر تا الان شما با آن‌ها آشنا شده، و مرتبط بوده‌اید، خواه الان نیز هنوز ارتباط را داشته باشید، و خواه این ارتباط الان به کلی قطع شده باشد. اگر جواب شما ۳۰۰ نفر باشد، شبکه‌ی کلی شما این عدد خواهد بود. در این نوع شبکه، ممکن است افرادی معرفی شوند که برای مدت‌های طولانی هم‌دیگر را ندیده باشند، یا از هم‌دیگر آگاهی کافی نداشته، و یا با هم‌دیگر در تماس نباشید.

۲- شبکه‌ی فعال<sup>۲</sup>: در این نوع تعریف، افرادی معرفی می‌شوند، که فرد توانایی برقراری تماس با آن‌ها را دارد. به عبارت دیگر، فرد به صورت فعالانه‌ای با این افراد در ارتباط است، و در صورت هر گونه نیاز، می‌تواند با وی ارتباط- تلفنی، حضوری، ایمیلی، یا هر روش ارتباطی دیگر- برقرار کند.

۳- شبکه‌ی حمایتی<sup>۳</sup>: در این نوع شبکه، شما افرادی را معرفی می‌کنید که ارتباط بسیار نزدیک با آن‌ها دارید و در صورت نیاز یا بروز هر گونه مشکل، شما می‌توانید روی حمایت آن‌ها حساب باز کنید. بر اساس این سه تعریف، می‌توان فهمید که شبکه‌ی حمایتی از شبکه فعال، و شبکه فعال از شبکه کلی کوچک‌تر است. دی‌گرام شماره‌ی ۱ اندازه این سه شبکه را به صورت کلی نشان داده است.

در مطالعه بسط شبکه‌ای، بر اساس بررسی متون انجام شده، تمرکز اصلی روی شبکه فعال افراد است. حال که مفهوم شبکه اجتماعی را متوجه شدیم، باید بدانیم به چه صورت می‌توان آن را محاسبه یا برآورد کرد. می‌توان گفت نخستین قدم در استفاده از روش بسط شبکه‌ای برای تعیین برآورد پدیده‌ها یا بیماری‌های مورد مطالعه، تعیین اندازه شبکه‌ی اجتماعی افراد- که با حرف C نمایش داده می‌شود است. برای برآورد C از دو روش استفاده می‌شود (۱۴).

**روش مستقیم:** در این روش از افراد پرسیده می‌شود که در شبکه فعال خود چند نفر را می‌شناسند. برای سادگی کار، شاید

<sup>۱</sup> Global network

<sup>۲</sup> Active network

<sup>۳</sup> Support network

<sup>۴</sup> Representative

صورت می‌توان آن را اندازه‌گیری یا برآورد کرد. نوبت آن است، که به بحث اصلی چگونگی برآورد گروه‌های پنهان یا بیماری‌هایی که به دنبال برآورد آن هستیم؛ پرداخته شود. در روش بسط شبکه‌ای، از افرادی که به صورت تصادفی از جامعه انتخاب شده بودند، پرسیده می‌شود، که آیا در درون شبکه خود، فردی مشابه آن مشکلی که دنبال آن هستیم- بر اساس تعریف شناختن می‌شناسند یا نه. مثلاً بر اساس تعریفی استاندارد که قبلاً مشخص کرده ایم، از افراد می‌پرسیم آیا شما در شبکه‌ی خود کسی را می‌شناسید که مصرف کننده مواد باشد (در این مرحله شما هم با ما همراه باشید و لحظه‌ای تأمل کنید، که در شبکه‌ی فعال خود کسی را می‌شناسید، که این مشکل را داشته باشد). در صورتی که فرد جواب «بلی» را ارایه دهد، در ادامه پرسیده می‌شود، که چند نفر را با چنین شاخصه‌ای مصرف کننده مواد می‌شناسند. به حالت اول که آیا فردی را می‌شناسند یا نه جواب بلی یا خیر، روش احتمالی<sup>۲</sup> برآورد گفته می‌شود، و به حالت دوم که چند نفر را می‌شناسند افرادی که جواب بلی را داده‌اند روش فراوانی<sup>۳</sup> برآورد گفته می‌شود. در کنار این دو روش، روش دیگری به اسم روش حداکثر درست‌نمایی نیز وجود دارد، که در ادامه اهمیت آن و چرایی استفاده از آن نیز ارایه خواهد شد (۱۴).

### روش (۱) روش فراوانی

اگر  $T$  کل افرادی باشد که برآورد به دست آمده به آن جمعیت تعمیم داده می‌شود،  $C$  شبکه اجتماعی و فعال افراد باشد و  $m$  تعداد افرادی بیماری (مثلاً معتاد) که افراد مورد مطالعه می‌شناسند، عدد  $e$  یا برآورد زیر گروه مورد نظر (مثلاً گروه معتادان) که محقق به دنبال برآورد آن است، بر اساس فرمول زیر محاسبه می‌شود (۱۰).

فرمول (۱)

$$m/C=e/T$$

این فرمول دقیقاً مشابه فرمولی است که در برآورد اندازه‌ی شبکه‌ی اجتماعی به روش غیر مستقیم به کار برده شد. در این فرمول سه عدد  $T$  (کل جمعیت)،  $C$  (اندازه‌ی شبکه‌ی اجتماعی افراد)، و  $m$  (تعداد افرادی معرفی شده بیمار یا شاخصه بهداشتی مورد نظر) که توسط افراد مورد مطالعه ارایه شده‌اند، معلوم

(این عدد را با  $m$  نشان می‌دهیم). با استفاده از فرمول زیر می‌توان شبکه اجتماعی فرد را به صورت غیر مستقیم برآورد کرد. در این فرمول  $C$  اندازه‌ی شبکه اجتماعی افراد است. محققان بر اساس آمارهای موجود در مراکز ثبت احوال می‌توانند تعداد کل جمعیت ( $T$ ) و همچنین در آن جمعیت کل چند نفر به اسم مثلاً سعید وجود دارد ( $e$ ) را به دست آورند. (درصد تعداد اسامی با نام سعید برابر  $e/T$ ). بر اساس فرمول زیر عدد  $C$  یعنی شبکه‌ی اجتماعی فرد بر آورد می‌شود (۲۰).

فرمول (۱)

$$m/C=e/T$$

که،  $e/T$  فراوانی نسبی (درصد) اسم سعید (که به اسم گروه رفرنس نیز شناخته می‌شوند) به صورت واقعی در آن جامعه است (در این مثال یک صدم یا ۱ درصد). شاخص  $m$  هم تعداد افرادی هستند که توسط هر فرد ذکر شده‌اند (این جا ۵ نفر به اسم سعید). اندازه‌ی شبکه‌ی اجتماعی این فرد بر اساس این فرمول برابر ۵۰۰ نفر خواهد بود. این اندازه‌ی شبکه‌ی اجتماعی برای همه‌ی افراد مورد مطالعه، محاسبه خواهد شد و میانگین آن‌ها به عنوان برآورد نهایی اندازه‌ی شبکه‌ی اجتماعی کل افراد لحاظ خواهد شد. در این روش از طریق جمعیت‌هایی که دسترسی به آن‌ها آسان است،  $C$  برآورد می‌شود، و از طریق این  $C$  به دست آمده جمعیت‌هایی که دسترسی به آن سخت است -بیماری یا مشکلات بهداشتی که به دنبال برآورد آن‌ها هستیم- را برآورد کنیم. نکته‌ی بسیار مهمی که باید در این روش متوجه آن باشیم، این است که پیشنهاد شده است برای افزایش دقت برآورد اندازه‌ی شبکه‌ی اجتماعی به روش غیر مستقیم، باید از بیش از یک گروه رفرنس استفاده کرد. برخی مطالعه‌ها بیش از ۱۵ گروه رفرنس را پیشنهاد کرده‌اند، که در نهایت اطلاعات همه‌ی این گروه‌های رفرنس با هم ترکیب می‌شوند، تا با استفاده از برآورد حداکثر درست‌نمایی<sup>۱</sup>، اندازه شبکه افراد مورد مطالعه به دست آید. در زیر توضیحات این روش ارایه شده است. پیشنهاد شده است، که شیوع این پدیده‌ها یا گروه‌های رفرنس بین ۵-۱۰ درصد باشد.

### برآورد گروه‌های پنهان

حال تا این جا یاد گرفتیم شبکه اجتماعی چیست، و به چه

<sup>۱</sup>.Probability

<sup>۲</sup>.Frequency

<sup>۱</sup> Maximum likelihood

هستند و فقط عدد  $e$  (برآورد نهایی زیر گروه یا بیماری مورد نظر) مجهول است، که محاسبه و تعیین خواهد شد (۲۸).

## روش ۲) روش احتمالی

همچنان که در بالا ذکر شد، در این روش از افراد پرسیده می‌شود، که آیا حداقل یک فرد که دارای ویژگی یا بیماری مورد مطالعه باشد را می‌شناسند. گفته می‌شود که در این روش چون از افراد پرسیده می‌شود آیا حداقل یک نفر را می‌شناسید که در شبکه‌ی مورد نظر خود، داری چنین بیماری یا مشکل بهداشتی است، راحت‌تر جواب ارایه می‌شود، تا این که از آن‌ها پرسیده شود که چند نفر بیمار را می‌شناسند. این محققان فرمول زیر را برای برآورد عدد  $C$  با این روش پیشنهاد داده‌اند (فرمول ۲):

فرمول ۲)

$$(1-PT)C=(1-Pr)$$

که بر اساس این فرمول می‌توان عدد  $C$  را به این صورت به دست آورد:

فرمول ۳)

$$C=\ln(1-Pr)/\ln(1-PT)$$

در فرمول ۴،  $Pr$  نسبت افرادی است که ذکر می‌کنند که حداقل یک نفر را می‌شناسند (چند درصد حداقل یک نفر را می‌شناسند) و  $PT$  هم نسبتی از افراد است که در آن جمعیت با آن اسم (که در روش محاسبه  $C$  به صورت غیر مستقیم بیان شد)، وجود دارد (یعنی مقدار  $e/T$ ). بنابراین، به صورت خلاصه‌تر، برای برآورد اندازه بیماری مورد مطالعه (عدد  $e$ ) از فرمول زیر (فرمول ۴) استفاده خواهد شد (۲۸):

فرمول ۴)

$$e=T(1-(1-p)^{1/C})$$

هنوز مشخص نیست که کدام یک از این دو روش برآورد دقیق‌تری را ارایه می‌کنند. در مطالعه‌ها و جوامع مختلف ممکن است دقت آن‌ها یکسان نباشد، و این یکی از جاهایی است که نیاز به پژوهش‌های بیشتر در مورد آن حس می‌شود.

روش ۳) روش حداکثر درست‌نمایی ((ML) Maximum likelihood)

فرمول‌های فوق، برآوردی از اندازه‌ی شبکه بر اساس پاسخ به سؤال در مورد یک گروه رفرنس را ارایه می‌دهد. طبیعی است که، برای افزایش دقت برآورد، بهتر است به جای یک گروه رفرنس

چندین گروه رفرنس را برگزینیم، و بر اساس ترکیب آن‌ها یک برآورد دقیق‌تر ارایه دهیم. به صورت کلی پیشنهاد می‌شود حداقل بیش از ۱۵ گروه رفرنس برای برآورد اندازه‌ی شبکه داشته باشیم، اما یک مشکل جدی با انتخاب گروه‌های رفرنس متعدد پیش می‌آید، هر گروه رفرنس یک اندازه‌ی شبکه را ارایه خواهد داد، اما چگونه باید این اندازه‌ها را با هم ترکیب نمود؟ شاید یک پاسخ این باشد که میانگین اندازه شبکه‌های به دست آمده با گروه‌های رفرنس مختلف را مبنای قرار دهیم. این پیشنهاد بدی نیست و قابل قبول است، اما شاید از نظر آماری بهترین محاسبه نباشد. اگر فرضاً شما ۱۷ گروه رفرنس داشته باشید و ۱۶ گروه، اندازه‌ی شبکه را چیزی بین ۳۳۰-۳۰۰ برآورد کنند، و یکی از آن‌ها عدد ۵۰۰ را ارایه دهد، آیا باز میانگین گرفتن از همه این گروه‌ها کار درستی است؟

برای حل این مشکل، یک روش آماری پیشنهاد می‌شود، و آن استفاده از روش ML است (۱۴). در این روش، سعی می‌شود مقدار اندازه‌ی شبکه‌ی عددی انتخاب شود، که بهترین برازندگی را برای برآورد جمعیت‌های گروه‌های رفرنس مختلف را داشته باشد. به زبان ساده‌تر، اگر مقدار اندازه‌ی شبکه مثلاً ۳۱۰ در فرمول گذاشته شود، برای هر یک از گروه‌های رفرنس، جمعیتی را پیش‌بینی می‌کند، این مقدار برآورد شده برای جمعیت گروه‌های رفرنس است، حال ما مقدار واقعی جمعیت گروه‌های رفرنس را داریم، و مقدار برآورد شده را نیز محاسبه کرده‌ایم. طبیعی است که بین این دو اختلاف باشد، در روش ML سعی می‌شود، مقدار اندازه‌ی شبکه در یک دامنه وسیع تغییر کند، و مقداری از آن پذیرفته شود که در مجموع فاصله بین مقادیر مشاهده شده و مقادیر پیش‌بینی شده برای جمعیت گروه‌های رفرنس را به حداقل برساند. البته محاسبات این روش چندان آسان نیست، و نیاز است از نرم‌افزارهای آماری استفاده شود و با برنامه‌نویسی این فرایند به سادگی انجام شود. نگارندگان این مقاله تجربه نوشتن چنین برنامه‌ای را در نرم‌افزار stata و هم‌چنین در اکسل دارند. هم‌چنین برخی دیگر از مطالعه‌ها از رویکردهای رگرسیونی با روش حداقل مربعات نیز استفاده کرده‌اند.

## نکات مهم در روش بسط شبکه‌ای

این روش نیز مثل سایر روش‌های برآورد نیازمند است، که به دقت مورد استفاده قرار گیرد و به پیش‌فرض‌ها و خطاها یا سوگرایی‌های احتمالی و بالقوه آن توجه ویژه‌ای شود. این‌جا به برخی نکات بسیار مهم به هنگام استفاده از این روش اشاره

خواهیم کرد.

**پیش‌فرض‌ها:** در زیر برخی پیش‌فرض‌های مهم برای این روش ارایه شده است. اگر این پیش‌فرض‌ها رعایت شوند، می‌توان انتظار داشت برآوردهای به‌دست آمده هم از اعتبار و هم از تعمیم‌پذیری بالایی برخوردار خواهند بود (۱۴)، که شامل:

الف) هر شخصی از جمعیت T احتمال مشابه و یکسانی برای شناختن یک نفر در جمعیت E داشته باشد.

ب) هر شخصی از جمعیت T همه آشنایانش<sup>۱</sup> را بشناسد، و بداند آیا به گروه پنهانی متعلق هستند یا خیر، به عبارتی شبکه‌ی اجتماعی فعال خود را به خوبی بشناسد.

ج) به یاد آوردن تعداد افراد شبکه‌ی اجتماعی خیلی برای افراد مشکل نباشد، به عبارتی به راحتی بتوانند تعداد شبکه‌ی اجتماعی و فعال خود را به یاد بیاورند.

**مزیت‌ها:** این روش از آنجایی که کم‌هزینه، سریع، عدم نیاز به نمونه‌ی نیروی انسانی زیاد، استفاده از پرسش‌گری، برآورد تعداد زیادی از گروه‌های هدف مورد نظر، آسان بودن نمونه‌گیری به‌طور تصادفی از جامعه و کاربرد بودن، بسیار مورد توجه نه تنها کشورهای پیشرفته مثل ایالات متحده آمریکا قرار گرفته است، بلکه برخی کشورهای در حال توسعه نیز توانسته‌اند از این روش و از نتایج به‌دست آمده آن استفاده کنند.

**خطاها یا محدودیت‌ها:** با این حال، با توجه به پیش‌فرض‌های بالا، در زیر مهم‌ترین سوگرایی‌های بالقوه‌ای که ممکن است این روش را تهدید کند را مورد بحث قرار داده‌ایم، که محققان باید به دقت به آن‌ها توجه داشته باشند (۲۰، ۲۹):

۱) **تورش انتقال اطلاعات:**<sup>۲</sup> وقتی از افراد مورد مطالعه در مورد تعداد اندازه شبکه اجتماعی وی سؤال می‌شود، ممکن است افرادی معرفی شوند که صرفاً وی را شناخته اما از رفتارهای واقعی وی آگاهی کامل نداشته باشند. مثلاً ممکن است یکی از دوستان خود را به عنوان یکی از افراد شبکه‌ی اجتماعی خود معرفی کند، اما از رفتار وی که ممکن است مصرف کننده دارو باشد، آگاهی نداشته باشد. همچنین ممکن است تعریف زیر گروه هدفی که دنبال آن هستیم، از دیدگاه افراد کاملاً متفاوت باشد (هر چند در هنگام مطالعه، محققان تعریفی واحد و یکسان از زیر گروه‌های هدف برای همه افراد پاسخ دهنده ارایه می‌دهند، اما باز هم بیم

آن می‌رود که این افراد بر اساس بینش خود یک فرد FSW را معرفی کنند)، یا این‌که تعریف این افراد در فرهنگ‌های مختلف متفاوت باشد. راه حل جالبی برای حل این مشکل پیشنهاد شده است. گفته شده است که با افراد گروه هدف (مثلاً IDUها یا FSWها) نیز ارتباط برقرار کنید و از آن‌ها بپرسید که با چه احتمالی خویشاوندان شما از رفتار شما آگاهی دارند. به احتمال زیاد همه خویشاوندان وی از رفتار این افراد گروه هدف آگاهی ندارند. با دانستن احتمال شناختن گروه‌های هدف توسط خویشاوندان می‌توان تا حدی به اندازه کم‌شماری احتمالی گروه هدف نیز پی برد. به منظور تعیین این‌که چه اندازه تعریف‌های متفاوت منجر به برآورد اندازه گروه‌های هدف خواهد شد، مطالعه‌ای در چین روی افراد تن‌فروش و مشتریان<sup>۳</sup> آن‌ها در مکان‌های (پاتوق) خاصی که رابطه جنسی برقرار می‌کردند، انجام گرفت (۳۰). این مطالعه نشان داد که فقط ۴۴ درصد از افرادی که در ازای برقراری رابطه‌ی جنسی پول دریافت می‌کردند، خود را به عنوان یک تن‌فروش معرفی کرده بودند (بر اساس تعریف حداقل یک رابطه جنسی در ازای پول در ۱۲ ماه اخیر). مطالعات در سایر کشورها روی سایر گروه‌های هدف (مثلاً IDUs) نیز نتایج مشابهی از خطای کم‌شماری خودشناختن<sup>۴</sup> را نشان داد. این خطا همچنین روی افراد پاسخ‌دهنده نیز تأثیر خواهد گذاشت.

۲) **اندازه‌ی شبکه نسبی:**<sup>۵</sup> باید توجه داشته باشیم که ممکن است اندازه شبکه‌ی اجتماعی افرادی که در گروه‌های هدف قرار دارند، با افراد جمعیت عمومی متفاوت باشند. مثلاً اگر اعضای گروه هدف (مثلاً IDUsها) اندازه شبکه اجتماعی کوچک‌تری در مقایسه با جمعیت عمومی داشته باشند، برآورد این گروه هدف کم‌شماری<sup>۶</sup> خواهد شد. در این مورد نیز راه‌حل آن است که از افراد گروه‌های هدف نیز در مورد اندازه‌ی شبکه‌ی آن‌ها سؤال شود تا این‌که اندازه‌ی خطاها مشخص شود. اصلاح کردن یا تعدیل کردن برای میزان خطای انتقال اطلاعات<sup>۷</sup> و اندازه‌ی شبکه‌های متفاوت، رویکرد بسط تعمیمی<sup>۸</sup> نامیده می‌شود. باید توجه داشته باشیم که پیش‌فرض آن است که این دو نوع اصلاح یا تعدیل از هم مستقل هستند. اگر این دو مورد با هم مرتبط باشند، روی

<sup>۲</sup>. Patrons

<sup>۳</sup>. Self-identification

<sup>۴</sup>. Relative network size

<sup>۵</sup>. Under estimate

<sup>۶</sup>. Information transmission error

<sup>۷</sup>. Generalized scale up approach

<sup>۱</sup>. Acquaintance

<sup>۲</sup>. Transmission bias

برآورد تأثیر خواهند گذاشت.

**۳) سوگرایی گزارش‌دهی<sup>۱</sup>:** این نوع خطا به علت خطای «درجه‌ی اشتیاق یا تمایل اجتماعی<sup>۲</sup>» اتفاق می‌افتد. در این حالت؛ به علت وجود برخی شرایط خاص در برخی زیرگروه‌های هدف، مثل FSW (مثلاً، از انگ اجتماعی برخوردار هستند)، افراد پاسخ دهنده ممکن است به سؤال پاسخ ندهند. این عدم پاسخ یا بی‌میلی در پاسخ به چنین سؤالاتی ممکن است منجر به سوگرایی گزارش‌دهی شود. این نوع خطا در بین کشورهای مختلف، و حتی در درون کشورها بر اساس قومیت‌ها و نژادهای مختلف، متفاوت خواهد بود. راه‌حل آن است که در چنین حالتی نتایج بر اساس برخی متغیرهای مهم مانند سن، شهری یا روستایی، جنسیت، و یا حتی درجه احترام<sup>۳</sup> به گروه‌های هدف، تعدیل شود. این تعدیل‌ها در مطالعات کشورهای مولداوی<sup>۴</sup> و اوکراین صورت گرفت، اما پیشنهاد شده است که برای کاهش این خطا بهتر است روی بهبود روش‌های جمع‌آوری اطلاعات و کاهش خطاهای درجه‌ی تمایل پاسخ دهنده برای پاسخ دقیق به سؤالات تمرکز داشته باشیم، تا این‌که فقط به اندازه‌گیری میزان «انگ» گروه‌های هدف بسنده کنیم.

**۴) اثرات موانع<sup>۵</sup>:** این خطا، ممکن است به علت ترکیبی از تصادفی و غیر تصادفی بودن افراد<sup>۶</sup> و توزیع نابرابر جمعیت-موانع فیزیکی مثل تفاوت در ژئوگرافی مردم و یا موانع اجتماعی مثل جنسیت یا نژاد- باشد. در این حالت، احتمالات متفاوتی برای شناختن افراد برای برآورد اندازه شبکه اجتماعی وجود خواهد داشت. تفاوت احتمال در اندازه‌ی جمعیت بر اساس پاسخ‌دهنده‌های متفاوت ممکن است خطاهایی را به‌وجود بیاورد. متعاقباً، احتمال شناختن گروه‌های هدف بر این اساس ممکن است متفاوت باشد. برای به حداقل رساندن این خطا، افراد پاسخ‌دهنده انتخاب شده برای مطالعه باید نماینده‌ی گویایی از جمعیت واقعی آن منطقه باشند، اما باز هم این احتمال می‌رود که به علت برخی تغییرات جمعیتی در گذر زمان، انتخاب تصادفی افراد که نماینده‌ی گویایی از جمعیت کل باشد، اتفاق نیافتد (۳۱).

**۵) کم-یادآوری افراد شبکه<sup>۷</sup>:** گفته می‌شود، افراد پاسخ‌دهنده تمایل به کم‌شماری اسامی گروه‌های رفرنس با شیوع

بالا دارند، به عبارتی دیگر گروه‌های رفرنسی که از شیوع بالا برخوردار هستند، با احتمال بیش‌تری منجر به کم‌شماری توسط افراد پاسخ‌دهنده خواهند شد. راه‌حل این مشکل، تهیه یک منحنی درجه‌بندی<sup>۸</sup> است (۳۱). فرض کنید افراد پاسخ دهنده دارای یک شبکه واقعی<sup>۹</sup> هستند، اما به هنگام مصاحبه و بدون سوگرایی، تعدادی افراد را به عنوان شبکه خود معرفی می‌کنند<sup>۱۰</sup>. شبکه‌ی معرفی شده زیر مجموعه‌ای شبکه‌ی واقعی افراد است. برای گروه‌های کوچک- مانند اسامی که از شیوع کم‌تری برخوردار هستند- یک ارتباط خطی، و برای گروه‌های بزرگ‌تر - مانند اسامی که از شیوع بیش‌تری برخوردار هستند- یک ارتباط ریشه دوم بین اندازه‌ی شبکه‌ی معرفی شده و اندازه‌ی شبکه‌ی واقعی افراد وجود دارد. نوع ارتباط به نوع داده‌های مورد مطالعه بستگی دارد. استفاده از منحنی درجه‌بندی، شبکه‌ی معرفی شده توسط افراد را به شبکه‌ی واقعی افراد تعدیل می‌کند. میزان این تعدیل در گروه‌های کوچک‌تر، کم‌تر و در گروه‌های بزرگ‌تر بسیار بیش‌تر خواهد بود. توجه داشته باشیم، که استفاده از منحنی درجه‌بندی، اندازه‌ی شبکه‌ی اجتماعی افراد را بالا خواهد بود.

**۶) شفافیت متفاوت گروه‌های رفرنس<sup>۱۱</sup>:** نکته‌ی بعدی که باید به آن توجه کرد؛ این است که به هنگام تعیین اندازه‌ی شبکه‌ی اجتماعی افراد به روش غیر مستقیم (C غیر مستقیم)، به مشکلی به اسم شفافیت بر می‌خوریم (۲۷). هم‌چنان که در بالا هم اشاره شد، از افراد پاسخ‌دهنده در مورد یک سری پدیده‌های اجتماعی- که به اسم گروه رفرنس شناخته می‌شود- برای برآورد شبکه‌ی اجتماعی افراد به روش غیر مستقیم سؤال می‌پرسیم. هم‌چنان که در بالا نیز اشاره شد، پیشنهاد می‌شود گروه‌های رفرنس بیش از ۱۵ گروه باشد، تا این‌که برآورد به دست آمده با روش ML از دقت بالاتری برخوردار باشد. حال گروه‌های رفرنس پرسیده شده ممکن است از شفافیت- یا سادگی مفهوم- متفاوتی برخوردار باشند، برخی از شفافیت بالاتر (پس به راحتی فهمیده می‌شوند و شمارش می‌شوند، مثل بودن اسم سعید در شبکه اجتماعی خود- و برخی هم از شفافیت پایین‌تر- پس به راحتی فهمیده نمی‌شوند و در شمارش یا شناسایی آن دچار مشکل خواهند شد، مثلاً شمارش افرادی که خواهر یا برادر دوقلو دارند-

<sup>۸</sup>.Calibration curve

<sup>۹</sup>.Actual network

<sup>۱۰</sup>.Recalled network

<sup>۱۱</sup>.Transparency

<sup>۱</sup>.Reporting bias

<sup>۲</sup>.Social desirability error

<sup>۳</sup>.Level of respect

<sup>۴</sup>.Moldova

<sup>۵</sup>.Barrier effects

<sup>۶</sup>.Non-random mixing

<sup>۷</sup>.Under-recall



۳) آن‌ها چه چیزی را عمداً به شما می‌گویند؛ ۴) این افراد چه چیزی را غیر عمدی به شما می‌گویند؛ و ۵) نتیجه‌گیری از گفته‌های مجموعه‌ای از افراد پاسخ دهنده به افراد گروه هدف (بر خلاف سایر مطالعات دیگر که نتیجه‌گیری از افراد مصاحبه شونده به خود آن‌ها خواهد بود). شما ممکن است بگویید که از برخی روش‌ها مانند تعدیل کردن بر اساس برخی متغیرها، برآوردگرهای بسط تعمیمی و یا برقراری تماس با خود افراد گروه هدف- در شرایط مورد نیاز- استفاده خواهید کرد، اما در این تعدیل کردن‌ها، نیازمند توجیه خطای تخمینی<sup>۱</sup> خواهید بود. روش‌های پیشنهادی پیشنهادی برای توجیه آن استفاده از چند نمونه گرفته شده از جمعیت عمومی برای ایجاد یک جمعیت عمومی تکرارپذیر است، که از این دو یا چند نمونه مجدد برای برآورد واریانس و برآوردگرهای بسط تعمیمی استفاده می‌شود.

**حجم نمونه**<sup>۲</sup>: سؤال این است که چه اندازه حجم نمونه برای انجام این مطالعات لازم است؟ میزان خطای ناشی از نمونه‌گیری با افزایش بیش‌تر تعداد نمونه‌ها کم‌تر خواهد شد، اما روی خطاهای مربوط به غیر نمونه‌گیری تأثیر نخواهد گذاشت. خطای مربع میانگین<sup>۳</sup> شامل دو خطای واریانس و تورش است، هر چند که تورش بسیار مهم‌تر از خطای مربوط به واریانس است. از آنجایی که در مورد تورش‌های بالقوه به‌طور چیزی نمی‌دانیم، استفاده از MSE برای کاهش تورش خیلی کاربردی نیست، اما سؤال دیگری که در مورد اندازه جمعیت پیش می‌آید این است که آیا می‌توان روش بسط شبکه‌ای را با حجم نمونه کم‌تری انجام داد؟ جواب این است که به‌صورت یک قاعده کلی، هر اندازه که گروه هدف (مثل FSW) در درون جامعه کوچک‌تر و یا نادرتر باشد، به حجم نمونه بیش‌تری به منظور گرفتن برخی از این افراد در نمونه خود، برای برآورد آن گروه هدف نیازمند است.

برخوردار هستند. شفافیت پدیده‌ها -گروه‌های رفرنس- می‌توانند در برآورد اندازه‌ی شبکه‌ی اجتماعی و در نتیجه برآورد گروه‌های هدف تأثیر بگذارند. اگر بخواهیم شبکه‌ی اجتماعی را فقط بر اساس پدیده‌ای که از شفافیت بالایی برخوردار هستند؛ تعیین کنیم، و بر اساس آن اندازه گروه‌های هدف را برآورد نماییم (مثل زیر گروه IDUs) با کم برآوردی روبرو خواهیم شد. راهحل این مشکل این است که ما اندازه‌ی شبکه‌ی اجتماعی را بر اساس پدیده‌ای -گروه رفرنسی- تخمین بزنیم، که شفافیت آن مشابه گروه هدفی باشد، که قصد برآورد آن را داریم. هم‌چنین پیشنهاد می‌شود در صورتی که از تعادل میزان شفافیت در گروه هدف و گروه رفرنس آگاهی نداریم، بهتر است گروه‌های رفرنس متفاوت و با شفافیت متفاوت انتخاب و اطلاعات آن‌ها جمع‌آوری شود.

**۷) خطای برآورد واریانس و غیر نمونه‌گیری:** خطاهای یک مطالعه به خطای نمونه‌گیری و خطای غیر نمونه‌گیری مرتبط می‌شوند. دو روش برای برآورد واریانس از یک نمونه وجود دارد، که شامل: ۱) روش تحلیلی: استفاده از فرمول‌های استاندارد که به شما فاصله اطمینان را خواهد داد، که بر اساس آن می‌توان میزان خطای واریانس را درک کرد (که برای این روش بسط شبکه‌ای خیلی کاربردی نیست)؛ و ۲) Bootstrap کردن، یک روش بسیار انعطاف پذیرتر است، که اشاره به این مورد دارد که چه اندازه نتایج نزدیک یا دورتر از آن چیزی است که می‌توانست باشد. اما باید توجه داشته باشیم که Bootstrap خطای نمونه‌گیری را پوشش خواهد داد نه خطای غیر نمونه‌گیری. در روش بسط شبکه‌ای، به هنگام استفاده از Bootstrap برای توجیه خطای نمونه‌گیری، فاصله اطمینان‌های معتبری به دست نخواهیم آورد، که این احتمالاً به علت انواع متفاوت خطاهای غیر نمونه‌گیری- خطای اثرات موانع، اثرات انتقال اطلاعات، خطای پاسخ‌دهی و ...- باشد. مطالعاتی در برزیل، آمریکا و رواندا نشان دادند، که به هنگام استفاده از Bootstrap برای برآورد واریانس، به جای ۹۵ درصد اطمینان، فقط می‌توان با ۲۵-۱۰ درصد اطمینان، عدد برآورد شده در دامنه اطمینان مورد نظر قرار گیرد، به عبارتی دیگر شانس این‌که عدد برآورد شده در فاصله اطمینان قرار بگیرد، از ۲۵-۱۰ درصد، بر اساس مطالعه‌های مختلف، متفاوت است. پیشنهاد شده است که این اتفاق به علت خطاهای غیر نمونه‌گیری زیادی است که وجود دارد. منابع خطای زیادی وجود دارد که شامل: ۱) فرد مصاحبه شونده؛ ۲) این افراد چه چیزی را می‌دانند؛

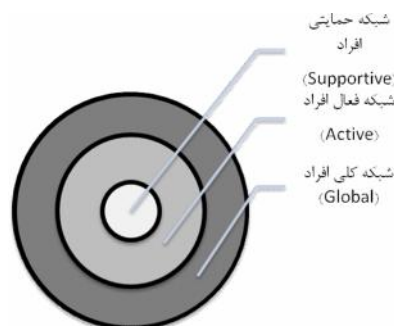
<sup>۱</sup>.Uncertainty

<sup>۲</sup>.Samoling error

<sup>۳</sup>.Mean square error

جدول شماره ۱- روش‌های برآورد اندازه جمعیت اصلی با مزایا و محدودیت‌های آن‌ها

| روش                   | چگونگی کارکردن  | ملزومات                                 | مزایا   | معایب  | سایر موارد   |
|-----------------------|---|---|---|--|--|
| سرشماری               | شمارش تمامی افراد گروه‌های پنهان  | تماس مستقیم با کل جمعیت گروه پنهان      | مستقیم و بدون واسطه   | گران و هزینه‌بر<br>اندازه را بی‌اهمیت جلوه می‌دهد  | زمانی که جمعیت پنهان از نظر جغرافیایی پراکنده است، به خوبی انجام نمی‌شود   |
| شمارش                 | شمارش کسری از گروه پنهان  | تماس مستقیم با بخشی از جمعیت گروه پنهان | مستقیم و بدون واسطه   | گران و هزینه‌بر<br>اندازه را بی‌اهمیت جلوه می‌دهد  | گروهی که مورد مطالعه قرار می‌گیرد باید یک نمونه تصادفی از جمعیت پنهان باشد |
| صید- باز<br>صید       | تعداد تصاویر در دو مرتبه  | دو تصویر مستقل از گروه پنهان            | استفاده آسان و به کار بردن فرمول  | متکی بر فرضیاتی است که برقرار بودن آن‌ها مشکل است.<br>منابع ممکن است نادرست باشند                    | -  |
| ارجاع                 | تماس با بخش قابل مشاهده و پرسش از آن‌ها برای ارائه اطلاعات از دیگر اعضا | دسترسی محدود به گروه پنهان              | مفید برای فعالیت‌های پیش‌نظارتی   | بخش قابل مشاهده ممکن است دیگر اعضا را گزارش ندهند، و با دیگران متفاوت باشند.<br>انتخاب همراه با تورش | -  |
| مدل<br>ضریبی          | مقایسه دو منبع مستقل داده‌ها برای گروه‌های پنهان                        | دو منبع مستقل داده در مورد گروه پنهان   | در صورتی که داده‌ها در دسترس باشند؛ ساده است  | نیاز به منابع مستقل داده با تعریف مشابهی از جمعیت دارد.<br>منابع ممکن است نادرست باشند               | به میزان زیادی به کیفیت داده‌ها بستگی دارد                                 |
| روش<br>بسط<br>شبکه‌ای | فراوانی رفتارهای پرخطر در اندازه شبکه‌ای از پاسخگو                      | اندازه‌ی شبکه باید شناخته شده باشد      | به منظور برآورد اندازه بسیاری از گروه‌ها در یک مطالعه نیاز به تماس با گروه‌های پنهان نیست | شفافیت و اندازه خطای حصار بسیاری از گروه‌های مرجع مورد نیاز است                                      | شفافیت اجتماعی باید برآورد شود   |



نمودار شماره ۱- مقایسه سه نوع شبکه‌ی حمایتی، فعال و کلی از لحاظ اندازه‌های ممکن

- Tate JE, Hudgens MG. Estimating population size with two- and three-stage sampling designs. *American journal of epidemiology* 2007; 165: 1314-1320.
- Vandepitte J, Lyerla R, Dallabetta G, Crabbé F, Alary M, Buvé A. Estimates of the number of female sex workers in different regions of the world. *Sexually Transmitted Infections* 2006; 82: iii18-iii25.
- Zhang D, Wang L, Lv F, Su W, Liu Y, Shen R, et al. Advantages and challenges of using census and multiplier methods to estimate the number of female sex workers in a Chinese city. *AIDS care* 2007; 19: 17-19.
- Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. *Epidemiologic Reviews* 1995; 17: 243-264.
- Laska EM Meisner M. A plant-capture method for estimating the size of a population from a single sample. *Biometrics* 1993: 209-220.
- Vadivoo S, Gupte MD, Adhikary R, Kohli A, Kangusamy B, Joshua V, et al. Appropriateness and execution challenges of three formal size estimation methods for high-risk populations in India. *Aids* 2008; 22: S137-S148.
- Stimson GV, Hickman M, Rhodes T, Bastos F, Saidel T. Methods for assessing HIV and HIV risk among IDUs and for evaluating interventions. *International Journal of Drug Policy* 2005; 16: 7-20.
- Hickman M, Taylor C, Chatterjee A, Degenhardt L, Frischer M, Hay G, et al. Estimating drug prevalence: Review of methods with special reference to developing countries. *UN Bulletin of Narcotics* 2003; 54: 15-32.
- Pisani E. Estimating the Size of Populations at Risk for HIV: Issues and Methods updated July 2003. 2003: UNAIDS/WHO working group on HIV/AIDS/STI surveillance. 12.
- Paniotto V, Petrenko T, Kupriyanov O, Pakhok O. Estimating the size of populations with high risk for HIV using the network scale-up method. Ukraine: Kiev International Institute of Sociology 2009.
- Rees Davis W, Johnson BD, Randolph D, Liberty HJ. An enumeration method of determining the prevalence of users and operatives of cocaine and heroin in Central Harlem. *Drug and alcohol dependence* 2003; 72: 45-58.
- Guidelines on estimation the size of population most at risk to hiv, UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance.
- Bernard HR, Hallett T, Iovita A, Johnsen EC, Lyerla R, McCarty C, et al. Counting hard-to-count populations: the network scale-up method for public health. *Sexually Transmitted Infections* 2010; 86: ii11-ii15.
- Killworth PD, Johnsen EC, McCarty C, Shelley GA, Bernard HR. A social network approach to estimating seroprevalence in the United States. *Social Networks* 1998; 20: 23-50.
- Killworth PD, McCarty C, Bernard HR, Shelley GA, Johnsen EC. Estimation of seroprevalence, rape, and homelessness in the United States using a social network approach. *Evaluation Review* 1998; 22: 289-308.
- Kadushin C, Killworth PD, Bernard HR, Beveridge AA. Scale-up methods as applied to estimates of heroin use. *Journal of Drug Issues* 2006; 36: 417-440.
- McCarty C, Killworth PD, Bernard HR, Johnsen EC, Shelley GA. Comparing two methods for estimating network size. *Human Organization* 2001; 60: 28-39.
- Shokoohi M, Baneshi MR, Haghdoost AA. Estimation of the active network size of Kermanian Males. *Journal of Addiction and Health* 2010; 2: 81-88.
- Rastegari A, Haji-Maghsoudi S, Haghdoost AA, Shatti M, Tarjoman T, Baneshi MR. The Estimation of Active Social Network Size of the Iranian Population. *Global Journal of Health Science* 2013; 5: 217-227.
- Shokoohi M, Baneshi MR, Haghdoost AA. Size Estimation of Groups at High Risk of HIV/AIDS using Network Scale Up in Kerman, Iran. *International Journal of Preventive Medicine* 2012; 3: 471.
- Haghdoost AA, Baeshi MR, Mohebbi E, Haji-maghsoudi S, Rastegari A. Estimation of the prevalence of Alzheimer's , Parkinson's , Multiple sclerosis an Epilepsy using Network Scale up Method in Kerman. 2013: [Unpublished Report; kerman University of Medical Sciences]
- Mohebbi E, Baeshi MR, Haji-maghsoudi S, Haghdoost AA, . The application of Network Scale up to estimate the prevalence of disabilities ; 2014 [Under Review on Journal of Research in Health Sciences] :
- Haghdoost AA, Baeshi MR, Mohebbi E, Haji-maghsoudi S, Rastegari A. Estimation of the prevalence of Breast , ovary and uterine, prostate, Bladder Cancers using Network Scale up Method in Kerman. 2013: [Unpublished Report; kerman University of Medical Sciences]
- Rastegari A, Haghdoost AA, Baneshi MR, Nakhvae N, Eslami M, Malekafzali H, et al. Estimation of the annual incidence of abortions in Iran applying Network Scale Up approach. 2014 [Accepted on Iranian Red Crescent Medical Journal]
- Nikfarjam A, Shokoohi M, Haghdoost AA, Baneshi MR, Maghsoudi S, Rastegari A, et al. National Survey on Population Size Estimation of Illicit Drug users through the Network Scale-up Method in Iran. 2014 [Under Review on Drug and Alcohol Dependence]
- Killworth PD, Johnsen EC, Bernard HR, Ann Shelley G, McCarty C. Estimating the size of personal networks. *Social Networks* 1990; 12: 289-312.
- Brewer DD. Forgetting in the recall-based elicitation of personal and social networks. *Social Networks* 2000; 22: 29-43.
- Johnsen EC, Bernard HR, Killworth PD, Shelley GA, McCarty C. A social network approach to corroborating the number of AIDS/HIV+ victims in the US. *Social Networks* 1995; 17: 167-187.
- Report from the consultation on Network scale-up and other size estimation methods from general population surveys: TECHNICAL REPORT AND RECOMMENDATIONS. 2012, UNAIDS & the US Office of the Global AIDS Coordinator: New York City.
- Huang Y, Henderson GE, Pan S, Cohen MS. HIV/AIDS risk among brothel-based female sex workers in China: Assessing the terms, content, and knowledge of sex work. *Sexually transmitted diseases* 2004; 31: 695-700.
- McCormick TH, Salganik MJ, Zheng T. How many people do you know?: Efficiently estimating personal network size. *Journal of the American Statistical Association* 2010; 105: 59-70.

**Review Article**

# The Introduction of Network Scale-up Method: An Indirect Method to Estimate the Hard-to-Reach Populations

Shokoohi M<sup>1</sup>, Mohebbi E S<sup>2</sup>, Rastegari A<sup>3</sup>, Hajimaghsoudi S<sup>4</sup>, Haghdoost AA<sup>5</sup>, Baneshi MR<sup>6</sup>

1- MS in Epidemiology, Regional Knowledge Hub, and WHO Collaborating Centre for HIV Surveillance, Kerman University of Medical Sciences, Kerman, Iran

2- DVM, Research Center for Health Services Management, Institute of Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran

3- MS in Biostatistics, Research Center for Health Services Management, Institute for Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran

4- MS in Biostatistics, Medical Student Research Committee, Kerman University of Medical Sciences, Kerman, Iran

5- PhD in Epidemiology, Research Center for Modeling in Health, AND Regional Knowledge Hub, and WHO Collaborating Centre for HIV Surveillance, Institute of Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran

6- PhD in Biostatistics, Social Determinant of Health Research Center, Institute of Futures Studies in Health, Kerman University of Medical Sciences, Kerman, Iran

**Corresponding author:** Haghdoost AA., ahaghdoost@kmu.ac.ir

Knowing the population size of rare diseases or special subpopulations like injection drug users (IDUs) is one of the most important challenges in public health and health surveillance systems but it is difficult to estimate these groups. During the last few years, new methods have been suggested to estimate hidden or hard-to-reach populations, one of which is the network scale-up method (NSUM). The NSUM itself includes measuring the personal network size and estimating the prevalence of hidden and hard-to-count populations. In this paper, we basically discussed the indirect methods of calculating the population size, and the history of NSUM and its concepts, and then addressed the estimation of hidden populations with NSUM and the applicable notes for such populations.

**Keywords:** Indirect methods, Network scale-up, Personal social network, Methodology, Hidden sub-populations