

تحلیل چندسطحی؛ راهکاری برای خطاهای حاصل از تجمیع داده‌ها: استفاده از داده‌های سطح دانش‌آموز و معلم تیمز ۲۰۱۱

زهرا نقش*

چکیده

مشکل بوم‌شناسی در تحلیل داده‌های تجمیع، مشکل آماری جدی در اغلب پژوهش‌هاست. در این مقاله ضمن توضیح مشکلات مربوط به تحلیل‌های تجمیع و تجمیع‌زدایی، مدل‌یابی چندسطحی به‌ویژه مدل‌یابی خطی چندسطحی (HLM) به‌عنوان روشی جایگزین معرفی شده است. پژوهش حاضر از نوع همبستگی است و رابطه میان متغیرها در ۶۰۲۹ دانش‌آموز پایه هشتم شرکت‌کننده در آزمون تیمز (۲۸۱۶ دختر و ۳۲۱۳ پسر) که پرسشنامه استاندارد آن را تکمیل کرده بودند، بررسی شد. نتایج بررسی رابطه مقدار تکلیف و مدت‌زمان انجام آن بر عملکرد ریاضیات دانش‌آموزان نشان داد که بین متغیر مقدار تکلیف و مدت‌زمان انجام آن با عملکرد ریاضیات در سطح دانش‌آموز رابطه معنی‌داری وجود دارد ولی این رابطه در سطح معلم معنی‌دار نبود. نتایج همچنین سودمندی تحلیل چندسطحی در داده‌های آشیانه‌ای را نشان داد.

واژگان کلیدی: مدل‌یابی چندسطحی، تجمیع، خطای بوم‌شناسی، مقدار تکلیف، مدت‌زمان انجام تکلیف، عملکرد ریاضیات

مقدمه

ساختار داده‌ها در علوم اجتماعی و علوم تربیتی اغلب سلسله‌مراتبی^۱ است. بدین معنا که متغیرهایی در سطح فردی وجود دارد و خود افراد نیز در واحدهای بزرگ‌تری گروه‌بندی شده‌اند. مثلاً در آموزش و پرورش، دانش‌آموزان در درون کلاس‌ها گروه‌بندی می‌شوند برخی متغیرها ویژگی‌های دانش‌آموزان را توصیف و برخی توصیف‌کننده ویژگی‌های کلاس‌ها هستند. این متغیرهای توصیف‌کننده کلاس‌ها ممکن است تجمیع^۲ شده متغیرهای سطح دانش‌آموز باشد و یا می‌توانند متغیرهایی باشند که از سطح کلاس به دست آمده‌اند. تجمیع داده‌ها فرایندی است که در آن داده‌های سطح پایین جمع شده و متغیر سطح بالاتر را می‌سازند. مثلاً وقتی محقق وضعیت اجتماعی-اقتصادی دانش‌آموزان یک کلاس را جمع می‌کند و متوسط وضعیت اجتماعی اقتصادی کلاس را به دست می‌آورد، متغیر وضعیت اجتماعی-اقتصادی را تجمیع کرده است (بریک و رادنبوش^۳، ۱۹۹۲). هدف اصلی این پژوهش، معرفی شیوه آماری درست در برخورد با متغیرهای سطوح مختلف است. دو رویکرد رایج در مواجهه با این داده‌ها وجود دارد که هر دو دارای مشکلات زیادی هستند. اول، رویکرد "تجمیع‌زدایی"^۴ است؛ در این رویکرد تمام متغیرهای سطح بالاتر در سطح فردی در نظر گرفته می‌شوند. ویژگی‌های معلم، کلاس و مدرسه در سطح فردی در نظر گرفته شده و تحلیل در سطح فردی صورت می‌گیرد. مشکل این رویکرد در این است که دانش‌آموزان در یک کلاس تأثیرات یکسانی از هر یک از متغیرهای سطح کلاس می‌گیرند و این رویکرد این مسئله را در نظر نمی‌گیرد، لذا فرضیه استقلال مشاهدات را که اساس تکنیک‌های آماری کلاسیک است در نظر نمی‌گیرد. رویکرد دوم "تجمیع" است؛ در این رویکرد متغیرهای سطح فردی در سطوح بالاتر تجمیع می‌شوند و انجام تحلیل در سطوح بالاتر انجام می‌شود. بنابراین ویژگی‌های دانش‌آموز در کلاس احتمالاً با وزن‌دهی^۵ نسبت به اندازه هر کلاس تجمیع و یک تحلیل سطح کلاسی انجام می‌گیرد (بریک و رادنبوش، ۱۹۹۲). در ادامه، رویکرد دوم

1. Hierarchical

2. Aggregation

3. Bryk & Raudenbush

4. Disaggregate

5. Weighting

یعنی تجمیع کردن متغیرها و خطاهای آن بیشتر تشریح می‌شود و تکنیک چندسطحی نیز به‌عنوان گامی بزرگ در جهت رفع مشکلات دو رویکرد قبلی معرفی خواهد شد. به‌طور کلی دو روش برای ارزیابی اثرات سطوح وجود دارد: تحلیل تجمیع‌زدایی^۱ و تحلیل تجمیعی^۲ که هر دو تحلیل، اثر متقاطع^۳ سطوح متغیرها را در سطح فردی بررسی می‌کنند. در تحلیل تجمیع‌زدایی داده‌های سطح بالاتر، مدرسه، به واحدهای سطح پایین‌تر توزیع می‌شود و لذا واحدهای سطح پایین‌تر، دانش‌آموز، برای تحلیل استفاده می‌شوند. تحلیل تجمیع‌زدایی به دلیل نادیده گرفتن مفروضه‌های تحلیل‌های پارامتریک به‌ویژه توزیع نرمال داده‌ها، همگنی واریانس و استقلال مشاهدات نامناسب است. در نظر نگرفتن ساختار آشیانه‌ای داده‌ها باعث کاهش گمراه‌کننده خطای استاندارد می‌شود و در نتیجه فاصله اطمینان به‌غلط کوچک می‌شود. مهم‌تر اینکه، این روش خطای نوع اول را در ارزیابی اثرات متغیرهای مستقل افزایش می‌دهد و این بدین معنی است که فرض صفر بیشتر از حد مورد انتظار رد می‌شود (پاترسون^۴، ۱۹۹۱).

در تحلیل تجمیعی، داده‌های سطح پایین‌تر، دانش‌آموز، تجمیع می‌شوند و واحد سطح بالاتر، مدرسه را می‌سازند و لذا واحدهای سطح بالاتر به‌عنوان واحد تحلیل در نظر گرفته می‌شوند. تجمیع کردن داده‌ها در سطح بالاتر و سپس انجام تحلیل با استفاده از داده‌های تجمیع شده باعث بیش برآورد ضرایب خواهد شد (رابینسون^۵، ۱۹۵۰). در تحلیل داده‌های تجمیعی نیز مشکل اساسی این است که محققان فرض می‌کنند که روابط در سطح گروهی با سطح فردی یکسان است که معمولاً چنین نیست. لذا این نبود تمایز بین سطوح به خطای بوم‌شناسی^۶ منجر می‌شود (گرینلند^۷، ۲۰۰۱؛ گرینلند و مورگنسترن^۸، ۱۹۸۹؛ گرینلند و رابینز^۹، ۱۹۹۴؛ مورگنسترن و

1. Disaggregate analysis
2. Aggregate analysis
3. Cross-level impact
4. Paterson
5. Rabinson
6. Ecological fallacy
7. Greenland
8. Morgenstern
9. Robins

توماس^۱، (۱۹۹۳). اغلب محققان به مفهوم خطای بوم‌شناسی توجه ندارند. مثلاً برخی از محققان جدول‌های آماری مربوط به درآمد گروه‌های مختلف مردم را در نظر می‌گیرند و سپس شواهد را برای افراد در همان گروه به کار می‌برند و بیان می‌کنند فرد پولدار، پولدارتر می‌شود و فقیر، فقیرتر می‌شود (هاکر^۲، ۱۹۹۷). در حالی که شواهد حاکی از آن است که افراد با درآمد پایین با به دست آوردن تحصیلات و تجربه در طی زمان وارد گروه‌های بالاتر می‌شوند و برخی از افراد در گروه‌های بالاتر با افزایش سن و بازنشستگی درآمدهای خود را کاهش پیدا می‌کنند (سوول^۳، ۲۰۰۹). خطای بوم‌شناسی همچنین به تورش بوم‌شناسی^۴، تورش بین سطحی^۵، یا تورش تجمیعی^۶ نیز شناخته می‌شود (گرینلند، ۲۰۰۱؛ گرینلند و مورگنسترن، ۱۹۸۹؛ گرینلند و روبینز^۷، ۱۹۹۴؛ مورگنسترن و توماس^۸، ۱۹۹۳) که به تفاوت بین ضرایب به دست آمده در سطح تجمیع و ضرایب به دست آمده در سطح فردی اشاره دارد.

رابینسون (۱۹۵۰) که تأثیر زیادی در فهم خطای بوم‌شناسی داشته است (واژه‌ای که توسط سلوین^۹، ۱۹۵۸ ابداع شده است) بیان می‌کند که خطای بوم‌شناسی به الگوی روابط نمایش شده در سطح تجمیع اشاره دارد که نمی‌تواند به‌سادگی به رفتار افراد تعمیم داده شود. به‌ویژه، ضرایب بوم معمولاً بیشتر از ضرایب فردی هستند. از نظر مفهومی، اکز^{۱۰} (۲۰۰۹) نشان داده است که: اول، تحلیل‌های مبتنی بر اندازه‌گیری‌های تجمیع شده به‌سختی روابط درست در سطح فردی را ارائه می‌دهند. دوم، حتی با فرض نرمال بودن توزیع‌های سطح فردی، روابط سطح گروهی مساوی با روابط سطح فردی نیستند. در نتیجه به دلیل تفاوت‌های بین تحلیل‌های سطح فردی و سطح گروهی، داده‌های بوم‌شناسی به نتایج گمراه‌کننده برای سطح فردی منجر می‌شود و در

-
1. Thomas
 2. Hacker
 3. Sowell
 4. Ecological bias
 5. Cross-level bias
 6. Aggregation bias
 7. Robins
 8. Thomas
 9. Selvin
 10. Oakes

بهترین وضعیت، محقق می‌تواند فرض کند که تحلیل‌های بوم‌شناسی و سطح فردی برابر هستند.

دلایل مختلفی برای تورش و سوگیری در تحلیل‌های تجمیعی وجود دارد که از جمله این دلایل: اول، دشوار بودن شناسایی و کنترل متغیرهای مشتبه‌کننده^۱ است که در هر دو سطح فردی و زمینه‌ای می‌تواند اتفاق بیفتد. در سطح فردی برآوردهای بوم‌شناسی زمانی رخ می‌دهد که افراد به‌طور مساوی به وسیله متغیرهای مشتبه‌کننده تحت تأثیر نیستند و ممکن است در برخی گروه‌ها اثرات مشتبه‌کننده‌ها مثبت و در برخی دیگر منفی باشد. حتی اگر اثر متغیر مشتبه‌کننده در گروه‌ها یکسان باشد برآوردهای بوم‌شناسی باز هم تورش‌دار هستند. به عبارتی چون متغیرهای مشتبه‌کننده به‌طور مشابه در همه اعضای گروه‌ها توزیع نشده‌اند به سوگیری در تحلیل‌های بوم‌شناسی منجر می‌شود (گرینلند، ۲۰۰۱). در سطح زمینه‌ای، تورش بوم‌شناسی وقتی رخ می‌دهد که متغیرهای آمیخته زمینه‌ای به‌طور متفاوت در گروه‌ها توزیع شده باشند، حتی اگر به‌عنوان متغیر آمیخته در سطح اول هم نباشند. تغییر معنی متغیرها بعد از تجمیع کردن آنها دلیل دیگر خطای بوم‌شناسی است. یک متغیر تجمیع شده ممکن است سازه متفاوتی را از سطح اول اندازه‌گیری کند (گرینلند و مورگنسترن، ۱۹۸۹).

همان‌طور که گفته شد هر دو روش تجمیع و تجمیع‌زدایی دارای خطاهایی هستند لذا روش مدل‌یابی چندسطحی به‌عنوان روشی جایگزین مطرح است. اساس روش تحلیل چندسطحی نسبتاً ساده است. مدل‌های چندسطحی مشخص می‌کند که چطور متغیرهای اندازه‌گیری شده در هر سطح تجمیع بر توزیع پیامدها در سطح دانش‌آموز اثر می‌گذارد (بارستین، لین و کیپل^۲، ۱۹۷۸). در نتیجه، در تحلیل چندسطحی، اثرات مدرسه فقط مشخص‌کننده واریانس بین مدرسه‌ها^۳ نیست بلکه واریانس بین دانش‌آموزان^۴ در هر مدرسه را نیز مشخص می‌کند (براون و ساکس^۵، ۱۹۸۱). تکنیک مدل‌یابی چندسطحی، اثر زمینه گروه بر افراد ارزیابی و همچنین تفاوت‌های فردی، مطالعه می‌شود. در مقایسه با تحلیل تجمیع، این تکنیک افراد درون یک گروه را متفاوت می‌داند و به زمینه گروه واکنش متفاوتی نشان می‌دهد. تکنیک مدل‌یابی

1. Confounding factors

2. Burstein, Linn & Capell

3. Between schools variance

4. Between students variance

5. Brown & Saks

چندسطحی در مقایسه با تحلیل تجمیع‌زدایی، افراد درون یک گروه را شبیه‌تر به هم نسبت به افراد گروه‌های دیگر می‌داند. مدل‌یابی چندسطحی مقایسه با تحلیل تجمیع و تجمیع‌زدایی دارای مزیت‌هایی است: اول، مزیت آشکار مدل‌یابی چندسطحی این است که محققان می‌توانند به فرمول‌بندی و آزمون مدل‌های بین سطحی بدون نگرانی در مورد واحد تحلیل اقدام کنند. مدل‌یابی چندسطحی ساختار سلسله‌مراتبی داده‌ها را در نظر گرفته و ارائه‌دهنده زیرمدل^۱ برای هر سطح است. این زیرمدل‌ها بیان‌کننده روابط بین متغیرها در درون سطح فرض شده است و مشخص می‌کند که چطور متغیرها در یک سطح بر روابط بین متغیرها در سطح دیگر اثر می‌گذارد. همچنین داده‌های وارد شده در مدل در سطح خاص و مدل‌ها به محققان امکان بررسی واریانس هر دو سطح پایین‌تر و بالاتر را می‌دهد. بنابراین، مدل‌یابی چندسطحی بر مشکلات تجمیع‌زدایی متغیرهای گروه به افراد و تجمیع داده‌های فرد به گروه غلبه می‌کند. دوم، مدل‌یابی چندسطحی به محققان امکان برآورد مؤلفه‌های واریانس و کواریانس داده‌های آشیانه‌ای غیر متعادل^۲ را می‌دهد؛ چراکه در این روش امکان جداسازی تغییرپذیری به مؤلفه‌های درون و بین گروهی امکان‌پذیر است. مؤلفه درون‌گروهی، روابط درون گروه‌ها را توصیف می‌کند که امکان مقایسه افرادی که در گروه مشابه هستند مثل کلاس‌ها یا مدرسه‌ها را فراهم می‌کند. مؤلفه بین‌گروهی به تفاوت بین گروه‌ها اشاره دارد و امکان مقایسه بین کلاس‌ها و مدرسه‌ها را برای محقق فراهم می‌کند. به‌طور طبیعی، گام اول در تحلیل چندسطحی، جداسازی مؤلفه‌های واریانس درون و بین گروهی است. در حقیقت تحلیل چندسطحی تفاوت‌های اساسی بین مدرسه‌ها را مشخص می‌کند که تحلیل‌های مرسوم قادر به کشف آن نیستند. نهایتاً، مهم‌ترین مزیت مدل‌یابی چندسطحی افزایش دقت هر دو ضرایب رگرسیون و خطاهای استاندارد است. داده‌های خوشه‌بندی شده، به دلیل ماهیتشان، مفروضه‌های همگنی و استقلال را نقض می‌کنند که این مفروضه‌ها اساس تحلیل‌های حداقل مجذورات معمولی^۳ هستند. ولی این نقض در تحلیل چندسطحی مشکلی ایجاد نمی‌کند. چراکه تحلیل چندسطحی به مفروضه‌های اساسی (مثل نرمال بودن، خطی

1. Submodel

2. Unblanced

3. Ordinary least squares

بودن^۱، فاصله مساوی^۲) که شبیه به مفروضه‌های مورد نیاز در تحلیل‌های پارامتریک مرسوم است نیاز کمتری دارد و مفروضه‌های همگنی و استقلال که برای تحلیل‌های چندسطحی مناسب است؛ چراکه در این تحلیل‌ها معادلات برای هر گروه به‌طور جداگانه برآورد می‌شود و مفروضه همگنی واریانس ممکن است در درون هر گروه صحیح باشد ولی ضرورتاً در بین گروه‌ها صحیح نیست. همچنین یک اثر تصادفی واحد برای هر گروه درون طرح آماری ترکیب می‌شود. همچنین اگر گروه‌ها از نظر تعداد غیر متعادل هستند، تحلیل چندسطحی برآورد اثرات گروه‌های کوچک را با تعدیل آنها برای شبیه شدن به اثرات نمونه کل بهبود می‌بخشد و در نهایت می‌توان گفت که تحلیل چندسطحی برآوردها را در مقایسه با طرح‌های تجمیع و تجمیع‌زدایی بهبود می‌بخشد (اتینگتون^۳، ۱۹۹۷)

بریک و رادنبوش (۱۹۹۲) با مقایسه کاربرد سه مدل تجمیع، تجمیع‌زدایی و تحلیل چندسطحی، رابطه ویژگی‌های مدرسه و احساس خودپنداره معلم را بررسی کردند. مقایسه سه روش نشان داد که در تحلیل تجمیع‌زدایی ضرایب برآورد شده قوی هستند و عموماً شبیه به ضرایبی هستند که با کاربرد تحلیل چندسطحی به دست می‌آید. اندازه اثرات همان متغیر بر متغیرهای دیگر شبیه هستند. ولی اثر به‌احتمال بیشتری در تحلیل سطح فردی از نظر آماری معنی‌دار است چراکه خطاهای استاندارد به علت وابستگی بین افراد در گروه‌ها کمتر برآورد می‌شوند. لذا، محقق به‌احتمال بیشتری نتیجه می‌گیرد که یک اثر از نظر آماری معنی‌دار است در حالی که واقعاً چنین نیست و به نتیجه‌گیری غلط منجر می‌شود. در کاربرد تحلیل تجمیع، ضرایب قوی نیستند، لذا محققان نمی‌توانند به برآوردهای پارامتر اعتماد کنند. از سوی دیگر، تحلیل‌های چندسطحی، دقیق‌ترین برآوردها را از ضرایب و بهترین برآوردها از خطاهای استاندارد از درجه وابستگی درون گروه‌ها و عدم تعادل کوچک بین گروه‌ها را فراهم می‌کنند.

با وجود تفاوت زیاد بین این سه روش در برآوردهای نسبت واریانس تبیین شده در مثال بریک و رادنبوش (۱۹۹۲) در کاربرد تحلیل فردی، فقط ۵/۴ در صد از واریانس در متغیر وابسته به وسیله ویژگی‌های مدرسه تبیین می‌شود؛ چراکه در آن

1. linearity
2. Equal interval
3. Ethington

سطح، تغییرپذیری بین و درون‌گروهی را ترکیب می‌کند. در کاربرد تحلیل تجمیع، نسبت واریانس تبیین شده متغیر وابسته ۴۲/۶ بود چراکه تغییرپذیری بین افراد در نظر گرفته نشد و تنها تغییرپذیری بین گروهی، ارزیابی می‌شد. ولی در تحلیل چندسطحی، ۶۳/۱ درصد از واریانس متغیر وابسته تبیین شد چون برخلاف تحلیل تجمیع‌زدایی، از وابستگی بین افراد در درون گروه‌ها تأثیر نمی‌پذیرد و برخلاف تحلیل تجمیع، تحلیل چندسطحی از عدم اعتبار ایجاد شده به وسیله تجمیع داده‌ها تأثیر نمی‌پذیرد. مزیت‌های تحلیل چندسطحی نسبت به سایر تحلیل‌ها در پژوهش‌های مختلف شبیه‌سازی شده حمایت شده است (مثل بارسیکوسکی^۱، ۱۹۸۱؛ باسینگ^۲، ۱۹۹۳؛ گولدستین^۳، ۱۹۹۵؛ گولدستین و راسباش^۴، ۱۹۹۶، کیم^۵، ۱۹۹۰؛ کیش^۶، ۱۹۶۵؛ کرفت^۷، ۱۹۹۶؛ رودریگز و گلدمن^۸، ۱۹۹۵، تات و وانگباندھیت^۹، ۱۹۸۳؛ ون در لییدن و باسینگ^{۱۰}، ۱۹۹۴). در ایران نیز پژوهش‌هایی در زمینه کاربرد تحلیل چندسطحی در داده‌های تیمز اجرا شده است که از جمله می‌توان به پژوهش نقش و مقدم (۱۳۹۱)؛ کیامنش، حجازی و نقش (۱۳۹۵) اشاره کرد.

روش

پژوهش حاضر از نوع توصیفی و به‌طور دقیق‌تر همبستگی است. جامعه این پژوهش را دانش‌آموزان ایرانی پایه هشتم در سال تحصیلی ۹۰-۱۳۸۹، تشکیل می‌دهد که در مطالعه تیمز ۲۰۱۱ شرکت کرده‌اند. تیمز ۲۰۱۱ شبیه به مطالعات قبلی تیمز (۱۹۹۵، ۱۹۹۹ و ۲۰۰۳ و ۲۰۰۷) برای اطمینان از اینکه داده‌های نمونه معرف جامعه دانش‌آموزان ملی است از روش نمونه‌گیری خوشه‌ای طبقه‌ای دو مرحله‌ای^{۱۱} استفاده

1. Barcikowski
2. Busing
3. Goldstein
4. Rasbash
5. Kim
6. Kish
7. Kreft
8. Rodriguez & Goldman
9. Tate & Wongbunhit
10. Van Der Leeden & Busing
11. Two-stage Stratified Cluster Design

می‌کند (به این معنا که سهم هر یک از خوشه‌ها^۱ در نمونه، متناسب با حجم آن در جامعه است). افزون بر آن با بهره‌گیری از وزن‌های نمونه‌گیری^۲ اطمینان حاصل می‌شود که شاخص‌های آماری به دست آمده از نمونه، معرف جامعه مورد نظر هستند. در مرحله اول، مدرسه‌ها با روش احتمال متناسب با حجم^۳ نمونه‌گیری شدند سپس در هر مدرسه منتخب، از میان همه کلاس‌های پایه هشتم یک کلاس با روش تصادفی سیستماتیک انتخاب شده و در نهایت، همه دانش‌آموزان با احتمال مساوی از کلاس‌های نمونه‌گیری شده در آزمون شرکت کردند. در ایران در سطح مدرسه یا سطح ۲، تنها یک کلاس از هر مدرسه انتخاب می‌شود، لذا تعداد مدرسه‌ها با تعداد کلاس‌ها مساوی است. تعداد نمونه تیمز ۲۰۱۱ از ۶۰۲۹ دانش‌آموز (۲۸۱۶ دختر و ۳۲۱۳ پسر) تشکیل شده است.

تیمز برای جمع‌آوری اطلاعات در مورد زمینه‌های آموزشی برای تدریس و یادگیری ریاضیات و علوم، سه پرسشنامه دارد: پرسشنامه دانش‌آموز، پرسشنامه معلم و پرسشنامه مدرسه. در این پژوهش، به منظور جمع‌آوری اطلاعات از پرسشنامه دانش‌آموز و معلم استفاده شده است. متغیر وابسته این پژوهش پیشرفت ریاضیات دانش‌آموزان پایه هشتم است. متغیرهای مستقل در سطح دانش‌آموز متغیر مقدار تکلیف در هفته و مدت زمان انجام تکلیف است و در سطح مدرسه/معلم شامل متغیرهای مقدار تکلیفی که معلم به دانش‌آموز می‌دهد و مدت زمانی است که معلم برای انجام تکلیف به دانش‌آموز می‌دهد.

محاسبه اعتبار^۴ و روایی^۵ سؤال‌های تیمز بر اساس شاخص‌های روان‌سنجی از طریق انجام آزمون‌های مقدماتی^۶ در کشورهای شرکت‌کننده انجام می‌گیرد و پس از تعیین درجه دشواری و قدرت تشخیص برای هر یک از سؤال‌های چندگزینه‌ای و پاسخ باز به تفکیک هر یک از کشورها در قالب گزارش آماری منتشر می‌شود و در اختیار کشورهای عضو قرار می‌گیرد. پس از تجزیه و تحلیل شاخص‌های روان‌سنجی، سؤال‌هایی که از نظر اعتبار و روایی، شرایط لازم را نداشته باشند حذف و سؤال‌های

1. Strata
2. Sampling Weights
3. Probability Proportional to Size (PPS)
4. Reliability
5. Validity
6. Field test

دیگر جایگزین می‌شود. بنابراین، شاخص‌های آماری مربوط به روایی و اعتبار سؤال‌های تیمز برای تمام کشورهای شرکت‌کننده از جمله ایران محاسبه می‌شود. در این پژوهش از تحلیل چندسطحی (دوسطحی) با استفاده از نرم‌افزار HLM و به‌منظور وزن‌دهی از روش MATWGT استفاده شده است.

نتایج

در جدول (۱) میانگین و انحراف معیار متغیرهای پژوهش ارائه شده است.

جدول (۱) میانگین و انحراف معیار زمان انجام تکلیف و عملکرد ریاضیات

سطح	متغیر	میانگین	انحراف معیار
دانش‌آموز	عملکرد ریاضیات	۴۱۴/۹۶	۹۰/۳۸
دانش‌آموز	مقدار تکلیف در هفته	۲/۵۶	۰/۸۸۳
دانش‌آموز	مدت‌زمان انجام تکلیف	۳/۵۴	۱/۱۰۴
معلم	مقدار تکلیف در هفته	۳/۱۴	۰/۷۶۳
معلم	مدت‌زمان انجام تکلیف	۲/۷۵	۰/۸۳۳

در ابتدا و قبل از اجرای تحلیل چندسطحی و به‌منظور مقایسه سه رویکرد تجمیع، تجمیع‌زدایی و چندسطحی بدون توجه به اینکه داده‌های متغیرهای پژوهش در تیمز در دو سطح جمع‌آوری شده‌اند و صرفاً به‌منظور مقایسه سه رویکرد مورد مطالعه با توجه به متغیر مدت‌زمان انجام تکلیف، سه تحلیل اجرا شد. اول، رویکرد تجمیع‌زدایی به‌منظور تحلیل داده‌ها استفاده شد. در این رویکرد تمام متغیرهای سطح بالاتر در سطح فردی در نظر گرفته می‌شوند. نمره مدت‌زمان انجام تکلیف که معلم مشخص کرده بود برای همه دانش‌آموزان آن معلم در نظر گرفته شد و تحلیل در سطح دانش‌آموز انجام گرفت. در مرحله بعد برای تحلیل تجمیعی، نمره مدت‌زمان انجام تکلیف که توسط دانش‌آموز پاسخ داده شده بود تجمیع و نمره‌های واحد سطح بالاتر یعنی معلم با تجمیع داده‌ها ساخته شد و سطح معلم به‌عنوان واحد تحلیل در نظر گرفته شد. در سومین تحلیل یعنی تحلیل چندسطحی، مدت‌زمان انجام تکلیف در دو سطح در نظر گرفته شد.

در مقایسه برآورد اندازه اثر ثابت و خطای استاندارد سه تحلیل، نتایج نشان داد که برآورد اثر ثابت در هر سه تحلیل مشابه است. تحلیل سطح دانش‌آموزی یا تجمیع زدایی کوچک‌ترین (۰/۱۲) و تحلیل سطح کلاسی (۰/۱۶) بزرگ‌ترین اثر را ایجاد می‌کند و مقدار آن در برآورد سلسله مراتبی است که در بین این دو عدد قرار می‌گیرد (۰/۱۴۸).

در مقایسه خطاهای استاندارد در تحلیل تجمیع‌زدایی خطاهای استاندارد به علت وابستگی بین افراد درون گروه‌ها کمتر برآورد شده است (۰/۰۶۴). بنابر این، محقق به احتمال بیشتری نتیجه می‌گیرد که یک اثر از نظر آماری معنی‌دار است در حالی که واقعا چنین نیست و به نتیجه‌گیری غلط منجر می‌شود. در تحلیل تجمیع یا تحلیل کلاس ضریب خطای استاندارد ۰/۱۲۷ به دست آمد. خطای استاندارد در تحلیل چندسطحی ۰/۰۹ به دست آمد. خطای استاندارد در تحلیل سلسله‌مراتبی همان‌طور که انتظار می‌رفت کمتر از خطای استاندارد در سطح کلاسی است.

در مقایسه نسبت واریانس تبیین‌شده در هر سه مدل در تحلیل تجمیع‌زدایی یا سطح دانش‌آموز، فقط ۴/۴ درصد از واریانس در متغیر وابسته به‌وسیله ویژگی‌های مدرسه تبیین می‌شود؛ چراکه در آن سطح، تغییرپذیری بین و درون‌گروهی ترکیب می‌شود. در تحلیل تجمیع، نسبت واریانس تبیین شده متغیر وابسته ۳۲/۲ بود چون تغییرپذیری بین افراد در نظر گرفته نشد و تنها تغییرپذیری بین گروهی مورد ارزیابی قرار می‌گرفت. ولی در تحلیل چندسطحی، ۴۳/۸۱ درصد از واریانس متغیر وابسته تبیین شد چون برخلاف تحلیل تجمیع‌زدایی از وابستگی بین افراد درون گروه‌ها تأثیر نمی‌پذیرد و برخلاف تحلیل تجمیع، تحلیل چندسطحی از عدم اعتبار ایجاد شده به‌وسیله تجمیع داده‌ها تأثیر نمی‌پذیرد.

در ادامه با توجه به اینکه داده‌های تیمز در هر دو سطح جمع‌آوری شده‌اند و به‌منظور بررسی دقیق‌تر مراحل و خروجی‌های تحلیل چندسطحی، تحلیل داده‌ها با رویکرد دوسطحی اجرا می‌شود. به این ترتیب هر یک از سؤال‌های پژوهش، مطرح و سپس با استفاده از داده‌های آماری، تحلیل می‌شوند. درنهایت، با استفاده از نتایج به دست آمده، به مسئله پژوهش پاسخ داده می‌شود.

آیا عملکرد ریاضیات دانش‌آموزان پایه هشتم در مدرسه‌ها متفاوت است؟ در این پژوهش، ابتدا برای پاسخ به پرسش بالا یک تحلیل غیر شرطی (مدل آنوا یک‌راهه با

اثرات تصادفی^۱ اجرا می‌شود (مدل A). این مدل برآوردی از نسبت واریانس بین کلاس‌ها در عملکرد را فراهم می‌آورد که همان ضریب همبستگی بین کلاسی (ICC) است. تحلیل واریانس یک‌راهه با تأثیرات تصادفی، اطلاعات مقدماتی مفیدی را راجع به چگونگی تغییرات در پیامدهای نهفته درون و بین مدرسه‌ها و اعتبار هر یک از میانگین نمونه مدرسه‌ها به صورت برآوردی از میانگین جمعیت آن ارائه می‌دهد.

جدول (۲) نتایج مدل آنوا یک راهه (مدل A)

		SE	ضرایب	اثر ثابت
		۴/۲۸۹۹۸	۴۱۳/۸۴۰۸۱۵	متوسط میانگین کلاس
سطح معنی‌داری	خی دو	درجه آزادی	مؤلفه واریانس	اثرات تصادفی
۰/۰۰۱	۵۵۱۷/۶۲۴۴	۲۲۶	۳۹۹۵/۶۴۴۲۴	میانگین کلاس
			۴۲۳۵/۱۶۹۷۸	تأثیر سطح دانش‌آموز

در مدل آنوا یک‌راهه با اثرات تصادفی ضریب همبستگی بین مدرسه‌ها با توجه به فرمول $\rho = \tau_{00} / (\tau_{00} + \sigma^2)$ به دست آمد.

$$۳۹۹۵/۶۴۴۲۴ / (۳۹۹۵/۶۴۴۲۴ + ۴۲۳۵/۱۶۹۷۸) = ۴۸/۵\%$$

بنابراین کلاس‌ها حدود ۴۹ درصد از واریانس عملکرد را تبیین می‌کنند و ۵۱ درصد برای عوامل سطح دانش‌آموزان تبیین شده است. همچنین با توجه به اینکه مقدار خی دو χ^2 به دست آمده در جدول (۲) در سطح مدرسه‌ها معنی‌دار است می‌توان گفت متوسط عملکرد ریاضی دانش‌آموزان مدرسه‌های مختلف به طور معنی‌داری با هم تفاوت دارد. مقدار اعتبار به دست آمده ۰/۹۵۲ نشان می‌دهد که میانگین نمونه مورد نظر معتبر بوده و می‌تواند به عنوان شاخصی از میانگین‌های کلاس‌های واقعی باشد. نتایج مدل غیر شرطی نشان داد که عملکرد ریاضی به طور معنی‌داری بین

¹ One way Anova with random effects model

کلاس‌ها متفاوت است. مدل‌های بعدی برای تبیین واریانس در عملکرد ریاضی با پیش‌بینی کننده‌های سطح دانش‌آموز و سطح کلاس/معلم اجرا شدند. چه مقدار از واریانس عملکرد ریاضی دانش‌آموزان پایه هشتم کشور ایران، مربوط به متغیرهای مقدار تکلیف در هفته و مدت زمان انجام تکلیف در سطح دانش‌آموز است؟ در مدل (B) که مدل عرض از مبدأ تصادفی^۱ با متغیرهای سطح دانش‌آموزان است به این سؤال پاسخ می‌دهد. از آنجایی که هیچ فرض قبلی درباره تفاوت بین مدرسه‌ها و متغیرهای پیش‌بین در این مطالعه وجود ندارد، بخش تصادفی در شیب‌ها وابسته نیست. به عبارت دیگر، فقط عرض از مبدأ در همه مدرسه‌ها متفاوت است اما دیگر ضرایب سطح دانش‌آموز در مفهوم باقی می‌ماند.



¹. Random – intercept model with only student- level variable

جدول (۳) اثرات پیش‌بینی کننده‌ها بر عملکرد دانش‌آموزان برای مدل B و C

مدل C			مدل B					
سطح معنی داری	SE	B	سطح معنی داری	SE	B	اثر ثابت		
						سطح ۱		
۰/۰۰۱	۱/۱۷۲۸۴۱	۹/۳۶۲۵۲۲	۰/۰۰۱	۱/۱۷۲۸۴۱	۹/۳۶۲۵۲۲	مقدار تکلیف در هفته		
۰/۰۰۱	۰/۸۵۰۳۵۵	۴/۸۷۲۳۸	۰/۰۰۱	۰/۸۵۰۳۵۵	۴/۸۷۲۳۸	مدت‌زمان انجام تکلیف		
						سطح ۲		
۰/۱۰۹	۷/۰۶۸۸۱۵	۱۱/۳۶۸۴۱				مقدار تکلیف در هفته		
۰/۹۳۹	۵/۱۱۱۰۴۴	۰/۳۹۵۸۹				مدت‌زمان انجام تکلیف		
سطح معنی داری	خی دو	درجه آزادی	مؤلفه واریانس	سطح معنی داری	خی دو	درجه آزادی	مؤلفه واریانس	اثر تصادفی
۰/۰۰۱	۵۴۶۶۳	۲۲۴	۳۹۶۷/۸۳۸	۰/۰۰۱	۵۶۱۶/۹۳۷	۲۲۶	۳۹۹۹/۲۳۶	میانگین مدرسه
			۴۱۶۰/۶۳۳				۴۱۶۰/۳۸۵	تأثیر سطح اول

جدول (۳) ارائه‌دهنده نتایج مدل B برای دانش‌آموزان پایه هشتم کشور ایران است. مدل غیر شرطی، اساس و پایه‌ای برای محاسبه نسبت کاهش واریانس در مدل حاضر و مدل‌های بعدی است. با مقایسه مؤلفه واریانس مدل B و مدل غیر شرطی، شاخص نسبت کاهش در واریانس یا واریانس تبیین شده در سطح دانش‌آموز از فرمول زیر استفاده می‌شود:

$$\left(\frac{\sigma^2 \text{ null model} - \sigma^2 \text{ estimated model}}{\sigma^2 \text{ null model}} \right)$$

با اضافه شدن عوامل پیش‌بین در سطح دانش‌آموز، مؤلفه‌های واریانس در سطح دانش‌آموز کاهش یافته است. نسبت کاهش واریانس در سطح دانش‌آموز به صورت زیر به دست آمد:

$$0/017 = 4235/16978 - (4235/16978 - 4160,38545)$$

که این نشان می‌دهد حدود ۲ درصد از واریانس کل عملکرد ریاضی در سطح دانش‌آموز توسط متغیر مقدار تکلیف و مدت انجام تکلیف تبیین می‌شود. همان‌طور که در جدول فوق مشاهده می‌شود رابطه هر دو متغیر با عملکرد ریاضی در سطح ۰/۰۰۱ معنی‌دار است.

چقدر رابطه بین عملکرد دانش‌آموز بین کلاس‌ها مشابه است؟ برای پاسخ به این سؤال، مدل بعدی یعنی مدل عرض از مبدأ تصادفی و شیب‌های تصادفی با متغیرهای سطح دانش‌آموز و کلاس اجرا شد. برای بررسی روابط بین عوامل فردی و عملکرد ریاضی، همه عوامل سطح دانش‌آموز و مدرسه/معلم به‌عنوان پیش‌بینی کننده وارد مدل شدند و مدل عرض از مبدأ تصادفی با متغیرهای سطح دانش‌آموز و کلاس (مدل C) اجرا شد.

در مدل ضرایب تصادفی با متغیرهای سطح دانش‌آموز و کلاس، متغیرهای سطح دانش‌آموز (مقدار تکلیف و مدت‌زمان انجام تکلیف) و همچنین متغیرهای سطح کلاس/معلم (مقدار تکلیف و مدت‌زمان انجام تکلیف) وارد مدل شدند. چون هیچ فرضیه‌ای راجع به تفاوت بین کلاس‌ها در متغیرهای پیش‌بینی کننده در این مطالعه وجود نداشت لذا شیب‌ها بین کلاس‌ها متفاوت نبوده و ثابت در نظر گرفته شدند. همچنین در این مطالعه اثرات تعاملی بین متغیرها بررسی نشده است. نتایج مدل ضرایب تصادفی با متغیرهای سطح دانش‌آموز و کلاس در جدول (۳) نشان می‌دهد که دو متغیر مقدار تکلیف و مدت انجام آن در سطح دانش‌آموز رابطه معنی‌داری با عملکرد ریاضی دارد ولی رابطه این دو متغیر در سطح مدرسه/معلم معنی‌دار نیست.

مقدار ضریب همبستگی بین کلاسی با توجه به فرمول $\rho = \tau_{00} / (\tau_{00} + \sigma^2)$ به دست آمد.

$$\rho = 3967/838 / (3967/838 + 4160/633) = 0/488$$

ضریب همبستگی بین کلاسی بعد از اضافه شدن متغیرها در هر دو سطح تغییر زیادی نکرده است.

مقدار خبی دو نشان می‌دهد که تفاوت معنی‌دار بین کلاس‌ها در عملکرد همچنان باقی است. در مجموع در مدل کامل یا مدل C، ۴۹ درصد و ۵۱ درصد از واریانس عملکرد ریاضی را به ترتیب در سطح کلاس/معلم و دانش‌آموز تبیین می‌کند. همان‌طور که مشاهده شد با انجام تحلیل دوسطحی و در نظر گرفتن داده‌ها در دو سطح، نتایج متفاوتی در دو سطح به دست آمد. همچنین مشاهده شد که هر چه قدر به دانش‌آموزان تکلیف بیشتری داده شود و مدت‌زمان بیشتری صرف انجام تکالیف داشته باشند، عملکرد بهتری دارند ولی این رابطه در سطح بالاتر دیده نشد یعنی مقدار تکلیف و مدت‌زمان انجام تکلیف که معلم به دانش‌آموزان می‌دهد رابطه‌ای با عملکرد دانش‌آموزان ندارد.

بحث و نتیجه‌گیری

پژوهش حاضر با هدف مقایسه روش‌های تجمیع، تجمیع‌زدایی و تحلیلی چندسطحی انجام گرفت. همان‌طور که بیان شد هر دو روش تجمیع و تجمیع‌زدایی دارای خطاهایی هستند. از این‌رو، روش مدل‌یابی چندسطحی به‌عنوان روش جایگزین مطرح است. مدل‌های چندسطحی مشخص می‌کند که چطور متغیرهای اندازه‌گیری شده در هر سطح تجمیع بر توزیع پیامدها در سطح دانش‌آموز اثر می‌گذارد (بارستین، لین و کپل، ۱۹۷۸). در نتیجه، در تحلیل چندسطحی، اثرات مدرسه فقط مشخص‌کننده واریانس بین مدرسه‌ها نیست بلکه واریانس بین دانش‌آموزان درون هر مدرسه را نیز مشخص می‌کند (براون و ساکس، ۱۹۸۱). مدل‌های چندسطحی مزیت‌های قابل ملاحظه‌ای در مقایسه با دو تحلیل تجمیع و تجمیع‌زدایی دارد. با توسعه روش‌های چندسطحی در بررسی اثر متقابل بین سطوح، اثرات متغیرهای فردی و گروهی در هر سطح تحلیل می‌شود. در تحلیل‌های چندسطحی در برآورد اثرات بین سطوح، متغیرهای هر سطح (فردی، مدرسه) در سطح مربوط به خود تعریف و اندازه‌گیری شوند. در حال حاضر، تعدادی طرح‌های مدل‌یابی چندسطحی برای ارزیابی مدل‌های چندسطحی و مدل‌های رگرسیون چندسطحی، به‌ویژه مدل‌های خطی چندسطحی وجود دارد (هاکس، ۲۰۰۲). در این مقاله، از مدل‌یابی خطی چندسطحی برای بررسی داده‌های تیمز ۲۰۱۱ استفاده شده است. نتایج نشان داد که در استفاده از مدل‌یابی چندسطحی می‌توان مدل‌های بین سطوح را ارزیابی کرد و نتایج قابل تفسیرتری از دو تحلیل تجمیعی و تجمیع‌زدایی به دست آورد. همچنین این مدل‌ها نسبت واریانس

متغیر وابسته را در هر سطح مشخص می‌کند. مهم‌ترین برتری مدل‌های چندسطحی افزایش دقت ضرایب برآورد شده است. وقتی داده‌ها با توجه به ماهیتشان خوشه‌بندی می‌شوند، از مفروضه‌های همگنی و استقلال که اساس تحلیل‌های خطی سنتی هستند تخلف می‌کنند. در مدل‌های چندسطحی، این تخلف‌ها مشکل‌زا نیستند؛ چراکه معادلات جدا برای هر گروه نوشته می‌شود و مفروضات همگنی واریانس و استقلال درون هر گروه صحیح هستند ولی ضرورتاً بین گروه‌ها صحیح نیستند. در مدل‌های چندسطحی اثر تصادفی برای هر گروه، تغییرپذیری در اثرات تصادفی و همچنین خطاهای استاندارد محاسبه می‌شوند؛ به این معنی که اثر پارامترها و واریانس‌های خطا به‌طور جداگانه برآورد می‌شوند و به همین دلیل مدل‌های چندسطحی واقع‌بین هستند (اتینگتون، ۱۹۹۷).

در این پژوهش از تحلیل چندسطحی برای بررسی داده‌های آشیانه‌ای تیمز استفاده شد و نتایج علاوه بر نشان دادن مزیت این تحلیل بر دو تحلیل تجمیع و تجمیع‌زدایی نشان داد که بین مقدار تکلیف و مدت‌زمان انجام آن با عملکرد ریاضی در سطح دانش‌آموز رابطه معنی‌داری وجود دارد ولی بین مقدار تکلیفی که معلم به دانش‌آموزان می‌دهد و مدت‌زمان انجام آن با عملکرد ریاضی در سطح معلم رابطه معنی‌داری وجود نداشت. رسیدن به این نتیجه یعنی تحلیل رابطه در سطوح مختلف با تحلیل چندسطحی امکان‌پذیر است. تحلیل‌های چندسطحی در تحلیل‌های مختلفی مثل داده‌های آشیانه‌ای، داده‌های طولی و رشد و نمو داده‌ها نیز به کار می‌روند. محققان باید بنا به اهمیت توجه به سطوح در بسیاری از پژوهش‌ها، همان‌طور که به نمونه‌ای از آن در پژوهش حاضر اشاره شد، هر چه بیشتر به استفاده از این تحلیل‌ها در پژوهش‌های خود روی آورند.

منابع

- نقش، زهرا و مقدم، اعظم (۱۳۹۱). کاربرد تکنیک‌های مدل‌یابی چندسطحی در تحلیل داده‌های تیمز ۲۰۰۷ و مقایسه آن با تحلیل یک‌سطحی. *اندازه‌گیری تربیتی*. ۸ (۳)، ۱۳۳ - ۱۵۴.
- کیامنش، علیرضا؛ حجازی، الهه و نقش، زهرا (۱۳۹۵). رابطه مدت‌زمان انجام تکلیف و پیشرفت ریاضی: کاربرد مدل‌یابی چندسطحی برای تحلیل داده‌های تیمز ۲۰۱۱. *نوآوری‌های آموزشی*، ۶۰ (۱۵)، ۶۱ - ۷۴.
- Barcikowski, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics*, 6 (3), 267-285.
- Bronfenbrenner, U. (1976). The experimental ecology of education. *Educational Researcher*, 5 (9), 5-15.
- Brown, B. W. & Saks, D. H. (1981). The microeconomics of schooling. *Review of Research in Education*, 9 (1), 217-254.
- Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage.
- Burstein, L. (1980). The analysis of multi-level data in educational research and evaluation. *Review of Research in Education*, 8 (1), 158-233.
- Busing, F. (1993). *Distribution characteristics of variance estimates in two level models*. Leiden University, Department of Psychometrics and Research Methodology.
- Burstein, L.; Linn, R. L. & Capell, F. J. (1978). Analyzing multilevel data in the presence of heterogeneous within-class regressions. *Journal of Educational and Behavioral Statistics*, 3 (4), 347-383.
- Cronbach, L. J. (1976). *Research on classroom and schools: Formulation of questions, design and analysis*. Stanford, CA: Stanford Evaluation Consortium.
- Ethington, C. A. (1997). A hierarchical linear modeling approach to studying college effects. In J. Smart (ed), *Higher Education: Handbook of Theory and Research*, 12, (pp. 165-194). New York: Agathon.
- Goldstein, H. (1995). *Multilevel statistical models* (2nd ed). London, United Kingdom: Edward Arnold.
- Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement*, 8 (4), 369-395.
- Goldstein, H. & Rasbash, J. (1990). *Improved approximations for multilevel models with binary responses*. London, United Kingdom: University of London.

- Goodman, L. (1959). Some alternatives to ecological correlation. *American Journal of Sociology*, 64, 610-625.
- Greenland, S. (2001). Ecologic versus individual-level sources of bias in ecologic estimates of contextual health effects. *International Journal of Epidemiology*, 30 (6), 1343-1350.
- Greenland, S. & Morgenstern, H. (1989). Ecological bias, confounding, and effect modification. *International Journal of Epidemiology*, 18, 269-274.
- Greenland, S. & Robins, J. M. (1994). Ecologic studies-biases, misconceptions, and counterexamples. *American Journal of Epidemiology*, 139, 747-760.
- Hacker, A. (1997). *Money: Who has how much and why*. New York: Scribner.
- Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kim, K. S. (1990). *Multilevel data analysis: A comparison of analytical alternatives*. Los Angeles, CA: University of California.
- King, G. (1997). *A solution to the ecological inference problem: Reconstructing individual behavior from aggregate data*. Princeton, NJ: Princeton University Press.
- Kousser, J. M. (2001). Ecological inference from Goodman to King. *Historical Methods, Journal of Historical Methods*, 34 (3), 101-126.
- Kreft, Ita G. G. (1996). *Are multilevel techniques necessary? An overview, including simulation studies*. Los Angeles: California Stata University.
- Mischel, W. (1977). The interaction of person and situation. In D. Magnusson & N.S. Endler (Eds), *Personality at the crossroads. Current issues in interactional psychology* (pp. 333-352). Hillsdale, NJ: Erlbaum.
- Morgenstern, H. & Thomas, D. (1993). Principles of study design in environmental epidemiology. *Environmental Health Perspectives*, 101 (4), 23-38.
- Oakes, J. M. (2009). Commentary: Individual, ecological and multilevel fallacies. *International Journal of Epidemiology*, 38, 361-368.
- Paterson, L. (1991). An introduction to multilevel modeling. In S. W. Raudenbush & J. D. Willms (Eds). *Schools, classroom, and pupils* (pp.13-25). San Diego, CA: Academic Press, Inc.
- Raudenbush, S. W. & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education*, 59, 1-117.
- Rodriguez, G. & Goldman, N. (1995). An assessment of estimation procedures for multilevel models with binary responses. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 158 (1), 73-89.

- Robinson, W. S. (1950). Ecological correlations and behavior of individuals. *American Sociological Review*, 15, 351-357.
- Selvin, H. C. (1958). Durkheim's suicide and problems of empirical research. *American Journal of Sociology*, 63, 607-619.
- Shively, W. P. (1969). Ecological inference: The use of aggregate data to study individuals. *The American Political Science Review*, 63 (4), 1183-1196.
- Sowell, T. (2009). *The housing bottom and bust*. New York, NY: Basic Books.
- Tate, R. & Wongbundhit, Y. (1983). Random versus nonrandom coefficient models for multilevel analysis. *Journal of Educational Statistics*, 8, 103-120.
- Van Der Leeden, R. & Busing, F. (1994). *First iteration versus IGLS/RIGLS estimates in two-level models: A Monte Carlo study with ML3*. Leiden: Department of Psychometrica and Research Methodology, Leiden University.
- Webb, N. M. (1977). *Learning in individual and small group settings*. Stanford, CA: Stanford University School of Education.

