

عوامل مؤثر بر فقر با استفاده از شبکه‌های بیزی

پریسا نیلوفر*، مجتبی گنجعلی**

طرح مسأله: در این مقاله به بررسی علل و چگونگی تأثیر عوامل مؤثر بر فقر با استفاده از شبکه‌های بیزی پرداخته شده است.

روش: با روش نمونه‌گیری تصافی داده‌های ۵۰۰ خانوار از طرح هزینه و درآمد خانوارهای شهری انتخاب شد و مورد استفاده قرار گرفت. به منظور بررسی عوامل مؤثر بر فقر خانوارهای ایرانی، شبکه بیزی مناسب برای داده‌های فقر یافته شده و در تحلیل تشخیصی مورد استفاده قرار گرفته است. برخی جنبه‌های استدلالی (استدلال بالا - پایین و استدلال پایین - بالا) که با استفاده از آن‌ها عوامل مؤثر بر فقر بازنمایی می‌شوند نیز معرفی شده است.

یافته‌ها: سطح تحصیلات و بعد خانوار مهم‌ترین عوامل در مقوله تعیین سطح فقر خانوار می‌باشند. همچنین با استفاده از استدلال پایین - بالا دریافتیم در صورتی که خانواری فقیر مطلق محسوب شود، محتمل‌ترین علت برای آن سطح تحصیلات پایین یا همان بی‌سوادی سرپرست خانوار است.

نتایج: نسبت سرپرستان خانوار بی‌سواد (همچنین سرپرستان خانوار با سطح تحصیلات پایین) در کل جامعه با نسبت بی‌سوادان (همچنین سرپرستان خانوار با سطح تحصیلات پایین) در خانوارهای با فقر مطلق کاملاً متفاوت‌اند.

کلید واژه‌ها: استدلال بالا - پایین، استدلال پایین - بالا، شبکه‌های بیزی، فقر، گراف سودار غیردوری.

تاریخ پذیرش: ۸۷/۵/۷

تاریخ دریافت: ۸۶/۷/۲۰

* دانشجوی دکتری آمار دانشگاه شهید بهشتی <pniloofar@yahoo.com>

** دکتر آمار، عضو هیأت علمی دانشگاه شهید بهشتی

مقدمه

تعریف اقتصادی ارائه شده برای فقر یکی از متداول‌ترین تعاریف آن است که فقر را به دو صورت مطلق^۱ و نسبی^۲ بیان می‌کند. فقر مطلق بر این اساس تعریف می‌شود که هر فرد، حداقل نیازهای پایه‌ای برای معیشت دارد که در صورت عدم تأمین آن‌ها دچار مشکل می‌شود. در نتیجه، خط فقر مطلق عبارت است از درآمد یا هزینه قابل تصرف و قابل بیان برای رفع نیازهای واقعی، که می‌باید در زمینه‌های فرهنگی و اجتماعی و اقتصادی معین برای رفع نیازهای قطعاً ضروری در اختیار داشت (رئیس دانا، ۱۳۸۴: ۵۷). اما فقر نسبی به صورت ناتوانی در کسب متوسط سطح زندگی افراد یک جامعه تعریف شده است (Kakwani, 2003: 1). بنابراین فقر نسبی بیش از میزان مطلق درآمد افراد، با نابرابری توزیع درآمد در جامعه مرتبط است و عموماً با میانه درآمد (هزینه) کل در یک کشور ارتباط دارد (راغفر و ابراهیمی، ۱۳۸۶: ۶۰).

تاکنون تعاریفی متفاوت برای فقر و انواع آن ارائه شده است. یک تعریف جدیدتر از فقر که در دهه ۸۰ میلادی از سوی آمارتیا سن (۱۹۸۱) ارائه شد بر قابلیت دسترسی افراد به امکانات و خدمات تأکید دارد و آن را فقر قابلیت تعریف می‌کند. در این میان دیویس و مور (Davis, and Moore, 1945) از اولین نظریه‌پردازان علوم اجتماعی هستند که آشکارا به موضوع دیدگاه یا مدل کارکردگرایی فقر اشاره کرده‌اند و بعدها اسکار لوئیس (Lewis, 1996) مفهوم فرهنگ فقر را معرفی کرد. از طرفی فقر ابعاد گوناگونی دارد برای مثال می‌تواند به نوع خاصی از مصرف مربوط باشد؛ بعضی افراد ممکن است دچار فقر مسکن، بعضی دچار فقر غذایی و یا فقر سلامتی باشند.

همچنان که تعاریف مختلفی از فقر ارائه شده است، اندازه‌گیری‌های متفاوتی نیز برای خط فقر وجود دارد. سرفردریک ادن در قرن ۱۸ نخستین تلاش را برای اندازه‌گیری خط فقر انجام داد. برای مطالعه بیشتر تر در این زمینه می‌توان به ادن (Eden, 1797) مراجعه کرد.

1. Absolute Poverty
2. Relative Poverty

(۱) بیان مسأله

تاریخچه تعریف و اندازه فقر بسیار وسیع و بحث برانگیز است و سابقه آن به قبل از به وجود آمدن علم اقتصاد کنونی برمی‌گردد (Sen, 1983; Sen, 1981). آنچه مسلم است، تعریف فقر برای ملل مختلف تفاوت‌های قابل توجهی دارد. برای مثال کشورهای غنی نسبت به کشورهای فقیر عموماً استانداردهای سبک‌تری برای فقر دارند.

اولین گام در مطالعه فقر، تعیین و معرفی شاخص یا معیار سنجش فقر است. سه برداشت از شاخص تعیین فقر معرفی شده است که عبارت‌اند از سبدی از نیازهای اولیه، هزینه کل و درآمد کل. معیارهای بودجه‌ای از جمله روش‌هایی هستند که از ابتدا در اندازه‌گیری فقر به کار می‌رفته و هم‌اکنون نیز مورد توجه قرار می‌گیرند. اصولاً اندازه فقر براساس هزینه سبدی از کالاهای توافق شده تعیین می‌شود. در این روش با مراجعه به ترکیب خانوار و سطح هزینه‌ای آنها، مدل فقر را معرفی می‌کنند. این روشی است که ما در این مقاله برای تقسیم اعضای جامعه نمونه‌گیری شده به فقیر، شبه فقیر، غیرفقیر و شبه غنی یا غنی استفاده می‌کنیم.

مطالعاتی که فقر را در ایران اندازه‌گیری می‌کنند از هزینه و درآمد خانوارها که هر سال توسط مرکز آمار ایران جمع‌آوری می‌شود، استفاده می‌کند. در این مطالعات از خانوارها در مورد مخارج‌شان در ۳۰ روز و یا ۱۲ ماه اخیر سوالاتی می‌شود. در ایران اغلب نرخ فقر به صورت نسبی از خانوارها که زیر خط فقر زندگی می‌کنند، تعریف می‌شود.

بخش عمده‌ای از مطالعه پویش فقر، صرف پیدا کردن عوامل و متغیرهای اساسی است که در روند سطوح فقر تأثیرگذار می‌باشند (تقوی و نخجوانی، ۱۳۷۹). یکی از ابزارهایی که می‌توان از آن در یافتن این متغیرهای اساسی استفاده کرد، شبکه بیزی است. تا کنون از روش‌های آماری و مدل‌های گوناگونی برای اندازه‌گیری و مقایسه سطح فقر خانوارها استفاده شده است. در این میان، بررسی عواملی که می‌توانند بر سطح فقر یک خانوار ایرانی اثرگذار باشند، با استفاده از شبکه‌های بیزی، روشی نو در برخورد با این‌گونه مسائل است. با استفاده از این روش می‌توان نوع و قدرت رابطه بین مجموعه‌ای از عوامل (متغیرها) را به صورت هم‌زمان بررسی کرد که در مسأله سطح فقر، یکی از این عوامل یا متغیرها سطح فقر است و

دیگر متغیرها عواملی هستند که می‌توانند هر کدام به نوعی در اندازه‌گیری فقر مؤثر باشند. به این ترتیب با به‌کارگیری روش شبکه‌های بیزی می‌توان به پاسخی برای پرسش‌هایی از قبیل «آیا سطح فقر به بعد خانوار بستگی دارد؟»، «آیا نوع وضعیت مسکن بر سطح فقر اثرگذار است؟» و «آیا بی‌کار بودن باعث ازدیاد فقر می‌شود؟» دست یافت.

۲) چارچوب نظری

در این بخش سعی شده است با ارائه مطالبی روش به‌کار گرفته شده در انجام تجزیه و تحلیل‌ها یا همان شبکه‌های بیزی، بیش‌تر مورد بررسی قرار گیرد.

۲-۱) شبکه‌های بیزی

شبکه‌های بیزی در بررسی مسأله‌هایی با بعد زیاد بسیار سودمند واقع شده‌اند. این نوع شبکه‌ها پیچیدگی مسأله تحت مطالعه را کاهش می‌دهند. این کاهش، با ارائه روابط توأم بین مجموعه‌ای از متغیرها به‌وسیله روابط شرطی بین زیر مجموعه‌هایی از این متغیرها انجام می‌شود. اما با وجود تمامی مزیت‌هایی که شبکه‌های بیزی در تحلیل مسائل مختلف دارند مانند تمامی روش‌های دیگر، دارای محدودیت‌هایی نیز هستند. روش‌های مختلفی برای به‌دست آوردن شبکه بیزی از روی داده‌ها وجود دارد؛ از جمله آن‌ها روش‌های آماری است که در این مقاله از آن استفاده شده است؛ با وجود این، راهی که بتوان آن را بهترین روش برای مواجهه با انواع مسأله‌ها نام برد معرفی نشده است. شبکه‌های بیزی برای استفاده در سامانه‌هایی با حجم زیاد متغیرها وقتی تعداد ترکیبات متغیرها زیاد است عملی نیستند. این شبکه‌ها نیاز به نرم‌افزارهای مخصوص دارند که یا صرفه اقتصادی ندارد، یا به تخصص رایانه‌ای نیاز دارد. مطالعه شبکه‌های بیزی نیاز به دانش احتمالاتی دارد و گاهی بررسی جواب‌های شبکه مشکل است. هم‌چنین اگر متغیرها تاثیر متقابل بر هم داشته باشند این روش توانایی پاسخ‌گویی به آن را ندارد.

شبکه‌های بیزی در اصل نوعی فرمول‌بندی است که وابستگی میان متغیرها را به

صورت یک گراف و در یک زمینه چند متغیره نمایش می‌دهد. این نوع فرمول‌بندی تعریف کارایی از توزیع احتمال توأم برای هر زیرمجموعه‌ای از این متغیرها را امکان‌پذیر می‌سازد. شبکه‌های بیزی در ابتدا در سامانه‌های خبره مورد استفاده قرار گرفتند (Pearl, 1988). امروزه کاربردهای این شبکه‌ها شامل ژنتیک و استنباط‌های جنایی (Dawid, 2002) داروسازی (Rasmussen, 1995) و عیب‌یابی کامپیوتر (Heckerman, et al., 1994) است. کاربردهای شبکه‌های بیزی در آمار رسمی که کسانی مانند رامونی و سباستیان (Ramoni & Sebastiani, 2001) و هم‌چنین گتور و همکارانش (Getoor, et al.) در سال ۲۰۰۱ روی آن کار کرده‌اند، بسیار جدیدتر می‌باشند.

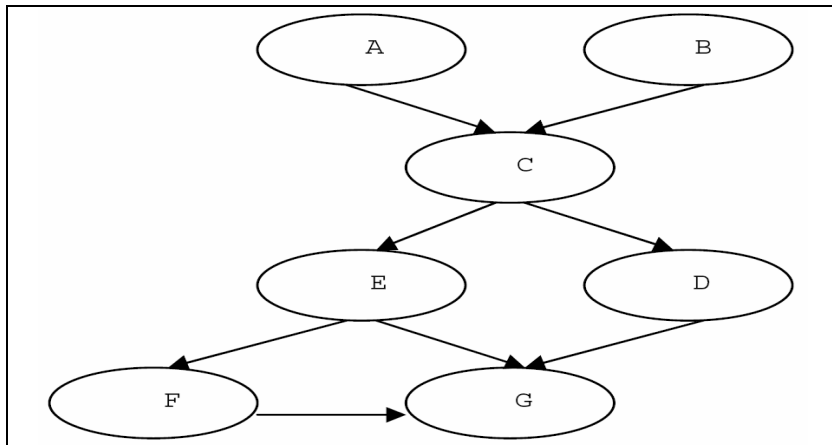
طبق تعریفی که یسن (Jensen, 1996) برای شبکه‌های بیزی ارائه داده است، این شبکه‌ها با سه عامل زیر تعریف می‌شوند:

الف) گره‌ها: هر گره نشان‌دهنده یک متغیر با تعداد متناهی حالت است.

ب) یال جهت‌دار: هر یال دو گره را به یکدیگر متصل می‌کند.

ج) هر متغیر یک توزیع احتمال شرطی دارد.

شکل ۱: نمایش یک شبکه بیزی با ۷ متغیر



دو عامل اول، ساختار شبکه‌های بیزی را مشخص می‌کنند. برای مثال هنگامی که دو متغیر A و B به وسیله یک یال به هم متصل‌اند، دو متغیر به طور احتمالی به هم وابسته‌اند. جهت یال‌ها عموماً به عنوان رابطه سببی بین دو متغیر تعبیر می‌شود. برای مثال هنگامی که V متغیر را مورد تجزیه و تحلیل قرار می‌دهیم، یکی از حالت‌های ممکن به صورت شکل شماره ۱ خواهد بود.

گره‌هایی که بلافاصله قبل از یک گره قرار دارند، والدین آن گره می‌باشند. برای مثال در شکل شماره ۱، A و B والدین C هستند که از این به بعد این مطلب را با $Pa(C) = \{A, B\}$ نشان می‌دهیم. در حالی که C والد D و E است. گره‌هایی که بعد از یک گره قرار دارند فرزندان می‌باشند. برای مثال G, F, E و D فرزندان C هستند. توجه داشته باشید که جهت یال‌ها هیچ نوع دوری را به وجود نمی‌آورد. این ساختار (گره‌ها و یال‌های سودار) معمولاً گراف‌های سودار غیر دوری (DAG)^۱ نامیده می‌شوند.

سومین عامل در تعریف شبکه‌های بیزی، نشان می‌دهد که هر گره با یک توزیع احتمال مشخص می‌شود که همان توزیع حاشیه‌ای گره به شرط والدین آن است. برای مثال در شکل شماره ۱ شبکه بیزی به گره C توزیع احتمال متغیر C ، به شرط متغیرهای $\{A, B\}$ را تخصیص می‌دهد. در این حالت شبکه بیزی، مدل وابستگی خاصی را میان متغیرها نشان می‌دهد، به طوری که یک شرط لازم و کافی برای این که یک گراف یک شبکه بیزی باشد این است که هر متغیر X به شرط والدینش، $Pa(X)$ ، مستقل از تمام متغیرهای غیر فرزندش باشد (Pearl, 1988). در نتیجه، شرطی کردن روی متغیر C در شکل شماره ۱، متغیر E را از متغیرهای A و B مستقل می‌سازد. به عبارت دیگر، ساختار احتمالی E فقط به اطلاعاتی راجع به C نیازمند است.

به این ترتیب با فرض معلوم بودن C ، سایر متغیرها در قسمت بالای گراف (غیر فرزندان) غیر لازم هستند.

به این ترتیب توزیع توأم کامل متغیرهای X_1, X_2, \dots, X_k به وسیله توزیع‌های شرطی

1. Directed Acyclic Graph

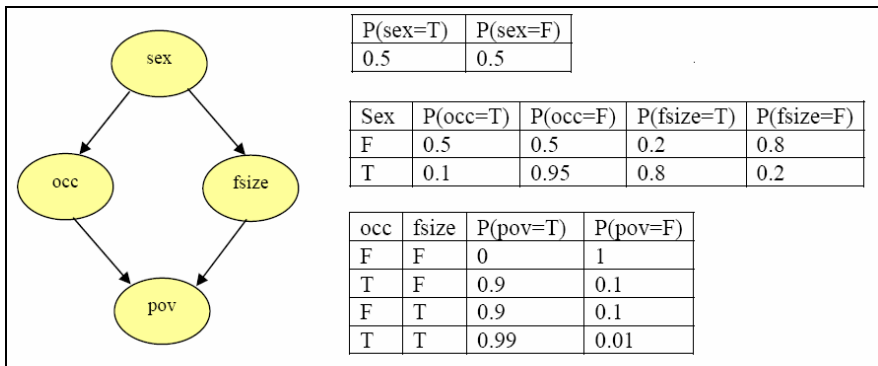
یک متغیره‌ای که به هر یک از گره‌ها در شبکه بیزی اختصاص یافته، تعیین می‌شود؛ یعنی طبق خاصیت زنجیری پیرل:

$$P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \prod_{i=1}^k P(X_i = x_i | Pa(X_i))$$

فرمول بالا دلیل اصلی فرمول‌بندی توزیع چند متغیره به وسیله شبکه‌های بیزی می‌باشد. تعریف توزیع چند متغیره توأم با حاصلضرب توزیع‌های شرطی تکی، باعث یک کاهش چشم‌گیر در تعداد پارامترها می‌شود. در نتیجه، برآورد کردن بسیار ساده خواهد بود. در واقع دلیل اصلی نام‌گذاری این شبکه‌ها به نام «شبکه‌های بیزی» استفاده از فرمول بیز برای به دست آوردن احتمال شرطی $P(X_i = x_i | Pa(X_i))$ است و هیچ نوع ارتباطی بین شبکه‌های بیزی و رویکرد بیزی که در آن اطلاعات پیشین نقشی مهم در تصمیم‌گیری ایفا می‌کند، وجود ندارد.

برای درک بهتر آنچه در مورد شبکه‌های بیزی آورده شد، ابتدا صورت بسیار ساده‌ای از شبکه بیزی‌ای را نشان می‌دهیم که می‌تواند به فهم مطالبی که در ادامه خواهد آمد، کمک کند.

شکل ۲: ساختار شبکه بیزی با ۴ گره و جداول احتمال مربوط به آن



شبکه‌ای را متشکل از ۴ گره دودویی در نظر بگیرید، یعنی هر گره دو حالت ممکن دارد که در این مثال با T و F نشان می‌دهیم. در این شبکه که در شکل شماره ۲ نشان داده شده است، گره sex نشان‌دهنده جنس می‌باشد، به این ترتیب که هرگاه $sex=T$ است یعنی فرد مورد نظر مرد است و هرگاه $sex=F$ یعنی فرد مورد نظر زن است. گره occ نشان‌دهنده وضعیت اشتغال است و هنگامی که در وضعیت F باشد یعنی فرد شاغل است و در غیر این صورت، بیکار است. در مورد گره $fsize$ که نشان‌دهنده بعد خانوار است، F مربوط به خانواری با ۴ یا کمتر از ۴ عضو و T مربوط به خانواری با بیش‌تر از ۴ عضو است؛ و گره pov که همان گره مربوط به متغیر سطح فقر است نیز دارای دو حالت فقیر و غنی است.

می‌توان از جدول سوم در شکل شماره ۲ که مربوط به احتمالات فقر به شرط بعد خانوار و وضعیت اشتغال سرپرست خانوار است، این‌طور استنباط کرد که بعد خانوار زیاد (در این‌جا منظور بالای ۴ نفر، $fsize=T$) و بیکاری سرپرست خانوار ($occ=T$) محتمل‌ترین علت‌ها برای پیشامد «فقیر بودن» ($pov=T$) محسوب می‌شوند. قدرت این رابطه در جداول شکل شماره ۲ آمده است.

با توجه به قانون زنجیره احتمال، احتمال توأم تمامی گره‌ها در شبکه بیزی بالا به صورت زیر است:

$$P(\text{sex}, \text{fsize}, \text{occ}, \text{pov}) = P(\text{sex}) \times P(\text{occ}|\text{sex}) \times P(\text{fsize}|\text{sex}, \text{occ}) \times P(\text{pov}|\text{sex}, \text{fsize}, \text{occ})$$

با استفاده از روابط استقلال شرطی می‌توان نوشت:

1. True
2. False
3. Occupation
4. Family Size
5. Poverty

$$P(\text{sex}, \text{fsize}, \text{occ}, \text{pov}) = P(\text{sex}) \times P(\text{occ}|\text{sex}) \times P(\text{fsize}|\text{sex}) \times P(\text{pov}|\text{fsize}, \text{occ})$$

عبارت سوم سمت راست معادله بالا به این دلیل که fsize به شرط والدش sex، از occ مستقل است، به این صورت نوشته شده؛ و در مورد عبارت چهارم سمت راست معادله مذکور نیز می‌توان گفت علت کاهش آن از $P(\text{pov}|\text{sex}, \text{fsize}, \text{occ})$ به $P(\text{pov}|\text{fsize}, \text{occ})$ این است که گره pov به شرط والدینش occ و fsize از گره sex مستقل است.

۲-۲) انواع استنباط در شبکه‌های بیزی

مرسوم‌ترین مسأله‌ای که امیدواریم با شبکه‌های بیزی حل کنیم استنباط احتمالی است. در مثال قبل دیدیم که دو دلیل برای فقیر محسوب شدن یک خانوار وجود داشت: اول، زیاد بودن بعد خانوار و دوم، بیکار بودن سرپرست خانوار.

سؤال اصلی این است که کدام یک از این دلایل محتمل‌تر است؟ برای پاسخ دادن به این پرسش به این صورت عمل می‌کنیم:

فرض کنیم غلط بودن یک پیشامد، یعنی اتفاق نیفتادن آن را با ۰ و درست بودن، یعنی به‌وقوع پیوستن آن را با ۱ نشان دهیم، در این صورت:

$$\begin{aligned} P(\text{occ}=1|\text{pov}=1) &= \frac{P(\text{occ}=1, \text{pov}=1)}{P(\text{pov}=1)} \\ &= \frac{\sum_{s,f} P(\text{sex}=s, \text{occ}=1, \text{fsize}=f, \text{pov}=1)}{P(\text{pov}=1)} = \frac{0/2781}{0/6471} = 0/4297 \end{aligned}$$

به همین ترتیب:

$$P(\text{fsize}=1|\text{pov}=1) = 0/6891$$

که در آنها

$$P(\text{pov}=1) = \sum_{s,o,f} P(\text{sex}=s, \text{occ}=o, \text{fsize}=f, \text{pov}=1) = 0/6471$$

بنابراین اگر خانواری فقیر باشد می‌بینیم که بیشتر احتمال دارد بعد آن خانوار زیاد باشد، تا این که سرپرست خانوار بی‌کار باشد. نوع دیگری از استنباط بدین صورت است که معلولی داریم و محتمل‌ترین علت را بر اساس آن استنباط می‌کنیم. به این روش، استدلال تشخیصی یا «پایین - بالا» گفته می‌شود، زیرا از معلول به سمت علت می‌رویم. این نوع استدلال در سیستم‌های خبره مرسوم است. شبکه‌های بیزی هم‌چنین می‌توانند برای استدلال سببی یا «بالا - پایین» استفاده شوند. برای مثال می‌توانیم احتمال فقیر بودن خانوار را به شرط مرد بودن سرپرست خانوار ($\text{sex}=T$) محاسبه کنیم. بنابراین شبکه‌های بیزی اغلب مدل‌های مبتکر نیز نامیده می‌شوند، زیرا این شبکه‌ها مشخص می‌کنند چگونه علت‌ها به وجود آورنده معلول‌ها هستند.

۳) روش مطالعه

در این مطالعه از داده‌های ثبت شده برای هر واحد خانوار شهری که شامل درآمد و هزینه (HIE)^۱ می‌باشد استفاده می‌کنیم. در این مطالعه، نمونه‌گیری توسط مرکز آمار ایران در سال ۱۳۸۴ انجام گرفته است و این پژوهش نوعی تحقیق بر مبنای نمونه‌گیری‌های تصادفی پیچیده با حجم قابل ملاحظه‌ای است که شامل مقدار زیادی اطلاعات از خانوارهای مورد بررسی است.

۳-۱) متغیرهای مطالعه

خط فقر مورد استفاده در این بررسی، خط فقر مطلق بر مبنای ۲۳۰۰ کالری و ۳۹۰۰۵۸ ریال به عنوان خط فقر در سال ۱۳۸۳ است (باقری و همکاران، ۱۳۸۳) که با توجه به نرخ تورم ۱۴ درصد سال ۱۳۸۴، خط فقر در سال ۱۳۸۴ به ۴۴۴۶۶۶/۱ ریال بالغ می‌شود. برای

1. Household Income and Expenditure

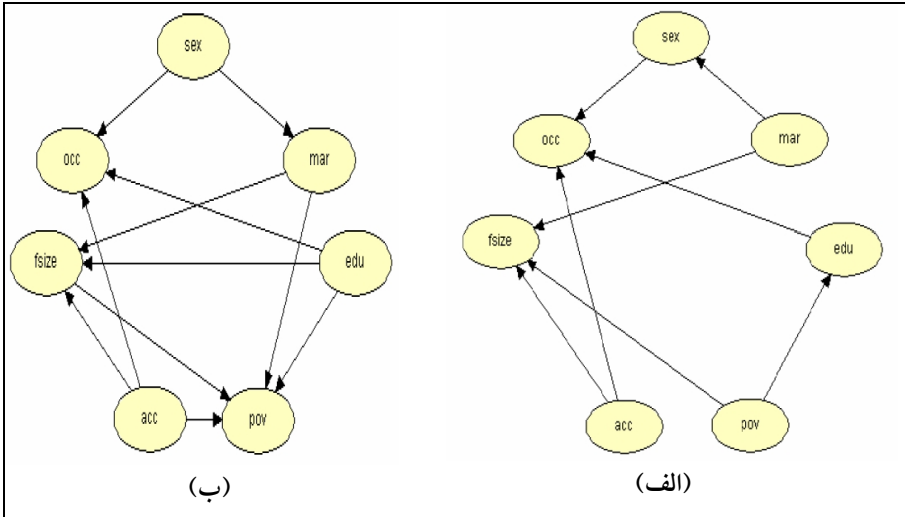
بررسی فقر به عنوان تابعی از هزینه کل، میزان هزینه کل را به صورتی که در مطالعه البرز و همکاران (۱۳۸۶: ۱۲۵) آمده است، دسته‌بندی نموده‌ایم. بنابراین، فرض می‌کنیم متغیر رده‌بندی شده سطح فقر که به صورت ترتیبی در نظر گرفته شده، یک متغیر ترتیبی است که مقدار ۱ نشان‌دهنده فقیر مطلق بودن، مقدار ۲ نشان‌دهنده شبه فقیر بودن، مقدار ۳ نشان‌دهنده غیر فقیر بودن و مقدار ۴ نشان‌دهنده شبه غنی یا غنی بودن فرد می‌باشد. متغیرهای دیگری که در نظر گرفته شده‌اند عبارت‌اند از: جنس، سطح تحصیلات، وضعیت فعالیت، وضعیت تأهل و زناشویی، وضعیت مسکن و بعد خانوار.

۲-۳) روش اجرا

همان‌طور که در بخش‌های قبل نیز به آن اشاره شد، یافتن متغیرهایی که بر سطح فقر اثرگذار هستند از اهمیت زیادی برخوردار است. روش شبکه‌های بیزی که در این پژوهش از آن استفاده شده است، به صورت هم‌زمان تأثیر تمامی عوامل را مورد بررسی قرار می‌دهد.

با در نظر گرفتن متغیر سطح فقر، ۷ متغیر ترتیبی در این مطالعه مورد بررسی قرار گرفته است. به دلیل ترتیبی بودن این متغیرها، شبکه بیزی‌ای با ۷ گره خواهیم داشت که دارای توزیع گسسته می‌باشند. به منظور یادگیری ساختار یک شبکه (یال‌ها و جهات آن‌ها) از روی داده‌های موجود از الگوریتم PC (Spirites, et al., 1993) استفاده شده است و برای یادگیری پارامترهای شبکه که در واقع همان جداول شرطی می‌باشند، الگوریتم EM به کار گرفته شده است. در این پژوهش از نرم‌افزار HUGIN که امکان پیاده‌سازی این اعمال را در اختیار کاربر قرار می‌دهد استفاده شده است. به این ترتیب شبکه بیزی اولیه مربوط به داده‌های مورد نظر به صورت شکل شماره ۳ قسمت (الف) خواهد بود.

شکل ۳: نمودار قسمت (الف) نشان‌دهنده شبکه بیزی اولیه برای داده‌های فقر و قسمت (ب) نشان‌دهنده شبکه بیزی داده‌های فقر پس از اعمال تغییر در جهت یال‌ها می‌باشد.



۳-۳ تجزیه و تحلیل داده‌ها

در این بخش با توجه به آنچه در بخش ۳ گفته شد، شبکه بیزی مربوط به داده‌های تشریح شده در بخش ۳-۱ را یافته و استنباطها و تجزیه و تحلیل‌های مناسبی انجام می‌دهیم تا مهم‌ترین علل به وجود آورنده فقر و اختلاف طبقاتی در جامعه تا حدودی مشخص شوند. به این ترتیب می‌توان استراتژی‌های لازم برای رفع و یا تعدیل این عوامل را پایه‌ریزی کرد. در شکل شماره ۳ قسمت (الف) گره sex نشان‌دهنده جنس سرپرست خانوار، mar^۱ مربوط به وضعیت تأهل، occ مربوط به وضعیت اشتغال، edu^۲ نشان‌دهنده وضعیت سواد،

1. Marriage
2. Education

fsize مربوط به بعد خانوار، acc¹ مربوط به وضعیت مسکن و pov نشان‌دهنده سطح فقر است. با توجه به مباحث نظری مربوط به یافتن ساختار در شبکه‌های بیزی، که در آن از الگوریتم PC استفاده می‌شود، جهت یال ارتباط دهنده بین بعضی متغیرها (در این مقاله یال ارتباط دهنده وضعیت تاهل و جنس) به صورت دلخواه توسط الگوریتم جهت‌دهی می‌شود و دلیل آن، این است که با استفاده از آزمون‌های آماری به کارگرفته شده در این الگوریتم، فقط وجود یال بین دو متغیر از اهمیت زیادی برخوردار است و جهت آن می‌تواند دلخواه باشد؛ اما در مورد متغیرهایی که جهت یال قویاً توسط الگوریتم تشخیص داده شده باشد، اعمال تغییر در جهت یال با استفاده از نرم افزار، یا امکان‌پذیر نمی‌باشد و یا با تغییر در جهت یال‌های دیگر همراه خواهد بود. در این مقاله، همان‌طور که می‌دانیم، متغیر جنس متغیر مستقل است؛ بنابراین به راحتی می‌توان جهت یال بین این گره و گره وضعیت تاهل را (که توسط الگوریتم به صورت دلخواه جهت‌دهی شده است) بدون هیچ مشکلی تغییر داد. همچنین به دلیل این‌که هدف یافتن عوامل مؤثر بر سطح فقر است، بنابراین با تغییر جهت یال‌های مربوط به این گره‌ها که با تغییر جهت در بعضی یال‌های دیگر و اضافه شدن چند یال به شبکه همراه است (Pearl & Verma, 1991)، شبکه بیزی موجود در شکل شماره ۳ قسمت (ب) حاصل خواهد شد. مشاهده می‌کنیم که هنگامی که به منظور اعمال این اهداف، جهت بعضی یال‌ها تغییر داده می‌شود، گرافی معادل حاصل می‌شود که دارای یال‌های بیشتر (Pearl & Verma, 1991) و در نتیجه پیچیده‌تر خواهد بود و این عامل باعث سخت‌تر شدن محاسبات می‌شود.

پس از به دست آوردن شبکه بیزی مربوط به این داده‌ها با توجه به مطالبی که در بخش ۲-۲ گفته شد، می‌توان محتمل‌ترین عامل اثرگذار بر این‌که خانواری غنی یا فقیر محسوب شوند را محاسبه کرد.

1. Accomodation

۴ یافته‌ها

درصد افراد در رده‌های مختلف متغیر پاسخ به ترتیب عبارت از ۱۷، ۳۳/۸، ۲۶/۸۰ و ۲۲/۴۰ است. این نشان‌دهنده آن است که ۵۰/۸ درصد خانوارها فقیر یا شبه فقیرند. متغیرهای دیگر شامل خصوصیتی از سرپرست خانواده هستند که به صورت زیر می‌باشند:

الف) جنس: مرد (۱) و زن (۰) که در این داده‌ها ۹۱/۸۰ درصد مرد و ۸/۲۰ درصد زن وجود دارد.

ب) سطح تحصیلات: باسواد (۱) و بی‌سواد (۰) که در این داده‌ها ۷۹/۶۰ درصد باسواد و ۲۰/۴۰ درصد بی‌سوادند.

ج) وضعیت فعالیت: شاغل (۱)، بیکار جویای کار (۲)، دانشجویان بدون کار (۳) و سایر (۴) که در این داده‌ها ۷۷/۸۰ درصد شاغل، ۱/۴۰ درصد بیکار جویای کار، ۱۷/۴۰ درصد دانشجویان بیکار و ۳/۴۰ درصد سایر می‌باشند.

د) وضعیت تأهل و زناشویی: دارای همسر (۱)، بی‌همسر بر اثر فوت (۲)، بی‌همسر بر اثر طلاق (۳) و هرگز ازدواج نکرده (۴) که در این داده‌ها ۸۹/۴۰ درصد دارای همسر، ۸/۲۰ درصد بی‌همسر بر اثر فوت، ۱/۲۰ درصد بی‌همسر بر اثر طلاق و ۱/۲۰ درصد هرگز ازدواج نکرده‌اند.

و همچنین خصوصیتی از خانواده از جمله:

ه) وضعیت مسکن: مالک (۱)، اجاره یا رهن (۲) و سایر (۳) که در این داده‌ها ۶۸/۶۰ درصد مالک، ۲۳ درصد مسکن اجاره‌ای و ۸/۴ درصد سایر می‌باشند.

و) بعد خانوار: یک نفر (۱)، دو نفر (۲) و ... و پنج نفر (۵) و حداقل ۶ نفر (۶) که در این داده‌ها ۵/۲ درصد خانواده‌ها یک نفره، ۱۲/۲ درصد دو نفره، ۱۸/۶۰ درصد سه نفره، ۲۱/۶۰ درصد چهار نفره، ۱۷/۴۰ درصد پنج نفره و ۲۵ درصد حداقل شش نفره می‌باشند.

نرم‌افزار توضیح داده شده در بخش ۳ امکان مشاهده جداول توزیع احتمال هر گره را در اختیار کاربر قرار می‌دهد. برای مثال جدول توزیع احتمال مربوط به متغیر وضعیت تأهل به شرط متغیر جنس به صورت آنچه در شکل شماره ۴ آمده است، می‌باشد. به‌طور

مثال می‌توان مشاهده کرد که با توجه به شکل شماره ۴، احتمال این‌که فردی دارای همسر باشد ($mar=1$) به شرط این‌که مرد باشد ($sex=1$) برابر 0.971678 است، در عین حال که احتمال این‌که بر اثر طلاق بدون همسر باشد ($mar=3$) به شرط این‌که زن باشد ($sex=0$) 0.146341 است.

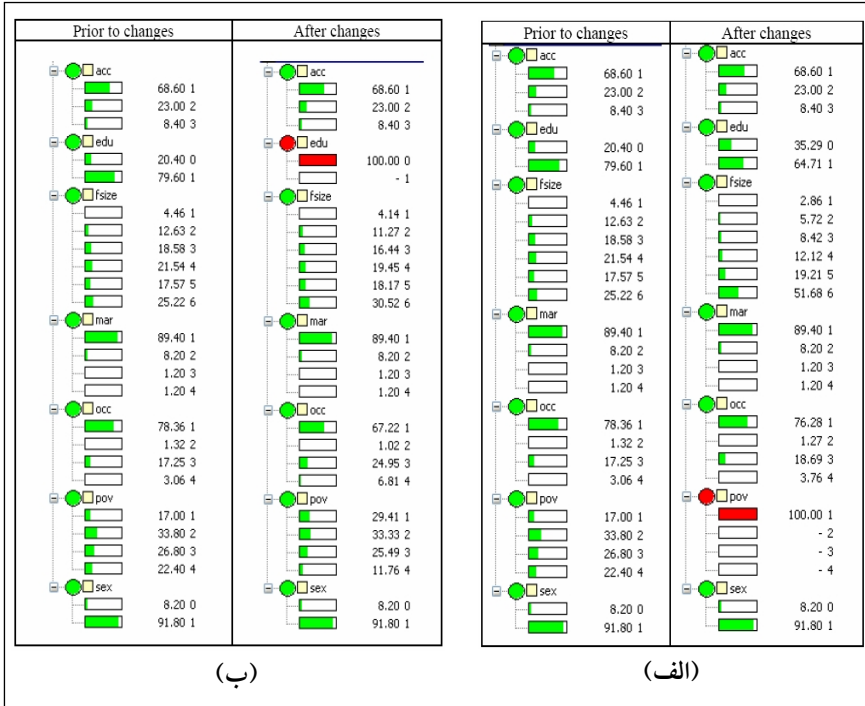
به این ترتیب با توجه به شکل شماره ۳ قسمت (ب) از بخش ۳ مشاهده می‌کنیم متغیرهای وضعیت مسکن، وضعیت سواد، وضعیت تأهل و بعد خانوار به طور مستقیم بر سطح فقر اثرگذار هستند و متغیرهای دیگر با تأثیر بر روی این متغیرها به‌طور غیرمستقیم بر سطح فقر اثر می‌گذارند.

شکل ۴: جدول توزیع احتمال متغیر وضعیت تاهل

mar		
sex	0	1
1	0.02439	0.971678
2	0.804878	0.017429
3	0.146341	0
4	0.02439	0.010893

فرض کنیم خانواری فقیر مطلق محسوب شود (به این منظور که متغیر فقر در سطح اول خود باشد، مستطیل پر شده در شکل ۵ قسمت (الف) ما را متوجه می‌کند که فرض کرده‌ایم تنها خانوارهای فقیر در نظر گرفته شده‌اند)، به این ترتیب احتمال‌های هر یک از سطوح مختلف عواملی که اثر مستقیم بر سطح فقر دارند به‌صورت آنچه در شکل شماره ۵ قسمت (الف) آمده است تغییر می‌کنند:

شکل ۵: (الف): جدول تغییرات در توزیع احتمال متغیرها - استدلال پایین بالا؛ (ب): جدول تغییرات در توزیع احتمال متغیرها - استدلال بالا پایین.



در شکل شماره ۵ قسمت (الف) ستون قبل از تغییرات^۱ (ستون سمت چپ) به معنی شرطی نکردن روی هیچ متغیری است و ستون سمت راست، بعد از تغییرات^۲ به معنی شرطی کردن روی سطح اول از متغیر فقر است. در هر قسمت اعداد سمت راست مربوط به سطوح متغیر و اعداد سمت چپ احتمال این سطوح به درصد را نشان می‌دهد.

1. Prior to changes
2. After changes

مشاهده می‌کنیم که در توزیع احتمال‌های متغیرهای جنس، وضعیت مسکن و وضعیت تأهل تغییری ایجاد نشده است. هم‌چنین در متغیر وضعیت اشتغال، تغییر قابل اغماضی صورت گرفته است؛ ولی تغییر قابل توجهی در احتمال‌های مربوط به متغیرهای وضعیت سواد و مخصوصاً بعد خانوار به وجود آمده است. به این ترتیب که مشاهده می‌کنیم احتمال مربوط به خانواری با تعداد اعضای ۶ یا بیش‌تر از ۶ نفر از $25/22$ درصد به $51/68$ درصد افزایش پیدا کرده است. هم‌چنین احتمال مربوط به خانواری با ۱ عضو از $4/46$ درصد به $2/86$ درصد کاهش پیدا کرده است. در کل، در مورد متغیر بعد خانوار می‌توان گفت تنها احتمال خانوارهایی با اعضای بیش‌تر از ۵ نفر افزایش و در بقیه سطوح کاهش در احتمال مشاهده شده است. آنچه در مورد متغیر سطح تحصیلات نیز مشاهده می‌شود جالب توجه است. احتمال این‌که بی‌سوادی نیز علتی برای فقر خانوار باشد از $20/40$ درصد به $35/29$ درصد افزایش یافته است.

استدلالی که انجام شد از نوع استدالات «پایین - بالا» محسوب می‌شود. نوع دیگری از استدلال می‌تواند به این صورت باشد که علت را معلوم فرض کنیم و در پی معلول باشیم (استدلال بالا - پایین). به این ترتیب حال فرض کنید می‌دانیم که سرپرست خانوار بی‌سواد است (سطح ۰ از متغیر edu). شکل شماره ۵ قسمت (ب) نشان‌دهنده تغییرات در توزیع احتمال باقی‌گروه‌هایی که به نوعی با این گروه به صورت مستقیم در تماس هستند، خصوصاً گروه فقر می‌باشد.

طبق آنچه در شکل شماره ۵ قسمت (ب) آمده است، مشاهده می‌کنیم که تنها در گروه‌های occ، fsize و pov تغییرات حاصل شده است. کاهش احتمال سطح ۱ از گروه وضعیت فعالیت (شاغلین) از $78/36$ به $67/22$ و افزایش احتمال سطح ۳ از این متغیر (دانشجویان بدون کار) از $17/25$ به $24/95$ ، نشان‌دهنده افزایش احتمال بیکاری برای سرپرست خانوارهایی است که بی‌سواد هستند.

در این میان، تغییراتی که در سطوح ۱ و ۴ از متغیر سطح فقر ایجاد شده جالب توجه است. افزایش احتمال سطح ۱، یعنی فقر مطلق خانوار، از ۱۷ به $29/41$ و کاهش احتمال

سطح ۴، یعنی غنی و شبه غنی، از ۲۲/۴۰ به ۱۱/۷۶ که در هر دو مورد نسبتاً زیاد است، نشان‌دهنده این امر است که بی‌سوادی سرپرست خانوار احتمال فقیر محسوب شدن خانوار را افزایش می‌دهد.

همین فرآیند را برای سه متغیر دیگر اثرگذار بر گره pov یعنی acc و $fsize$ و mar انجام می‌دهیم. تغییراتی که در سطوح مختلف از گره‌های وضعیت تأهل و وضعیت مسکن اعمال می‌شود کوچک‌ترین تغییری در احتمالات مربوط به سطوح متغیر سطح فقر ایجاد نمی‌کند. این امر نشان می‌دهد که میزان اثر این دو عامل بر متغیر سطح فقر قابل چشم‌پوشی است.

حال اگر فرض کنیم بعد خانواری بیش‌تر از ۶ است، این امر تأثیر بسزایی در احتمالات مربوط به گره سطح فقر می‌گذارد، به طوری که تعداد اعضای زیاد خانوار باعث کاهش شدید احتمال غنی بودن خانوار از ۲۲/۴۰ به ۶/۴۹ و افزایش بیش از دو برابری فقیر مطلق بودن خانوار از ۱۷ به ۳۴/۸۳ می‌شود.

بنابراین به طور خلاصه، شبکه بیزی مناسب برای تحلیل داده‌های طرح هزینه و درآمد خانوارهای شهری یافته شد و با استفاده از آن عوامل مؤثر بر فقر خانوارهای ایرانی مورد تحقیق قرار گرفت. در تحلیل‌های انجام شده در این مطالعه پی بردیم که متغیرهای وضعیت مسکن، وضعیت سواد، وضعیت تأهل و بعد خانوار به‌طور مستقیم بر سطح فقر اثرگذار هستند و متغیرهای دیگر با تأثیر بر روی این متغیرها به‌طور غیرمستقیم بر سطح فقر اثر می‌گذارند. از میان این متغیرها پی بردیم که سطح تحصیلات و بعد خانوار مهم‌ترین عوامل در مقوله تعیین سطح فقر خانوار می‌باشند. هم‌چنین با استفاده از استدلال پایین - بالا، پی بردیم در صورتی که خانواری فقیر مطلق محسوب شود، محتمل‌ترین علت برای آن، سطح تحصیلات پایین یا همان بی‌سوادی سرپرست خانوار است.

(۵) بحث و نتیجه‌گیری

پیدایش فقر به زمانی باز می‌گردد که جوامع انسانی به مازاد تولید دست یافتند و در سایه آن

سلسله مراتب اجتماعی بر اساس تمایزهای ناشی از دسترسی نابرابر به مازاد، شکل گرفت. در مطالعات مختلف، از جنبه‌های گوناگونی به بررسی عوامل تشدید کننده و یا اصولاً به وجود آورنده فقر در جوامع انسانی پرداخته شده است که از آن جمله می‌توان به موارد زیر اشاره کرد: رئیس دانا خدشه‌دار کردن بخش توسعه اجتماعی و انسانی و عدالت اجتماعی با بازارگرایی‌های افراطی و تعطیل برنامه‌ریزی و در سطح جهان، علل ازدیاد فقر و تبعیض را کار افتادن مکانیزم بازار و آرمان سیاسی نولیبرالیستی می‌داند (رئیس دانا، ۱۳۸۴)؛ میرزایی و غیاثوند دهک‌های درآمدی، بعد خانوار و نوع شغل سرپرست خانوار را در ایجاد تورم مؤثر دانسته‌اند (میرزایی و غیاثوند، ۱۳۸۶)؛ نصیری یکی از راه‌های مقابله با فقر را لزوم دخالت دولت در تأمین مسکن مطرح می‌کند (نصیری، ۱۳۸۶)؛ البرز و دیگران از دیدگاهی دیگر به این نتیجه رسیدند که شانس فقیر بودن برای زنان یا مردان با بعد خانوار ۱ بسیار کم‌تر از زنان یا مردان با بعد خانوار ۵ یا بیش‌تر است. همچنین وضعیت مستاجران با بعد خانوار بالا از نقطه نظر فقر بحرانی است (البرز و دیگران، ۱۳۸۶).

راهبرد شبکه‌های بیزی روشی است که اخیراً در مسائل مختلف از جمله پیش‌بینی و تحلیل تشخیصی به‌کار برده شده است. مقاله حاضر از شبکه‌های بیزی در مبحث فقر و بررسی عوامل تشدیدکننده آن استفاده کرده است. در این مقاله از داده‌های طرح هزینه و درآمد خانوارهای شهری در مرکز آمار ایران استفاده شده است. اطلاعات به‌دست آمده از طرح مذکور به ما این امکان را می‌دهد که مسأله فقر را با رویکردی متفاوت بررسی نماییم.

به این ترتیب به منظور بررسی عوامل مؤثر بر فقر خانوارهای ایرانی، که مربوط به مبحث تحلیل تشخیصی در مقوله شبکه‌های بیزی می‌شود، شبکه بیزی مناسب برای داده‌های فقر را به‌دست آوردیم. با استفاده از این شبکه بیزی برانده شده به داده‌های مذکور، در نگاه اول عوامل اثرگذار بر متغیر سطح فقر، سطح تحصیلات، وضعیت تأهل، بعد خانوار و وضعیت مسکن تشخیص داده شدند. ولی در تحلیل‌های ثانویه، از جمله استدلال بالا - پایین پی بردیم که وضعیت تأهل و وضعیت مسکن تأثیر چندانی بر سطح

فقر ندارند و در واقع سطح تحصیلات و بعد خانوار مهم‌ترین عوامل در مقوله تعیین سطح فقر خانوار می‌باشند. همچنین با استفاده از استدلال پایین - بالا دریافتیم در صورتی که خانواری فقیر مطلق محسوب شود، محتمل‌ترین علت برای آن سطح تحصیلات پایین یا همان بی‌سوادی سرپرست خانوار است.

- البرز، مسعود؛ افتخاری، سمانه و گنجعلی، مجتبی. (۱۳۸۶)، «عوامل تعیین‌کننده احتمال فقر افراد با استفاده از مدل لوژستیک تجمعی ترتیبی»، **فصلنامه علمی پژوهشی رفاه اجتماعی**، سال ششم، شماره ۲۴.
- باقری، فریده؛ حیدری، خلیل و پیمان، حسین. (۱۳۸۳)، **محاسبه خط فقر در ایران در طی سال‌های ۱۳۸۲-۱۳۸۰**، پژوهشکده آمار.
- تقوی، مهدی و غروی نجوانی، محمد. (۱۳۷۹)، «روند فقر در ایران طی دو دهه گذشته»، **فقر در ایران**، فریبرز رئیس دانا و دیگران، تهران، دانشگاه علوم بهزیستی و توان‌بخشی.
- رئیس دانا، فریبرز. (۱۳۸۴)، «اندازه‌گیری شاخص فقر در ایران»، **فصلنامه علمی پژوهشی رفاه اجتماعی**، سال چهارم، شماره ۱۷.
- راغفر، حسین و ابراهیمی، زهرا. (۱۳۸۶)، «فقر در ایران طی سال‌های ۱۳۸۳-۱۳۶۸»، **فصلنامه علمی پژوهشی رفاه اجتماعی**، سال ششم، شماره ۲۴.
- میرزایی، حسین و غیاثوند، ابوالفضل. (۱۳۸۶)، «عوامل اجتماعی و اقتصادی مؤثر بر نابرابری توری (۱۳۸۲-۱۳۷۴)»، **فصلنامه علمی پژوهشی رفاه اجتماعی**، سال ششم، شماره ۲۴.
- نصیری، معصومه. (۱۳۸۶)، «توزیع جغرافیایی فقر مسکن و پراکندگی زنان مطلقه خانه‌دار در مناطق ۲۲گانه تهران»، **فصلنامه علمی پژوهشی رفاه اجتماعی**، سال ششم، شماره ۲۴.
- David, A. P. (2002), "Influence Diagrams for Causal Modelling and Inference", **International Statistical Review**, 70.
- Davis, K. & Moore, W. (1945), "Some Principles of Stratification", **American Sociological Review**.
- Eden, F. M. (1797), **The State of the Poor or A History of the Labour Classes in England from the Conquest to the Present Period**, London, J. Davies, vol.4.
- Heckerman, D.; Breese, J. & Rommelse, K. (1994), "Troubleshooting Under Uncertainty", **Technical Report MSR-TR-94-07**, Microsoft Research. Redmond, Washington.
- Getoor, L.; Taskar, B. & Koller, D. (2001), "Selectivity Estimation Using Probabilistic Models", In: **Proceedings of the ACM SIGMOD International**

- Conference on Management of Data, (SIGMOD).**
- Jensen, F. V. (1996), *An Introduction to Bayesian Networks*, Springer-Verlag, New York.
 - Kakwani, N. (2003), "Issues in Setting Absolute Poverty Line", **Poverty and Social Development Papers**, Asian Development Bank.
 - Lewis, O. (1996), "The Culture of Poverty", In: G. Gmelch & W. Zenner, (eds.), **Urban Life**, Waveland Press.
 - Pearl, J. (1988), **Probabilistic Reasoning in Intelligent Systems**, Morgan Kaufmann, San Mateo. California.
 - Pearl, J., & Verma, T. S. (1991), "A Theory of Inferred Causation", In: J. A.; Allen, R. Fikes, & E. Sandewall. (eds.), **Principles of Knowledge Representation and Reasoning Proceedings of the Second International Conference**, Morgan Kaufmann, San Mateo. California.
 - Rasmussen, T. C. (1995), "Laboratory Characterization of Fluid Flow Parameters in a Porous Rock Containing a Discrete Fracture", **Geophysical Research Letters**, 22(11).
 - Ramoni, M. & Sebastiani, P. (2001), "Robust Learning with Missing Data", **Machine Learning**, 45(2).
 - Sen, A. K. (1981), **Poverty and Famines: An Essay on Entitlement and Deprivation**, Oxford University Press, Oxford.
 - Sen, A. K. (1983), "Poor, Relatively Speaking", *Oxford Economic Papers*, 35(2).
 - Spirtes, P.; Glymour, C. & Scheines, R. (1993), **Causation, Prediction, and Search**, Springer-Verlag, New York, 2nd ed., (2000). MIT Press, Cambridge, Massachusetts.