فصلنامهٔ علمی ـ پژوهشی
روش‌شناسی علوم انسانی
س۲۳، ش۹۱، تابستان ۱۳۹۶
صفحات ۶۷-۸۳

Methodology of
Social Sciences and
Humanities Journal
Vol.23, No.91, Summer 2017

کاربردپذیری استنباط آماری از دیدگاه بازنمایی مدل‌های علمی

مهدی عاشوری*

سیدمحمود طاهری**

چکیده

این مقاله با مفروض گرفتن نظریه‌ای ساختاری و نگاشتی در باب چیستی بازنمایی، فرضیه «توضیح کاربردپذیری براساس مفهوم بازنمایی» را در مورد «کاربردپذیری استنباط آماری در مدل‌های علمی» بررسی می‌کند. براساس دیدگاه نگاشتی، وجود رابطه‌ای متعین میان مدل و سیستم هدف، شرط کافی برای بازنمایی هدف توسط مدل است و این بازنمایی چرایی کاربردپذیری استنباط آماری، در مدل‌سازی را نیز توضیح می‌دهد. تحلیل کارکردی «برآورد» به‌عنوان یکی از مهم‌ترین الگوهای استنباط آماری نشان می‌دهد که ۱. استنباط و تحلیل آماری داده‌ها در ارزیابی مفروضات مندرج در مدل مفهومی کاربرد دارد، بدین معنا که با استنباط آماری می‌توان بازنمایی مدل مفهومی از سیستم هدف و موضوع مورد مطالعه را ارزیابی کرد، ولی ۲. نمی‌توان نشان داد که دستگاه معادلاتی که به‌عنوان مدل آماری «برآورد» شده «فی نفسه» روابطی را در موضوع مورد مطالعه بازنمایی می‌کند؛ زیرا دستگاه معادلات به‌طور کامل به وسیله برآورد «متعین» نمی‌شود

واژگان کلیدی: مدل علمی، کاربردپذیری آمار، بازنمایی علمی، تبیین، فلسفه آمار، برآورد پارامتر

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

* دکتری فلسفه علم، مؤسسه پژوهشی حکمت و فلسفه ایران؛ Email:ashoori@irip.ir
** استاد دانشکده فنی، دانشگاه تهران؛ Email: sm_taheri@ut.ac.ir
تاریخ دریافت: ۹۵/۱۱/۸ تاریخ تأیید: ۹۶/۲/۱۹

درآمد

امروزه استنباط آماری[1] نقش انکارناپذیری در فعالیت‌های علمی ما یافته است. یکی از مهم‌ترین عرصه‌های کاربرد استنباط آماری در علوم، به ویژه علوم انسانی، کاربرد آن در مدل‌سازی است. پژوهشگران، غالباً، فرضیه تحقیق را، که روابطی مفروض میان دو یا چند مفهوم است، به‌عنوان «مدل مفهومی»[2] در نظر می‌گیرند و مفاهیم مندرج در مدل را در قالب «متغیرهایی» تعریف می‌کنند. آنها در طراحی مدل مفهومی از میان روابط ممکن میان این مفاهیم در سیستم هدف برخی را مفروض می‌گیرند و از برخی از روابط ممکن چشم‌پوشی می‌کنند و در ادامهٔ تحقیق به نمونه‌گیری و گردآوری «داده»[3] می‌پردازند. مهم‌ترین گام در کاربرد استنباط آماری، تحلیل داده‌ها است.[4] در برخی از روش‌های آماری، مانند «تحلیل همبستگی»[5] و «تحلیل رگرسیون»،[6] «تحلیل مسیر»،[7] «تحلیل عاملی»[8] و «مدل‌سازی معادلات ساختاری»[9] معادله یا دستگاهی از معادلات به‌عنوان «مدل آماری» مرتبط با مدل مفهومی نیز معرفی می‌شود. برای توضیح مطلب به مثال زیر توجه کنید:

مثال ۱. در بررسی میزان اثر «سیاست ارزی انبساطی» در کنترل «نرخ ارز»، پژوهشگران لازم می‌بینند به متغیرهای دیگری مانند تورم، میزان سود مورد انتظار در بازار مسکن و مانند آن نیز بپردازند؛ زیرا این متغیرها ممکن است به صورت «مخدوش‌گر»[10] مطالعه را دچار سوگیری کنند، بنابراین، تأثیر آنها باید کنترل شود. ممکن است این متغیرها به امور دیگری نیز وابستگی داشته باشند، اما پژوهشگر با این فرض که اثرات آنها اندک و در عمل «قابل چشم‌پوشی» است، تمام این اثرات را به‌عنوان یک «متغیر اندازه‌گیری نشده» در مدل لحاظ می‌نماید. مرسوم است که روابط میان متغیرها را در قالب یک «نمودار»[11] یا «گراف» ترسیم کنند. همان‌گونه که در شکل ۱ ملاحظه می‌کنید، میان سه مفهوم «سیاست ارزی»، «تورم» و «ثبات نرخ ارز» روابط مختلفی را می‌توان در نظر

---

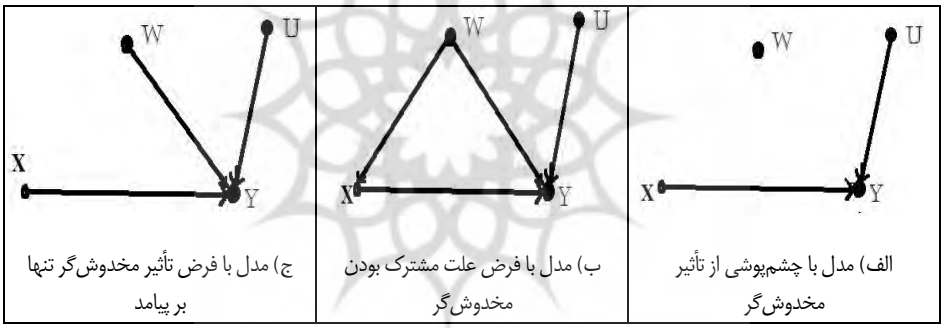
1. statistical inference
2. conceptual model
3. data

٤. در این تحقیق به روش‌های مبتنی بر مدل مفهومی پرداخته می‌شود. روش دیگر گردآوری داده‌ها بدون لحاظ نمودن مدل مفهومی است که در داده‌کاوی و مطالعات اکتشافی مورد توجه قرار می‌گیرد.

5. correlation analysis
6. regression analysis
7. path analysis
8. factor analysis
9. Structural Equations Modeling (SEM)
10. confounding
11. diagram

گرفت و یا از برخی از روابط چشم‌پوشی نمود که هر مجموعه روابط منجر به یک مدل خاص می‌شود.

اگر در مدل مفهومی پیش‌گفته، تنها رابطه مداخله (سیاست ارزی انبساطی) با پیامد (ثبات نرخ ارز) مورد مطالعه قرار گیرد و از اثر مخدوش‌گرها چشم‌پوشی شود (مدل ۱.الف)؛ آنگاه برای ساختن مدل کمّی و تحلیل داده‌های آزمایشی یا مشاهداتی چند روش وجود دارد که یکی از معمول‌ترین آنها استفاده از مدل‌های پارامتری است. برای مدل (ا.الف)، چنانچه نماد X را برای متغیر مداخله و نماد Y را برای متغیر پیامد در نظر بگیریم، آنگاه از معادله $E[Y|X]=\beta_0+\beta_1X$ برای تحلیل داده‌ها استفاده می‌شود که در آن β_0 و β_1 پارامترهایی هستند که باید «برآورد»[1] شوند. همچنین با استفاده از «آزمون‌های فرضیه»[2] می‌توان بررسی نمود که در نظر گرفتن معادله خطی برای رابطه آن بین X و Y، معنادار و پذیرفتنی است یا خیر؟ **در مثال فوق، می‌توان فواصل اطمینان برای میزان اثر «مداخله» بر «پیامد» را نیز تشکیل داد.**



الف) مدل با چشم‌پوشی از تأثیر مخدوش‌گر

ب) مدل با فرض علت مشترک بودن مخدوش‌گر

ج) مدل با فرض تأثیر مخدوش‌گر تنها بر پیامد

شکل ۱: سه مدل مفهومی مختلف میان چهار متغیر «X: مداخله ارزی»، «W: نرخ تورم»، «Y: نرخ ارز» و «U: اثرات اندازه‌گیری نشده»

مدل‌سازی آماری دو وجه ممیزه نسبت به مدل‌سازی‌های غیر آماری (از جمله مدل‌های تعینی ریاضی) دارد:

۱. مدل‌سازی آماری مبتنی بر این فرض است که در موضوعات مورد بررسی، مؤلفه «عدم حتمیت»[3] وجود دارد، بنابراین، از متغیرهای تصادفی در مدل‌های آماری استفاده می‌شود. در مثال فوق، نسبت دادن عدد به مفاهیم چون ترخ تورم، نرخ ارز و مانند آن همیشه همراه با تغییراتی است

1. stimation
2. hypothesis tests
3. uncertaity

که سبب می‌شود به صورت تصادفی نوساناتی داشته باشد؛

۲. در بیشتر تحلیل‌های آماری، به دلیل محدودیت دسترسی به کلّ جامعه یا پرهزینه بودن آن، ابتدا «نمونه»ای از جامعه اخذ می‌شود و سپس به روش‌های مختلف تلاش می‌شود، استنباط‌های آماری درباره جامعه انجام پذیرد. برای نمونه، در اندازه‌گیری نرخ ارز در بازار، بانک مرکزی مجبور است براساس یک نمونه از بازار نرخ ارز را برآورد نماید. براساس این مؤلفهٔ مدل‌های آماری است که مدل‌های آماری دارای ویژگی محتوا افزا و استنباط از جزء درباره کل باشند.

هنگامی که از اهمیت «مدل‌های علمی» بحث می‌شود، این دغدغه وجود دارد که آیا «مدل-ها» موضوع مورد مطالعه را «بازنمایی» می‌کنند، اگر بازنمایی مفهوم پایه‌ای در توضیح «کاربردپذیری» باشد، چنانچه در رویکردهای ساختارگرایانه مفروض است، توقع این است که ویژگی‌های یک مدل علمی براساس «ویژگی‌های بازنمایی» آن توضیح داده شود. مسئله اصلی این مقاله این است که آیا اصولاً می‌توان، «کاربردپذیری» استنباط آماری را در مدل‌سازی توضیح داد؟ این کاربردپذیری چه نسبتی با بازنمایی دارد؟ آیا ویژگی عدم حتمیت مفاهیم مورد بررسی در مدل‌های آماری و ویژگی استنباطی آن که براساس آن مدل مدعی بازنمایی خصایص کل جامعه است، برای بازنمایی مدل‌های آماری و کاربردپذیری آن مشکلی ایجاد نمی‌کند؟

۱. رویکرد ساختارگرایانه در توضیح بازنمایی و کاربردپذیری

آیا می‌توان «کاربردپذیری» مدل‌های آماری را به نقش آن در «بازنمایی» مدل علمی فرو کاست؟ در این بخش، نوشتگان فلسفی در توضیح بازنمایی و کاربردپذیری با رویکرد ساختاری به اختصار مرور می‌شود. سپس در بخش ۲، این رویکرد را در حوزه «کاربردپذیری مدل آماری در مدل-سازی» توضیح می‌دهیم.

۱-۱. چیستی بازنمایی

«بازنمایی» از مفاهیمی است که طی چند دههٔ اخیر در حوزه‌های مختلف فلسفی مورد توجه قرار گرفته است. برای نمونه، «بازنمایی زبانی» در فلسفه زبان، «بازنمایی تصویری» در فلسفه هنر و «بازنمایی علمی» در فلسفه علم مطرح شده است. انتظار فیلسوفان علمی که از مفهوم «بازنمایی علمی» استفاده می‌کنند این است که با این مفهوم می‌توانند کارکرد نظریه‌ها و مدل‌های علمی را توضیح دهند. در نظریه‌های بازنمایی تلاش می‌شود که مفهوم بازنمایی به مفاهیمی ساده‌تر و آشناتر

مربوط شود.[۱] بنابراین، یک نظریه در مورد بازنمایی، شرایط لازم و کافی را در مورد بازنمـایی علمـی با عبارتی مانند زیر بیان می‌کند: (یغمایی و شیخ‌رضایی، ۱۳۹۱)

مدل یا نظریه علمی M (به‌عنوان منبع)، هدف T را بازنمایی می‌کند اگـر و تنهـا اگـر (۱) ...، (۲) ...، (۳)... و...

به دیگر سخن، اگر M یک مدل علمی و T جنبه‌ای از واقعیت خارجی باشـد، رابطـه بازنمـایی علمی را می‌توان بـا $R_s(M,T)$ نشـان داد و نظریـه بازنمـایی علمـی جملـه زیـر را کامـل می‌کنـد: ($R_s(M,T)$ iff ...). این نظریه‌ها به ما پاسخ می‌دهند که منبع بازنمایی علمی، یعنی مدل علمـی، «به واسطه چه چیزی»[۲] هدف خود را بازنمایی می‌کنـد؟ بـرای مثـال از چـه جهـت اسـت کـه یـک اقتصاددان با مدلی آماری در باب وضعیت بـورس، رشـد یـا افـول برخـی از سـهام‌ها را پیش‌بینـی می‌کند؟

براساس دیدگاه ساختاری، مدل علمی یـک سـاختار نظریـه-مجموعـه‌ای[۳] اسـت کـه محمـول نظریه-مجموعه‌ای خاصی را برآورده[۴] می‌کند. میان اجزای مـدل و اجـزای واقعیـت مـورد مطالعـه، تناظری ساختاری برقرار است، یعنی مدل علمی M واقعیت T را بازنمـایی می‌کنـد، اگـر سـاختار «سیستم هدف» را «دربرداشته» باشد و «بنگـارد». ایـن «دربرداشـتن»[۵] و «نگاشـت»[۶] بـا تنـاظر یـا شباهتی ساختاری ماننـد «یـک‌ریختی»[۷] یـا «هم‌ریختـی»[۸] فـراهم می‌شـود، یعنـی رابطـه بازنمـایی تناظری میان مدل و هدف است که برای نشان دادن این تنـاظر از یـک نگاشـت اسـتفاده می‌شـود.[۹]

---

۱. با فرض اینکه رابطه بازنمایی مفهومی پایه (primitive) نباشد؛ این دیدگاه‌ها را دیـدگاه‌هـای قـائم بـه ذات (Substantive) نامیده‌اند (سوارز، ۲۰۱۰).

2. In virtue of
3. Set-Theoretic
4. satisfied
5. embracing
6. mapping
7. Isomorphism
8. homomorphism

۹. با توجه به تعریف فوق آشکار می‌شود که وجود یک نگاشت میـان مجموعـه وابسـته بـه منبـع (مـدل علمـی) و مجموعـه وابسته به هدف، برای وجود «بازنمایی علمی» کافی است. در این تعریف هیچ اشاره‌ای به «عامـل شـناختی» (Agent) و نقش وی نمی‌شود و بازنمایی، رابطه‌ای عینی میان دو ساختار است که مستقل از حضور یا شناخت عامل می‌باشد. در مقابل رویکرد فوق، برخی از فیلسوفان علم، بازنمایی را رابطه‌ای عینی و مستقل از عامل شناختی نمی‌دانند. این رویکرد شامل دیدگاه‌های مختلفی است که با عنوان دیدگاه «کارکردی» (Functional) می‌توان از همه آنهـا یـاد کـرد. براسـاس این دیدگاه‌ها در صورتی میان منبع و هدف، رابطه بازنمایی وجود دارد که «عامل» (سازنده مدل یا کاربر مدل) بتوانـد از مدل یا نظریه علمی در فهم و نتیجه‌گیری ویژگی‌های منبع و در کل برای تأمین اهداف شناختی خویش نسبت به جهـان بهره ببرد (چاکراورتی، ۲۰۰۹). براساس این رویکرد، رابطهٔ تناظر ساختاری میان مدل و هدف برای بازنمایی کافی نیست و حتی از نظر برخی شرط لازم برای بازنمایی نیز نمی‌باشد. بررسی بازنمایی مدل آماری براساس این دیـدگاه در مقالـه‌ای دیگر طرح شده است.

بنابراین، تعریف زیر برای بیان شرایط لازم و کافی بازنمایی در رویکرد نگاشتی ارائه می‌شود.

تعریف (۱-۱). مدل $M=<A,R>$ با مجموعه اجزای ساختاری $A=\{a_1, a_2, a_3,\dots\}$ و رابطه R میان این اجزا، واقعیت $T=<A',L>$ با مجموعه اجزا $A'=\{a_1',a_2',a_3',\dots\}$ و رابطه L میان این اجزا را بازنمایی می‌کند، اگر و تنها اگر نگاشت $f:A\rightarrow A'$ رابطه میان R و L را به صورت زیر برقرار سازد:

$$f(a_i)=a_i' \text{ and } f(a_j)=a_j'; L(f(a_i), f(a_j))\leftrightarrow R(a_i,a_j)$$

آنچه در تناظر ساختاری مقصود اصلی است، ایجاد تناظر میان هویت‌های رابطه‌ای[1] R و L است نه برقراری تناظر میان ویژگی‌های عینی مدل A و واقعیت خارجی A'. این نکته را در هنگام بحث از رابطه ساختاری در مدل آماری، با مثال (۲-۲) توضیح خواهیم داد.

۲-۱. چیستی کاربردپذیری

در نوشتگان فلسفی، به بحث از کاربردپذیری ریاضیات پرداخته شده است که می‌تواند الهام‌بخش بحث از کاربردپذیری آمار باشد. ایده اصلی رویکرد ساختاری در باب کاربردپذیری ریاضیات، این است که «ریاضیات منبعی غنی از «ساختار»ها است و هنگامی که یک نظریه ریاضیاتی در علوم کاربرد پیدا می‌کند، آشکار است که ریاضیات روابط ساختاری اصلیِ سیستم مورد مطالعه را پوشش می‌دهد». این دیدگاه که موسوم به «رویکرد نگاشتی»[2] است، کاربردپذیری ریاضیات را با شباهت ساختار میان مدل ریاضی و ساختار موضوع مورد مطالعه توضیح می‌دهد (پینکاک، ۲۰۰۷؛ بوئنو و کولیوان،[3] ۲۰۱۱).

عموماً رابطه تناظر ساختاری میان مدل و هدف رابطه یک‌ریختی دانسته می‌شود.[4] برای تعریف رابطه یک‌ریختی فرض می‌کنیم که M «دربرگیرنده» مجموعه M باشد که رابطه n-تایی R روی آن تعریف می‌شود، به صورت نمادین ساختار M را با $<M, R>$ نمایش می‌دهیم. به همین ترتیب، ساختار T را با $<T, L>$ نمایش می‌دهیم. با این مفروضات، رابطه یک‌ریختی بین مدل و هدف به صورت زیر تعریف می‌شود.

تعریف (۱-۲). دو ساختار با $M=<A, R>$ و $T=<A', L>$ یک‌ریخت هستند، اگر و تنها اگر

---

1. relational entities
2. mapping account
3. O. Bueno & M. Colyvan

٤. البته یکی از نقدهایی که به رویکرد نگاشتی در باب کاربردپذیری شده، این است که رابطه یک‌ریختی ممکن نیست میان مدل و موضوع مورد مطالعه برقرار باشد (بوئنو و کولیوان، ۲۰۱۱).

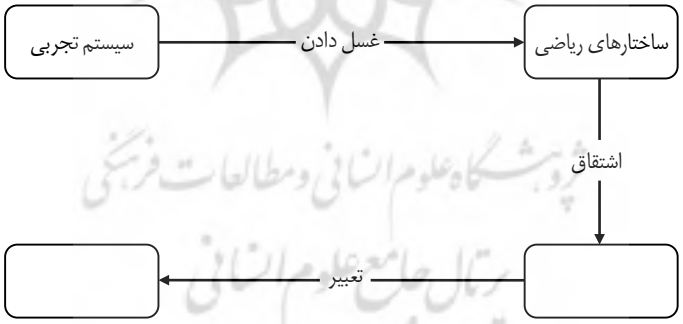
نگاشت پوشا و یک به یک f از M به T چنان وجود داشته باشد کـه بـه ازای هـر $(m_1, m_2, \ldots, m_n) \in M^N$ داشته باشیم:

رابطه (۱-۱) $R(m_1, m_2, \ldots, m_n) \leftrightarrow L(f(m_1), f(m_2), \ldots, f(m_n))$

که

$(t_1, t_2, \ldots, t_n) = (f(m_1), f(m_2), \ldots, f(m_n)) \in T^N$.

در تقریر فوق از رویکرد نگاشتی، وظیفـه ریاضـیات صـرفاً بـه نقشـه‌ای سـاختاری از «هـدف بازنمایی» و موضوع مورد مطالعه اکتفا شده است. برخی از منتقدان این رویکرد یادآور می‌شوند که عملیات‌های تبدیل ساختارها در «ریاضیات کاربردی» اهمیت بسیاری دارند. ساختارهای ریاضـی در ریاضیات به یکدیگر تبدیل می‌شوند و قدرت اصلی ریاضیات نیز به همـین «تبـدیلات» اسـت. براساس همین نکته الگوی پیشنهادی آنان در باب کاربرد ریاضیات در بازنمایی سیستم‌های هـدف شامل دو نگاشت و یک گام میانی برای «اشتقاق»[1] است، یعنی ریاضیات کاربردی شـامل سـه گـام می‌شود: ۱. نگاشتی از «سیستم مورد مطالعه یا تجربی»[2] به یک سـاختار مناسـب ریاضـی کـه آن را «غسل دادن»[3] می‌نامند؛ ۲. «اشتقاق» یا «استنتاج» ساختار ریاضی جدید از ساختار ریاضی اولـی با استفاده از روابط تبدیل ریاضی؛ ۳. نگاشتی از ساختار ریاضی جدید به سیستم مورد مطالعه کـه آن را نگاشت «تعبیر»[4] می‌نامند (بوئنو و کولیوان، ۲۰۱۱).



شکل ۲: مراحل کاربرد ریاضیات. نگاشت میان دامنه امور فیزیکی و امور ریاضی و استنتاج ساختارها در جهان یاضیات

1. derivation
2. empirical set-up
3. Immersion
4. Interpretation

نقش اصلی ریاضیات کاربردی، مربوط به گام دوم است. به وسیله این تبدیلات، ویژگی‌هایی که در ساختار اولیه پنهان بود، در ساختار تبدیل یافته آشکار می‌گردد. یعنی اجزایی در ساختار جدید، ظاهر می‌شود که باعث کشف ویژگی‌هایی بدیع در سیستم هدف می‌شود و همین امر سبب اهمیت «تبیین ریاضی» در علوم می‌گردد. به همین خاطر است که نقش ریاضیات بیش از نقشه‌برداری از جهان است و اگر تبدیلات ریاضی وجود نداشت، ریاضیات فایده چندانی برای علوم نداشت. اهمیت این تبدیلات و استنتاج در پیش‌بینی‌های «بدیعی» است که توسط ساختار جدید صورت می‌گیرد.

۲. کاربردپذیری استنباط آماری از دیدگاه تناظر ساختاری

همان‌گونه که در درآمد بیان شد، یکی از کاربردهای «استنباط آماری» تحلیل داده‌های مربوط به یک «مدل مفهومی» است که از آن برای مدل‌سازی و ارزیابی فرضیه‌ها استفاده می‌شود. اشاره شد که استنباط آماری مبتنی بر این فرض است که مفاهیم مطرح در موضوع مورد مطالعه، دارای مؤلفه «عدم حتمیت» هستند و آن را باید به صورت متغیرهای تصادفی لحاظ و از توابع توزیع احتمال و پارامترهای آن در مدل‌سازی آماری برای بیان روابط میان وضعیت‌های امور استفاده کرد. به دلیل عدم حتمیت در بسیاری از پدیده‌های طبیعی و مطالعات انسانی، کاربرد مباحث استنباط آماری، مانند «برآورد»، «آزمون فرضیه»،[1] «انتخاب مدل»،[2] «بازه‌های اطمینان»[3] و مانند آن در علوم مختلف، مانند مهندسی، پزشکی، کشاورزی، اقتصاد، جامعه‌شناسی، روان‌شناسی و... رایج است.

نگاه دانشمندان و پژوهشگرانی که از «آمار» در روش تحقیق استفاده می‌کنند، معطوف به جنبه تکنیکی «آمار» است و آن را «تکنولوژی استنباط اطلاعات از داده‌ها» می‌بینند؛ زیرا در این علم تکنیک‌هایی برای بررسی و ارزیابی ویژگی‌های یک جامعه آماری ارائه می‌شود. براساس این تصویر، «آمار» به مثابه جعبه‌ای از ابزارها و دستورالعمل‌های محاسبه مورد توجه قرار می‌گیرد که پژوهشگر با توجه به مسئله مورد نظر خویش ابزاری از درون آن را یافته و به کار می‌گیرد؛ اما روی دیگر آمار، علم و نظریه‌ای ریاضی است که مانند دیگر نظریه‌های ریاضیاتی صورت‌بندی و ارائه می‌شود، به‌گونه‌ای که ابزارهای مختلف با هم مرتبط هستند و «قضایای» آن نظریه مبتنی بر مجموعه‌ای از اصول اثبات می‌شوند و در حقیقت آمار «علم استنباط اطلاعات از داده‌ها» می‌باشد (هند، ۲۰۰۸، ص۲-۳). حال اگر «نظریه آمار» را نیز یک «نظریه اصل موضوعی» فرض کنیم، آیا

1. hypothesis testing
2. model selection
3. confidence intervals

می‌توان با رویکردی نگاشتی کاربردپذیری نظریه آمار را توضیح دهند؟ بررسی خود را محدود به استنباط پارامتری و مدل‌های ساختاری می‌کنیم که یکی از رایج‌ترین روش‌های آماری در علوم مهندسی و علوم انسانی است.

۱-۲. مدل‌های آماریِ ساختاری

در این مقاله، تناظر میان مدل و هدف در دسته‌ای از مدل‌های آماری بررسی می‌شود که رابطه میان متغیرها در منبع بازنمایی، به وسیله یک معادله یا دستگاهی از معادلات نشان داده می‌شود. مدل‌هایی چون، تحلیل همبستگی، رگرسیون، تحلیل عاملی و معادلات ساختاری که در بسیاری از موارد برای تحقیق آماری در باب فرضیات علّی استفاده می‌شود. آماردانان اذعان می‌کنند که همبستگی‌هایی که در چنین مدل‌هایی یافت می‌شود دلالت بر رابطه علّی نمی‌کند و همبستگی ممکن است غیر علّی باشد، اما پژوهشگران رشته‌های مختلف، مانند آمار زیستی، اقتصادسنجی، روان‌شناسی و...، تلاش کرده‌اند تکنیک‌هایی را توسعه دهند که با احتمال قابل قبولی نه تنها همبستگی‌های غیر علّی را کنار بگذارد، بلکه روش‌های رایج مورد اشاره در مقدمه مانند «تحلیل رگرسیون»، «تحلیل مسیر»، «تحلیل عاملی» و «مدل‌یابی معادله ساختاری» و مانند آن را نیز برای «استنباط علّی»[1] اصلاح و تکمیل کنند. در این روش‌ها مدل آماری در قالب یک دستگاه معادلات ارائه می‌گردد که براساس آن «میانگین تغییرات یک یا چند متغیر پاسخ براساس تغییرات یک یا چند متغیر پیش‌بین» بیان می‌شود. گاه به همه این‌گونه مدل‌ها، «مدل ساختاری» اطلاق می‌شود (پرل، ۲۰۰۹، ص۱۳۴).

هدف اولیه پژوهشگرانی که از این مدل‌ها استفاده می‌کنند، می‌تواند توصیف، تبیین، تأیید، پیش‌بینی و یا کنترل باشد (ویک‌فیلد،[2] ۲۰۱۳، ص۴). این روش‌ها تکنیک‌هایی برای تحلیل داده‌ها هستند که برای کشف و وارسی همبستگی یا ارتباط‌های خطی میان متغیرها استفاده می‌شوند و انتظار پژوهشگرانی که از آنها استفاده می‌کنند این است که تأثیرات خالص متغیرهای علّی بر یک پیامد را برآورد کنند.

در میان این مدل‌ها ساده‌ترین حالت «رگرسیون ساده خطی» است که در آن یک متغیر پیش‌بین و یک متغیر پاسخ وجود دارد. این تحلیل در مورد یک «مجموعه داده» صورت می‌گیرد که این داده‌ها را می‌توان به کمک نمودار پراکنش حاصل از مختصات این داده‌ها مشخص کرد و تحلیل رگرسیون،

1. causal inference
2. J. Wakefield

خط و معادله‌ای را در اختیار ما قرار می‌دهد که به مثابه رابطه‌ای برای پیش‌بینی مقادیر یک متغیر براساس مقادیر دیگر است. یعنی با تحلیل رگرسیونی ساده، متغیر پیشین y_i، ترکیبی خطی از متغیر مستقل x_i و ضرایب β_0 و β_1 خطی دانسته می‌شود که با معادله زیر نشان داده می‌شود:

$$Y=\beta_0+\beta_1 X+\varepsilon \qquad (۲-۱)$$

در این معادله، ضرایب β_0 و β_1 پارامترهایی هستند که براساس داده‌های مشاهدتی یا آزمایشی برآورد می‌شوند. فرض بر این است که در حوزه داده‌های مورد تحلیل، معادله خطی (۲-۱) تقریبی قابل قبول از رابطه میان X و Y است و ε یک خطای تصادفی است که ناشی از خطای سیستمی و متغیرهای لحاظ نشده در تحلیل است. به دیگر سخن، در صورت برقراری برخی شروط و مفروضات، این معادله حالتی «تقریبی» و «ایدئال‌سازی شده» از داده‌ها است. اگر n «داده» داشته باشیم، بر طبق معادله (۲-۱) هر داده منفرد را می‌توان به صورت

$$y_i=\beta_0+\beta_1 x_i+\varepsilon_i \quad i=1, 2, \ldots, n \qquad (۲-۲)$$

نوشت که y_i مقدار iام متغیر پاسخ Y، x_i مقدار iام متغیر پیش‌بین X و ε_i خطای در تقریب y_i را نشان می‌دهد. آماردان‌ها در تحلیل این داده‌ها چندین پرسش را دنبال می‌کنند؛ اول اینکه، آیا بنابر داده‌ها میان این دو متغیر همبستگی آماری وجود دارد و چه مقدار از متغیر پیش‌بین می‌تواند توسط متغیر مستقل تبیین شود؛ دوم اینکه، آیا مدل رگرسیون می‌تواند به‌طوری معنادار و مناسب، تغییرات متغیر وابسته را پیش‌بینی کند. آزمون‌های فرضیه مختلف به این پرسش پاسخ می‌دهند؛ سومین وظیفه استنباط آماری، برآورد ضرایب مدل ساختاری براساس داده‌های موجود است. هنگامی که بیشتر از یک متغیر مستقل در مدل داشته باشیم، مقدار β_i به ما کمک می‌کند که سهم نسبی هر متغیر را در پیش‌بینی متغیر وابسته مقایسه کنیم و به عبارتی بتوانیم تعیین کنیم که کدام متغیرها بیشترین تأثیر را بر متغیر وابسته دارند؛ چهارمین مسئله، ارزیابی برآوردهای صورت گرفته با آزمون‌های فرضیه متکفل ارزیابی برآوردها است. براساس فرایندی که خلاصه‌وار توضیح داده شد، مدلی علمی با استفاده از استنباط آماری فراهم می‌شود.

۲-۲. تناظر ساختاری میان مدل آماری و سیستم هدف

همان‌گونه که در توضیح تعریف (۱-۱) گفتیم، آنچه در تناظر ساختاری مقصود دیدگاه نگاشتی است، ایجاد تناظر میان هویت‌های رابطه‌ای ساختار منبع با روابط امور در سیستم هدف و نه صرفاً برقراری تناظر میان متغیرهای مدل با واقعیت‌های خارجی است. این نکته با یک مثال مبتنی بر یک مدل آماری توضیح داده می‌شود.

مثال (۲-۲). فرض کنید در مطالعه «رابطه نقش حجم نقدینگی و تورم»، معادله رگرسیونی (به صورت $R(Y,X): Y=\beta_0+\beta_1X+\varepsilon$) به دست آمده و بنابر آزمون فرضیه، خطی بودن معادله پذیرفتنی باشد. T نیز «تأثیر نقدینگی بر تورم» در واقعیت است. مدل آماری $M=<A,R>$ که $A=\{Y,X\}$ و R معادله رگرسیونی است که رابطه میان متغیرها را با پارامترهای β_0 و β_1 نشان می‌دهد. T نیز اجزایی مانند «حجم نقدینگی، m_2» و «نرخ تورم، m_1» دارد که انتظار داریم رابطه واقعی L میان آنها برقرار باشد. مدعای رویکرد نگاشتی این است که رابطه‌ای یک‌ریخت $R(Y,X)\leftrightarrow L(f(m_1), f(m_2))$ وجود دارد که شرط لازم و کافی برای بازنمایی T توسط M است، یعنی پارامترهای β_0 و β_1 که رابطه R را متعین می‌کند، در واقع، یعنی در L مابه‌ازاء دارند.

در مثال پیش، باید معادله رگرسیونی با رابطه‌ای حقیقی در موضوع مورد مطالعه متناظر باشد. در این حالت مدل آماری در قالب یک معادله (در برخی از مدل‌ها به صورت دستگاهی از معادلات) نمایش داده می‌شود و ضرایب ساختاری، این معادله را متعین می‌کنند. پرسشی که قابل طرح است، این است که این ضرایب که با استفاده از برآورد پارامتر محاسبه شده‌اند، چه چیزی را بازنمایی می‌کنند؟

در پاسخ به این پرسش باید توجه داشت که بنابر اصول آمار کلاسیک، هیچ اطمینان قطعی نسبت به درستی برآورد پارامتر وجود ندارد و خطای برآورد با استفاده ملاکی ممزوج از «صحت و دقت» و خطای قضاوت براساس میزان اطمینان از نزدیک بودن پارامتر برآورد شده به مقدار واقعی، محاسبه می‌گردد. این مسئله در پارادایم «بیز»[1] در استنباط آماری پیچیده‌تر می‌شود؛ زیرا براساس اصول این آمار، پارامترها مقداری واقعی ندارند و خود متغیری تصادفی هستند. در بخش بعد با تحلیل کارکرد «برآورد» در استنباط آماری نشان خواهیم داد که رویکرد نگاشتی نمی‌تواند توضیحی قانع‌کننده در مورد بازنمایی مدل‌های آماری ارائه دهد.

۳. برآورد پارامترها و تعین ناقص[2] روابط آماری

در بخش گذشته بازنمایی مدل آماری به مثابه تناظر ساختاری میان دستگاه معادلات برآورد شده و روابط ادعایی در موضوع مورد مطالعه صورت‌بندی شد. حال اگر کاربردپذیری نظریه آمار در ساختن مدل‌های آماری به تناظر ساختاری میان روابط ادعایی در مدل با روابط واقعی فروکاسته شود، باید به این پرسش پاسخ داد که ضرایب مدل آماری که «برآورد» می‌شوند، چه تعبیری دارند؟

1. Bayesian Paradigm
2. underdetermination

این ضرایب، هویت رابطه‌ای دارند و رابطه ادعایی در مدل آماری را متعین می‌سازند. پیش از بحث در باب تعبیر ضرایب برآورد شده می‌توان این پرسش را طرح کرد که اگر در بازنمایی، هدف برقراری تناظر میان مدل و موضوع مورد مطالعه باشد، این تناظر در چه سطح قابل طرح است:

۱. تناظر متغیرهای مدل با مفاهیم موضوع مورد مطالعه. این تناظر در مرحله طراحی مدل مفهومی صورت می‌گیرد و ممکن است که دچار خطاهایی شود، ولی این خطاها مربوط به تحلیل داده‌ها و استنباط آماری نیست؛

۲. تناظر روابط ادعایی مدل با روابط خارجی. روابط ادعایی در مدل، همان روابط ادعایی در مدل کیفی است و این تناظر پیش از تحلیل آماری مفروض است. برخی از روابط خارجی ممکن است توسط پژوهشگر مورد چشم‌پوشی قرار گیرد. پژوهشگران نیز اصولاً ادعا ندارند که تمام روابط را می‌خواهند، مدل کنند، زیرا مدل کردن تمام روابط در بسیاری از پژوهش‌ها، بیهوده و فاقد کارایی است. به عبارتی از این نظر صحیح این است که بازنمایی مدل‌های علمی در این سطح را «بازنمایی جزئی» بدانیم، اما این سطح نیز مانند سطح پیشین به نحوه کاربرد استنباط آماری و تحلیل داده‌ها مربوط نیست؛

۳. آنچه در مدل‌سازی آماری و کاربرد استنباط آماری به مدل اضافه می‌شود، برآورد مقدار برای پارامترهای مدل ساختاری است. البته در استنباط آماری برای «آزمون فرضیه» نیز محاسباتی صورت می‌گیرد که جهت ارزیابی مدل است، اما انتظار پژوهشگران این است که مدل آماری با پارامترهای برآورد شده، بیشتر از مدل کیفی موضوع مورد مطالعه را بازنمایی می‌کند، اما آیا این انتظار قابل توضیح است؟ آیا روابط حاصل از برآورد پارامترها و مدل، روابطی حقیقی را بازنمایی می‌کنند؟

۱-۳. اجزای تعین‌بخش مدل آماری

در بسیاری از مباحث آمار پارامتری، مدل آماری در دو مرحله متعین می‌شود:

۱. تعیین فرم تابعی[1] مدل، تأیید فرم انتخابی با انجام یک آزمون نیکویی برازش صورت می‌گیرد؛

۲. برآورد پارامتر که با استفاده از روش‌های مختلف برآورد و با استفاده از برآوردگر صورت می‌گیرد.

همان‌گونه که بیان شد، نقش استنباط آماری در مدل‌سازی، برآورد پارامترهای معادله متناظر با

1. functional form

مدل مفهومی و ارزیابی آن براساس داده‌ها و شواهد است. برای نمونه، در یک معادله رگرسیونی، کمیت‌های β_0 و β_1 و ε وجود دارد[۱] که آماردانان برای هر یک از این کمیت‌ها تعبیری را معرفی می‌کنند:

تعبیر ε مقدار «انحراف از میانگین» است و مؤلفه تصادفی مدل است و سهمی از y را نشان می‌دهد که x قادر به تبیین آن نیست.

تعبیر β_0 که عرض از مبدأ خط رگرسیون است، مقدار y را در نبود متغیر x نشان می‌دهد.

تعبیر β_1 که شیب خط رگرسیون است، میانگین تغییرات y را براساس تغییر یک واحد x نشان می‌دهد.

تعابیر فوق، تعبیر محض یک رابطه خطی میان چند متغیر است. هنگامی که از «بازنمایی مدل‌های آماری» سخن می‌گوییم، نظر به رابطه «مدل رگرسیونی» با «سیستم هدف» است. مدل رگرسیونی یک معادله با ضرایب خاص است که رابطه میان متغیرها را با مفاهیم متناظر با آنها در خارج را نشان می‌دهد.

هنگامی که از بازنمایی و تناظر میان مدل و هدف سخن می‌گوییم، انتظار اولیه این است که ضرایب معادله مورد نظر در مدل آماری در سیستم هدف دارای مابه‌ازاء باشند و روابطی واقعی میان مفاهیم را بازنمایی کنند. اما آیا این انتظار از استنباط آماری اصولاً صحیح است؟ پاسخ مثبت به پرسش فوق از دو جهت مورد خدشه قرار می‌گیرد: الف) از باب چیستی پارامترهای مدل از دیدگاه آمار بیز و ب) تعین ناقص در برآورد پارامتر.

۲-۳. چیستی پارامترهای مدل

در باب اصول و مبانی استنباط آماری پارادایم‌های مختلفی وجود دارد (باندی‌اوپادیهای و فورستر، ۲۰۱۱[۲]). دو پارادایم اصلی آمار بسامدی و آمار بیز رایج‌ترین الگوهای کلان استنباط آماری را تشکیل می‌دهند که اختلاف اصلی آنها در تعبیر و چیستی احتمالات ریشه دارد. این ایده که احتمالات بیان‌کنندهٔ وضعیتِ شانسیِ امور و مبتنی بر فرایندهای تصادفی است، در بسیاری از علوم رایج است. در این دیدگاه، امور احتمالی به «بسامدهایی»[۳] (فراوانی‌هایی) در دنباله‌ای از رویدادها مرتبط است یا آنکه «تمایل‌ها»[۴] یا «گرایش‌هایی»[۱] در سیستم‌ها هستند که این رویدادها

---

۱. کمیت‌های β_0 و β_1 پارامتر و ضریب هستند و ε خود یک متغیر تصادفی که برآورد می‌شوند.

2. P. Bandyopadhyay, & M. Forster

3frequencies

4tendencies

را متحقَّق[2]ساخته است، به چنین تعبیرهایی از احتمال تعبیرهای عینی یـا فیزیکـی گفتـه مـی‌شـود. ویژگی کلیدی رویکرد بسامدی، مشاهده‌پذیری و تکرارپذیری رویدادهای مورد مطالعـه اسـت. در مقابل این رویکرد، آمار بیز مبتنی بر تعبیر ذهنی از احتمال است، یعنی احتمالات نشان می‌دهد که میزان باور و اعتقاد یا تغییرات آنها را چگونه باید باشد.

بنابر آمار بسامدی آنچه در یک مدل آماری «تصادفی» است و از یک تابع توزیع پیـروی مـی-کند، «متغیر»های مدل است و پارامترهای مدل، مقداری ثابت اما مجهول دارند. بـر ایـن اسـاس، بحث از بازنمایی این مجهولات معنایی مُحصّل دارد. اما ایدهٔ اصلی بیزگرایـی در اسـتنباط آمـاری این است که پارامترهای یک مدل آماری، خـود متغیـری تصـادفی هسـتند و دارای توزیـع احتمـال موسوم به «توزیع پیشین»[3]می‌باشند و آماردان براساس اطلاعـات پیش‌زمینـه‌ای خـود، ایـن توزیـع پیشین را مشخص می‌کند. بدین ترتیب، پارامترهای مدل اصولاً قـرار نیسـت، چیـزی را در سیسـتم هدف بازنمایی کنند. به عبارت دیگر، مدل آماری استنباط شده از آمار بیز اصولاً به دنبال بازنمـاییِ «واقعیت» نیست که از بازنمایی درست یا بازنمایی نادرست سخن بگوییم.

۳-۳. تعیّن ناقص در برآورد پارامتر

مسئله دوم این است که آیا روش‌های بسامدگرایانه برآورد پارامتر، اصولاً، ممکـن اسـت کـه مقـدار واقعی ضریب را محاسبه کنند یا خیر؟ برای برآورد پارامترها روش‌های متعددی پیشنهاد شده اسـت که مهم‌ترین آنهـا «روش کمتـرین مربعـات»،[4]«روش بیشـینه درسـت‌نمایی»، «روش گشـتاوری»، «روش بیز» و مانند آن است (کسلّا و برگر،[5]۲۰۰۳، ص۳۱۲-۳۳۰). به دلیل بحث پیشـین، دیگـر از روش بیز بحث نمی‌کنیم؛ زیرا مفروض اصلی این روش این است که به دنبال کشف یـک مقـدار مجهول نیست، اما روش‌های دیگر فـرض را بـر بـرآورد و تخمـین یـک مقـدار ثابـت امـا مجهـول گذاشته‌اند.

«برآورد»[6]پارامتر از طریق «برآوردگر»[7]و با جایگذاری یک نمونـه تصـادقی nتـایی از متغیـر X صورت می‌گیرد. این نمونه تصادفی $\{x_1, x_2, \ldots, x_n\}$ باید با متغیر X هم‌توزیع باشـند کـه ایـن

1propensities
2realized
3prior distribution
4least squares
5Cassella & Berger
6estimation
7estimator

توزیع تابعی از X و پارامتر θ است. برآوردگر یا برآوردیاب، یک «تابع» مانند $U=f(X_1, X_2, \ldots, X_n)$ از نمونه تصادفی است که در آن پارامتر مجهول θ موجود نیست. هرگاه این تابع را با مقادیر یک نمونه جایگذاری کنیم، انتظار داریم، مقدار آن تابع به‌عنوان «برآورد نقطه‌ای» مقداری «تقریبی» برای پارامتر مجهول به ما بدهد (بهبودیان، ۱۳۸۲، ص۴۷-۴۸).

هر روش آماری برآوردگری را پیشنهاد می‌کند که ممکن است از نظر ویژگی‌هایی چون میزان «اریبی»،[1] «واریانس»،[2] «کارایی نسبی»[3] و... با برآوردگرهای پیشنهادی دیگر متفاوت باشد. یکی از وظایف آمار ریاضی معرفی ملاک‌هایی برای انتخاب برآوردگر از میان برآوردگرها است. برای برخی تحقیقات برآورد «نااریب» اهمیت دارد و برآوردگری با کمترین واریانس از دسته برآوردگرهای نااریب انتخاب می‌شود، اما در برخی از مسائل، ملاک‌های دیگری، مانند ملاک متوسط مجذور خطا[4] (MSE)، لحاظ می‌شود، که طبق آن ممکن است برآوردگری با اریبی اندک نیز مرجوح گردد (همان، ص۳۴۲). نکته مورد نظر این است که در برآورد پارامترها و مدل‌سازی آماری دقت و بازنمایی دقیقِ هدف، همیشه ارجحیت ندارد.

هنگامی که فرض می‌کنیم، پارامتر β مقداری ثابت اما مجهول است، یعنی مقداری مانند u وجود دارد و اگر با برآوردگر B و نمونه X، مقدار b را برآورد می‌کنیم، یعنی خطای این برآورد، |u-b| خطای تصادفی برآوردگر B است که خود یک متغیر تصادفی است. این خطا را نمی‌توان اندازه گرفت، ولی به تأسی از گوس[5] «امید ریاضی»[6] توان دوم این خطا را به‌عنوان ملاکی برای ارزیابی برآوردگر لحاظ می‌کنند که به ملاک (MSE) موسوم است و مهم‌ترین ویژگی آن امکان محاسبات ریاضی و فایده عملی آن است. ملاک مهم دیگر، نااریبی برآوردگر است. برآوردگر B را نااریب گویند، اگر برای هر b عضو فضای پارامتر و تمام نمونه‌های nتایی داشته باشیم که E(B)=b.

دو نکته فوق نشان می‌دهد که برآورد پارامتر، فی نفسه و جدای از ملاکی که آماردان به‌عنوان عامل شناسا لحاظ می‌کند، امری متعین نیست که حال بتوان از رابطه‌ای عینی و مستقل میان متغیرهای در معادله آماری سخن گفت. نتیجه آنکه معادله آماری R که قرار است یک طرف تناظر ساختاری f باشد، به صورت کامل تعین ندارد و دچار تعین ناقص است و برای تعین نیازمند دخالت عامل شناسا است. در نتیجه نگاشت و تناظر ساختاری میان مدل و هدف نیز بدون دخالت

1. biasedness
2. variance
3. relative efficiency
4. mean square error
5. Gauss
6. expected value

عامل شناسا تعین کامل نمی‌یابد. این تعین ناقص در مقداردهی به پارامترها و اختلاف میان برآوردگرها، ریشه در دو مؤلفه عدم حتمیت و استنباط کل از جزء دارد.

جمع‌بندی

بنابر رویکرد «نگاشتی» کاربردپذیری مدل‌های علمی، به یک رابطهٔ عینی میان مدل با موضوع مورد مطالعه تقلیل پیدا می‌کند و فایده مدل علمی مانند فایده یک نقشه دانسته می‌شود. اما در باب مدل آماری و کاربرد استنباط آماری در مدل‌سازی علمی، چنین تناظری میان مدل و سیستم هدف در موضوع مورد مطالعه به صورت فی نفسه و بدون دخالت عامل «تعین» پیدا نمی‌کند.

براساس مباحث مطرح در باب استنباط آماری، با یک بازنمایی «غیر قطعی» و با «تعین ناقص» در مورد ضرایب مدل مواجه هستیم، پس نمی‌توان «کاربردپذیری» مدل آماری را به تناظر معادله که توسط ضرایب متعین می‌شود با رابطه واقعی تقلیل داد. از سوی دیگر، آمار ریاضی و استنباط آماری در مدل‌سازی پدیده‌ها «کاربردپذیر» است و این شهود وجود دارد که غنای محتوایی مدل‌های آماری، بیش از مدل‌های کیفی است. آیا این کاربردپذیری با عدم بازنمایی قابل جمع است؟ پاسخ به این پرسش منوط به دیدگاه در باب بازنمایی ضرایب مدل است. در باب این ضرایب چند موضع می‌توان داشت:

الف) این ضرایب اصولاً قرار نیست چیزی را در موضوع مورد مطالعه بازنمایی کنند، یعنی این شهود که ضرایب تعبیری در موضوع مورد مطالعه دارند، محترم شمرده نمی‌شود و تحلیل آماری تنها برای ارزیابی بازنمایی مدل مفهومی است؛

ب) مدعی شویم که این پارامترها، بنابر تعبیری که برای آنها گفته شد، مابه‌ازاء دارند، اما به علت عدم تعین در برآورد، این عامل است که از این روابط موجود در معادله برای «بازنمایی» استفاده می‌کند و بازنمایی و در نتیجه کاربردپذیری قابل تقلیل به تناظر ساختاری نیست؛

ج) پارامترهای برآورد شده، مابه‌ازاء خود را بازنمایی می‌کنند، ولی ما نسبت به این بازنمایی، قطع و اطمینان نداریم و این عدم قطعیت، ریشه در خصیصه احتمالی موضوع مورد مطالعه دارد.

به دشواری‌های رویکرد کارکردی و استنباطی در باب بازنمایی مدل‌های آماری، در مقاله‌ای دیگر پرداخته شده است (عاشوری و طاهری، ۱۳۹۵). براساس رویکرد استنباطی، بازنمایی مدل آماری تنها به تناظر ساختاری تقلیل داده نشود و براساس رویکرد «کارکردی» مدعی شویم که عامل شناختی (مدل‌ساز یا کاربر مدل) چنان قصد، انتخاب یا تعبیر کند که با توجه به ویژگی‌های مدل (در اینجا ضرایب ساختاری) ویژگی‌هایی را در مورد سیستم هدف در موضوع مورد مطالعه استنباط (یا فهم) کند و در این امر موفق باشد.

منابع

۱. بهبودیان، جواد (۱۳۸۲)، آمار ریاضی، تهران: امیرکبیر.

۲. عاشوری، مهدی؛ طاهری، سیدمحمود (۱۳۹۵)، «بازنمایی مدل‌های آماری: دو رویکرد نگاشتی و استنباطی»، دوفصلنامه فلسفه علم، س۶، ش۲، ص۶۵-۸۴.

۳. یغمایی، ابوتراب؛ شیخ‌رضایی، حسین (۱۳۹۱)، «بازنمایی علمی»، دوفصلنامه فلسفه علم، س۲، ش۱، ص۱۱۵-۱۳۴.

4. Bandyopadhyay, P., and Forster, M. (eds.), (2011), *Philosophy of Statistics*. Handbook for the Philosophy of Science series, Elsevier.

5. Bueno, O. and Colyvan, M. (2011), "An inferential conception of the application of mathematics", Nouˆs, 45: 2, pp. 345-374.

6. Cassella, G. and Berger, R. (2002), *Statistical inference*, 2nd edition, Duxbury.

7. Chakravartty, A. (2010), "Informational versus functional theories of scientific representation", Synthese, 172: 197-213

8. Hand, D. J. (2008), *Statistics: A Very short Introduction*, Oxford University Press.

9. Pearl, J. (2009), 2nd Ed, Causality: Models, reasoning, and inference. *Cambridge: Cambridge University Press.*

10. Pincock, C., (2007), "A Role for Mathematics in the Physical Sciences", **Nous** 41, pp. 253-275.

11. Suarez, M. (2010), "Scientific representation", Philosophy Compass, 5 (1), pp. 91-101.

12. Wakefield, Jon , (2013), *Bayesian and Frequentist Regression Methods*, Springer.