



دانشگاه علوم پزشکی و خدمات بهداشتی درمانی کرمانشاه

معاونت تحقیقات و فناوری

کارگاه مقدماتی مدل یابی معادلات ساختاری

(Structural Equation Model)

با همکاری گروه اپیدمیولوژی و گروه آمار زیستی

دانشگاه علوم پزشکی کرمانشاه

پذیره‌ها- انواع مدل‌بندی معادلات ساختاری-
تفسیر ضرایب- انواع برازش- شاخص‌های نیکویی
برازش

شایان مصطفایی
استادیار آمار زیستی

Shayan.mostafaei@kums.ac.ir

- معرفی و مزایای مدلیابی معادلات ساختاری
- پذیره‌های معادلات ساختاری
- انواع مدل‌بندی معادلات ساختاری خطی
- انواع برازش مدل‌بندی معادلات ساختاری
- آزمون و تفسیر ضرایب مسیر
- اعتبار سنجی مدل براساس شاخص‌های نیکویی برازش

معرفی مدل‌بندی معادلات ساختاری

معادلات ساختاری ترکیبی از چند روش مدل‌سازی آماری است: رگرسیون چندگانه، تحلیل مسیر و تحلیل عاملی است بطوریکه در مدل رگرسیون وجود ارتباط خطی بین مجموعه‌ی متغیرهای اندازه‌گیری شده مستقل و متغیر وابسته آزمون می‌شود. تحلیل مسیری در کنار اثر مستقیم، اثر غیرمستقیم متغیر(های) مستقل بر متغیر(های) وابسته را نیز بررسی می‌کند. تحلیل عاملی بر اساس ماتریس واریانس کوواریانس مشخص می‌کند کدام مجموعه از متغیرهای مشاهده‌شده، سازه‌ها (یا متغیر پنهان) را می‌سازند. در تحلیل عاملی برخلاف رگرسیون، متغیر مستقل و وابسته وجود ندارد.

مدل معادلات ساختاری، تعمیم مدل‌های تحلیل مسیری هستند که در آن‌ها امکان حضور متغیرهای پنهان و بررسی روابط آن‌ها هم وجود دارد. در این روش ماتریس همبستگی یا ماتریس کوواریانس از داده‌های مورد نظر محاسبه و مدل‌سازی می‌شود.

پذیره‌های مدل‌بندی معادلات ساختاری

- مشاهدات در معادلات رگرسیونی، کمی و نرمال چندمتغیره باشند
- عدم حضور مشاهدات دورافتاده یا پرت
- عدم همخطی بین متغیرها رگرسیونی در هر معادله رگرسیونی
- ارتباط خطی بین متغیرهای مشاهده شده با سازه‌ها و بین سازه‌ها
- کفایت حجم نمونه
- عدم حضور داده‌های گمشده

مزایای مدل‌بندی معادلات ساختاری

- امکان برآورد اثرات مستقیم و غیرمستقیم با حضور متغیرهای پنهان با استفاده از برازش همزمان چندین معادله رگرسیونی
- اعتبار مدل‌بانی معادلات ساختاری در مقایسه با اجرای چندین معادله رگرسیونی در استنباط همزمان
- پایایی مدل‌بانی معادلات ساختاری به دلیل وارد کردن مولفه خطای اندازه‌گیری متغیرهای مستقل و مولفه خطای تصادفی برای متغیرهای پاسخ در هر معادله رگرسیونی
- قابلیت استفاده در ساختارهای پیچیده ارتباطات

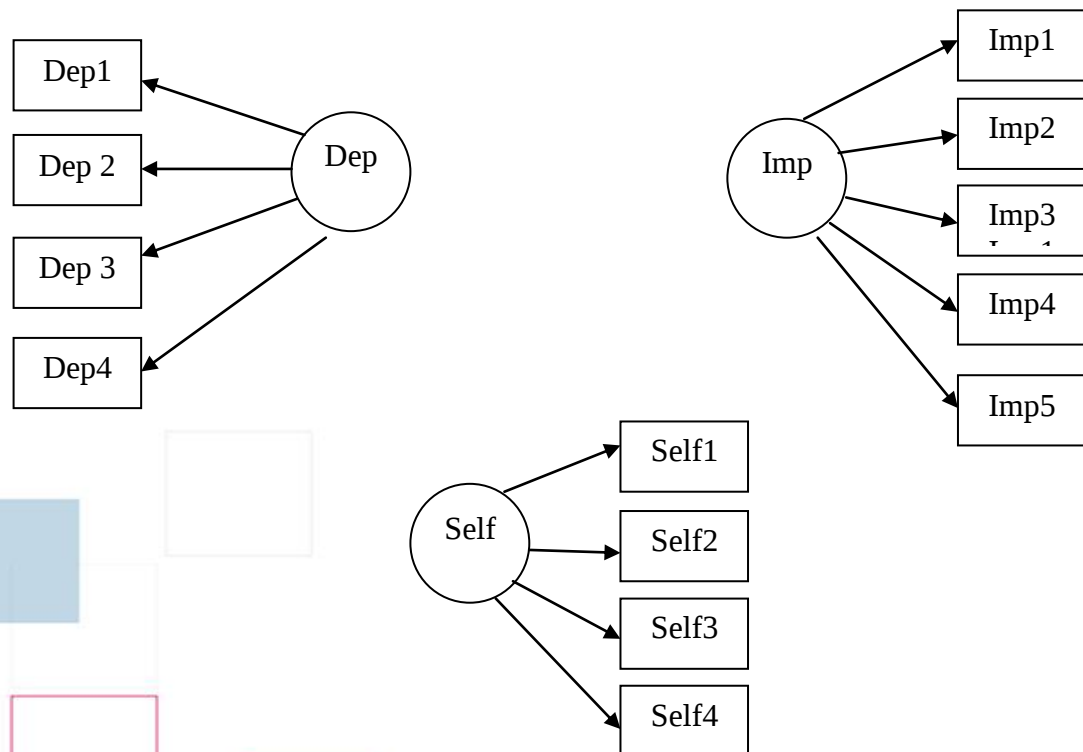
راه‌های مواجهه با محدودیت‌های معادلات ساختاری

- در سال‌های اخیر محدودیت این تحلیل در ناتوانی در کاربرد آن برای متغیرهای کیفی از طریق روش مولفه پنهان و برازش بیزی قابل حل شده است.
- محدودیت دیگر در زمینه گمشدگی و عدم کفایت حجم نمونه از طریق، جانهی کردن، روش‌های بیزی و بازنمونه‌گیری/بوت استرپ قابل حل شده است.

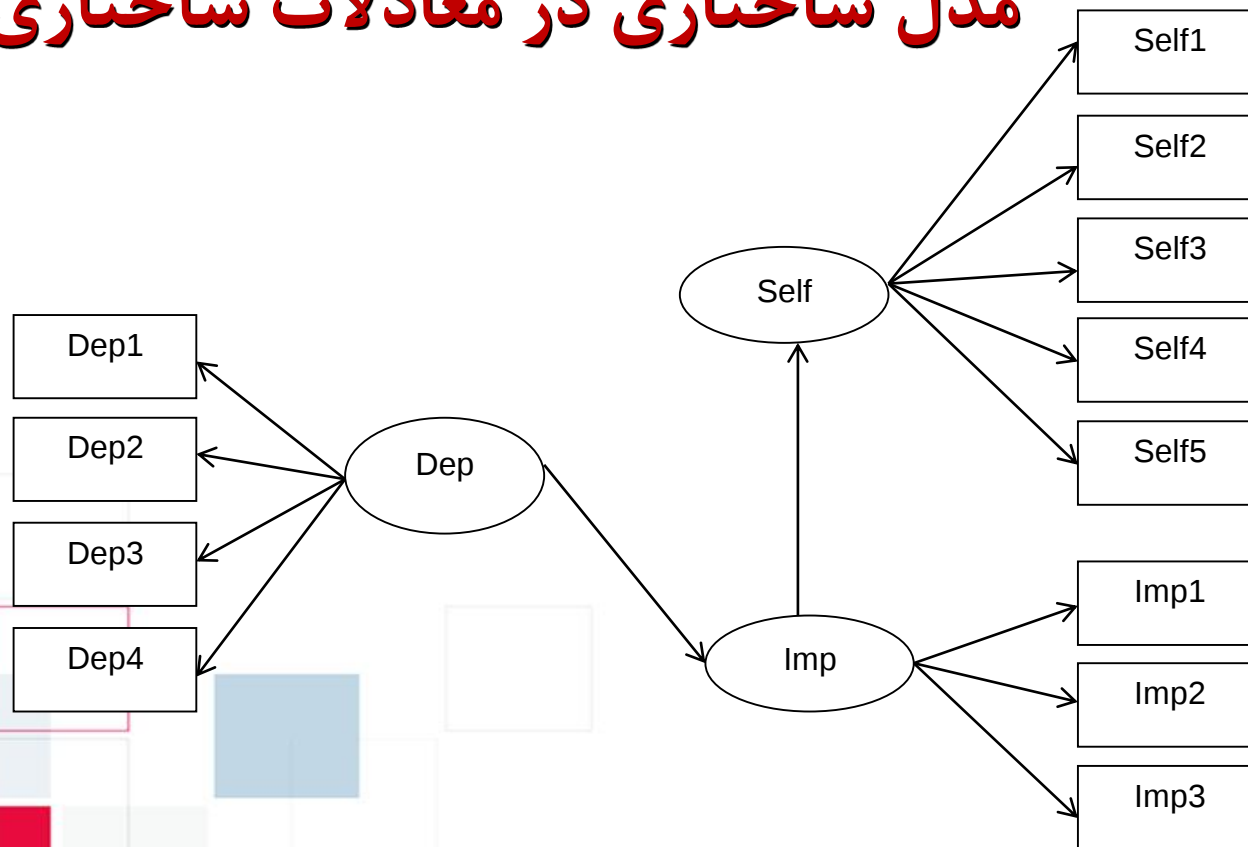
راه‌های مواجهه با محدودیت‌های معادلات ساختاری

- همچنین حساسیت این تحلیل‌ها به پذیره نرمالیتی نیز از طریق مدل‌بندی معادلات ساختاری با پاسخ غیرنرمال حل شده است. مثلاً یکی از استراتژی‌های این نوع از مدل‌بندی استفاده از تبدیل لگاریتم-مدل‌های لگ خطی، تبدیل لجیت و یا Latent Class Model است
- محاسبه مقدار احتمال وقوع هر متغیر کیفی دو حالتی با استفاده از رگرسیون لجستیک و سپس بجای متغیر مثلاً دو حالتی از مقادیر احتمال برآورد شده آن استفاده شود.

مدل اندازه‌گیری در معادلات ساختاری



مدل ساختاری در معادلات ساختاری



انواع مدل‌بندی معادلات ساختاری

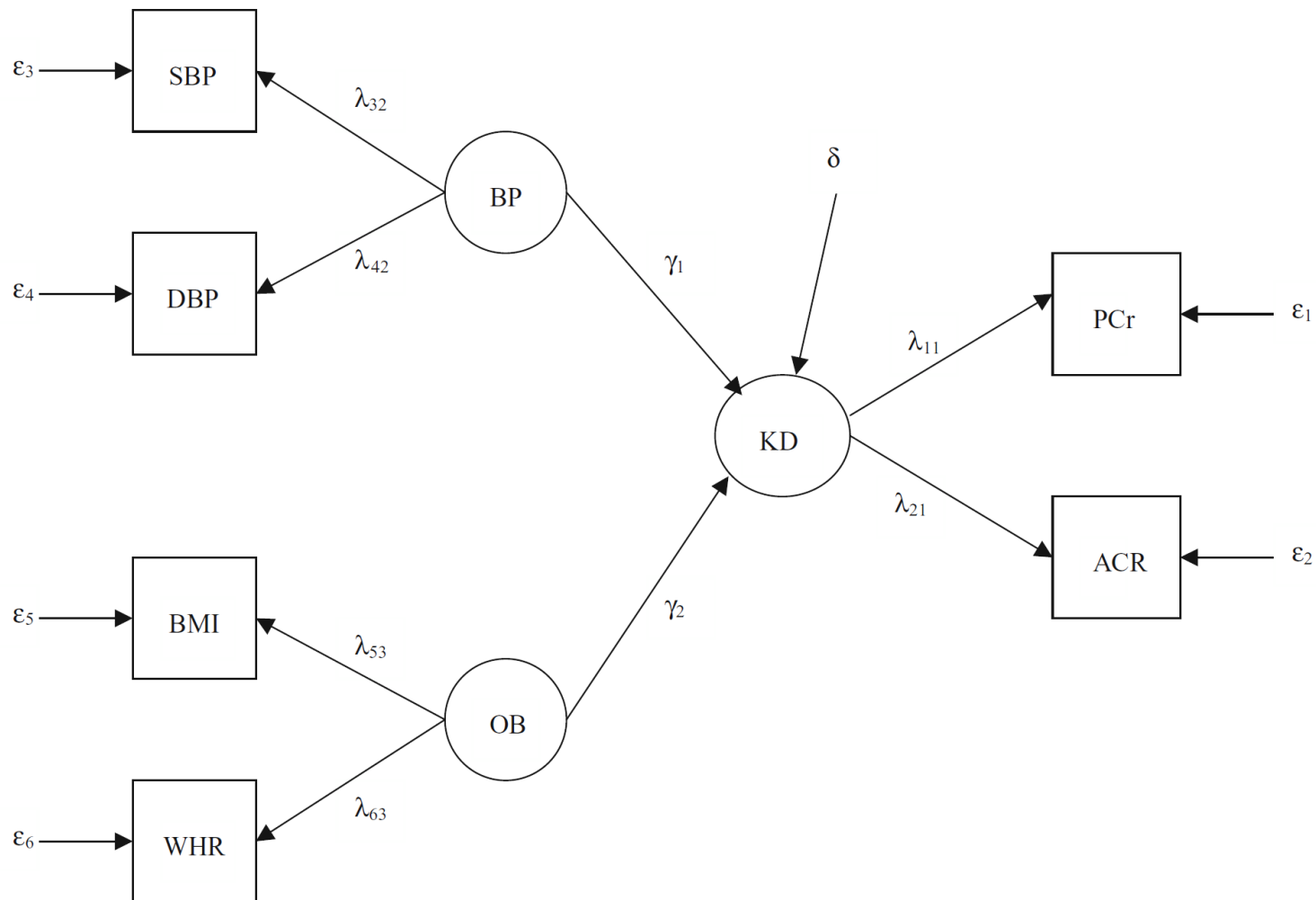
مدلیابی معادلات ساختاری خطی

رایج‌ترین نوع معادلات ساختاری است که فرض می‌کند که در مدل اندازه‌گیری، متغیرهای پنهان (فاکتورها) بصورت تابعی خطی از متغیرهای مشاهده‌شده است، تحت چنین شرایطی باید متغیرهای قابل مشاهده پیوسته، مستقل از هم و از توزیع نرمال پیروی کنند. نهایتاً تفسیر ضرایب مسیر همانند رگرسیون معمول، خطی است.

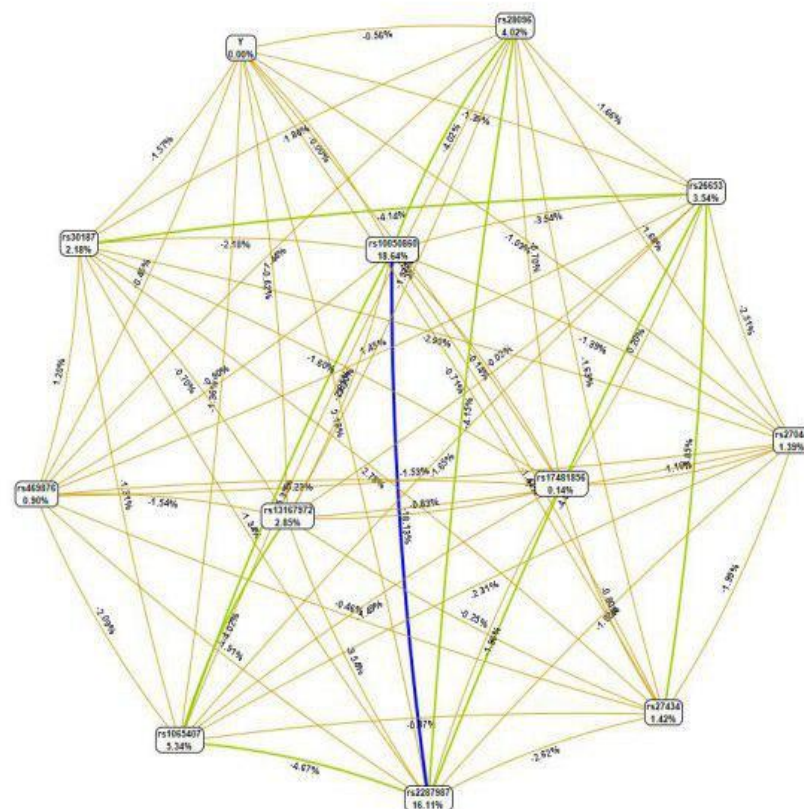
مدلیابی معادلات ساختاری غیرخطی

بر خلاف نوع قبلی، در این نوع از مدلیابی فرض می‌شود که در مدل اندازه‌گیری، متغیرهای پنهان بصورت تابعی غیرخطی از متغیرهای قابل مشاهده است و اساساً هدف ارزیابی روابط غیر خطی بین متغیرهای پنهان است. کاربرد این روش‌ها بیشتر در تحلیل‌های ژنتیکی است.

نمایش مدل مفهومی برای یک مدلیابی معادلات ساختاری خطی



A horizontal row of nine blue squares of varying shades, with the second square from the left being a lighter shade and outlined in white.

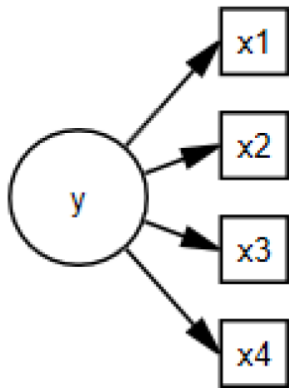


انواع مدل‌بندی معادلات ساختاری خطی

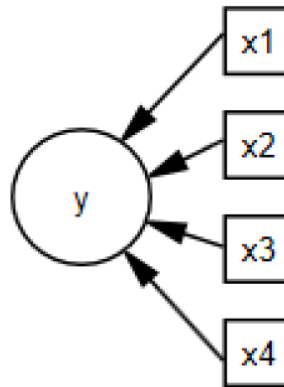
معادلات ساختاری بر حسب نوع تشکیل متغیرهای پنهان به سه نوع، ایجاد، انعکاسی و علیت چندگانه تقسیم می‌شوند:

1. در معادلات ساختاری ایجاد، متغیر پنهان با نقش وابسته و متغیرهای قابل مشاهده و اندازه‌گیری کننده آن در نقش مستقل هستند.
2. در معادلات ساختاری انعکاسی، متغیر پنهان با نقش مستقل و متغیرهای قابل مشاهده در نقش متغیرهای وابسته هستند.
3. در معادلات ساختاری علیت چندگانه، متغیر پنهان بطور همزمان نقش مستقل برای بعضی از متغیرهای قابل مشاهده و برای بعضی دیگر نقش وابسته را دارد.

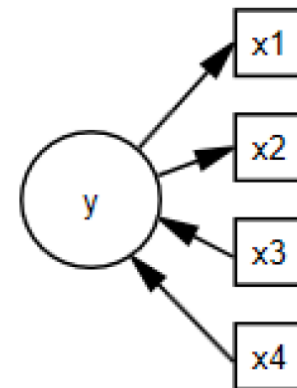
انواع مدل‌بندی معادلات ساختاری خطی



Reflective Model



Formative Model



Multiple Indicators Multiple Causes

انواع مدل‌بندی معادلات ساختاری خطی

- اگر در بین متغیرهای قابل مشاهده و اندازه‌گیری کننده متغیر پنهان همبستگی بالایی وجود داشته باشد می‌توان از هر دو مدل انعکاسی و ایجادی استفاده کرد اما با توجه با کارایی بیشتر مدل‌های انعکاسی در مقایسه با ایجادی توصیه می‌شود در چنین شرایطی از مدل‌های انعکاسی استفاده شود.
- در صورتی که در بین متغیرهای قابل مشاهده همبستگی پایینی وجود داشته باشد استفاده از مدل‌های ایجادی توصیه می‌گردد. اما در صورتی که بین بعضی از متغیرهای قابل مشاهده همبستگی بالا و در بعضی دیگر همبستگی پایینی وجود داشته باشد می‌توان از مدل علیت چندگانه استفاده کرد.
- معادلات ساختاری با مدل ایجادی را مدل واریانس و مدل‌های انعکاسی و علیت چندگانه را مدل کواریانس می‌شناسند و از ماتریس همبستگی داده‌ها استفاده می‌کند.
- زمانی که برای برازش معادلات ساختاری حجم نمونه کمی در اختیار باشد توصیه می‌شود از مدل واریانس یا ایجادی برای مدل‌بندی استفاده شود.

نقش متغیرها در مدلیابی معادلات ساختاری

- درونزا: متغیری است که از حداقل یک متغیر دیگر در مدل طراحی شده اثر می پذیرد در واقع متغیری است که حداقل یک فلش به آن وارد می شود/متغیر وابسته.
- برونزا: متغیر برونزا متغیری است که هیچ اثری از سایر متغیرهای مدل طراحی شده نمی پذیرد/متغیر مستقل رگرسیونی.
- میانجی: متغیری است که بطور همزمان برای بعضی متغیرها نقش وابسته و برای بعضی دیگر نقش مستقل را ایفا می کند.
- تعدیل گر/یا مداخله گر: متغیری است که جهت و شدت ارتباط بین دو متغیر دیگر اثر تعدیل گر دارد.

روش‌های برازش کلاسیک در معادلات ساختاری

در مدل‌بندی معادلات ساختاری بصورت پیش فرض از ماتریس واریانس-کواریانس متغیرها استفاده می‌شود. استفاده از این ماتریس آگاهی بخش تر از ماتریس همبستگی است و نشان داده شده است در این نوع از مدل‌بندی‌ها استفاده از ماتریس همبستگی باعث برآوردگرهای و انحراف معیار برآوردگرها نیز ناصحیح شده و باعث معنی‌داری کاذب پارامترها می‌شود.

برازش OLS به عنوان روش ابتدایی نیازی به پذیرهی توزیع احتمالاتی نداشته ولی دارای معایبی چون وابستگی به مقیاس متغیرهای مدل است. در حالی که برازش GLS و ML سازگار بوده و مستقل از مقیاس متغیرهای مدل هستند اما به حجم نمونه بیشتری به دلیل حساسیت به تخطی از پذیره نرمال چندمتغیره متغیرهای مشاهده شده است. نرم افزارهای چون لیزرل بصورت پیش فرض از برازش OLS استفاده می‌کند.

شاخص‌های نیکویی برازش در برازش کلاسیک معادلات ساختاری

Model Fit Criteria and Acceptable Fit Interpretation

<i>Model fit criterion</i>	<i>Acceptable level</i>	<i>Interpretation</i>
Chi-square	Tabled χ^2 value	Compares obtained χ^2 value with tabled value for given <i>df</i>
Goodness-of-fit (GFI)	0 (no fit) to 1 (perfect fit)	Value close to .95 reflects a good fit
Adjusted GFI (AGFI)	0 (no fit) to 1 (perfect fit)	Value adjusted for <i>df</i> , with .95 a good model fit
Root-mean-square residual (RMR)	Researcher defines level	Indicates the closeness of Σ to <i>S</i> matrix
Root-mean-square error of approximation (RMSEA)	<.05	Value less than .05 indicates a good model fit
Tucker–Lewis index	0 (no fit) to 1 (perfect fit)	Value close to .95 reflects a good model fit
Normed fit index	0 (no fit) to 1 (perfect fit)	Value close to .95 reflects a good model fit
Normed chi-square	1.0–5.0	Less than 1.0 is a poor model fit; more than 5.0 reflects a need for improvement
Parsimonious fit index	0 (no fit) to 1 (perfect fit)	Compares values in alternative models
Akaike information criterion	0 (perfect fit) to negative value (poor fit)	Compares values in alternative models



شاخص‌های نیکویی برازش کلاسیک در معادلات ساختاری

$$\text{GFI} = 1 - (\chi_{\text{model}}^2 / \chi_{\text{null}}^2)$$

$$\text{NFI} = (\chi_{\text{null}}^2 - \chi_{\text{model}}^2) / \chi_{\text{null}}^2$$

$$\text{RFI} = 1 - [(\chi_{\text{model}}^2 / df_{\text{model}}) / (\chi_{\text{null}}^2 / df_{\text{null}})]$$

$$\text{IFI} = (\chi_{\text{null}}^2 - \chi_{\text{model}}^2) / (\chi_{\text{null}}^2 - df_{\text{model}})$$

$$\text{TLI} = [(\chi_{\text{null}}^2 / df_{\text{null}}) - (\chi_{\text{model}}^2 / df_{\text{model}})] / [(\chi_{\text{null}}^2 / df_{\text{null}}) - 1]$$

$$\text{CFI} = 1 - [(\chi_{\text{model}}^2 - df_{\text{model}}) / (\chi_{\text{null}}^2 - df_{\text{null}})]$$

$$\text{Model AIC} = \chi_{\text{model}}^2 + 2q \text{ (number of free parameters)}$$

$$\text{Null AIC} = \chi_{\text{null}}^2 + 2q \text{ (number of free parameters)}$$

$$\text{RMSEA} = \sqrt{[\chi_{\text{M}}^2 - df_{\text{M}}] / [(N - 1)df_{\text{M}}]}.$$

انواع ضرایب در مدل‌بندی معادلات ساختاری

1. ضرایب مسیر مستقیم به اثر مستقیم یک متغیر مستقل بر یک متغیر پاسخ گفته می‌شود.
 2. ضرایب مسیر غیرمستقیم به اثر یک متغیر مستقل به واسطه متغیرهای واسطه دیگر بر روی یک متغیر پاسخ گفته می‌شود.
 3. ضرایب مسیر کل به مجموع اثرات مستقیم و غیرمستقیم یک متغیر مستقل بر روی یک متغیر غیرمستقیم اطلاق می‌شود.
- هر یک از ضرایب فوق می‌توانند بصورت استاندارد شده و مستقل از مقیاس متغیرها و یا غیراستاندارد شده و مبتنی بر مقیاس متغیرها باشند.

تفسیر ضرایب در مدل‌بندی معادلات ساختاری

تفسیر ضرایب استاندارد شده بصورت تغییر در انحراف معیار نسبت به میانگین هر متغیر است. مثلاً اگر ضریب مسیر مستقیم و استاندارد شده‌ی برابر 0.8 باشد یعنی تغییر به اندازه یک واحد انحراف معیار در متغیر مستقل بطور متوسط باعث 0.8 انحراف معیار تغییر در متغیر وابسته نسبت به میانگینشان می‌شود. در ضریب استاندارد شده بین هر متغیر مستقل و وابسته، انحراف معیارها واحد هستند. ضریب مسیر مستقیم استاندارد شده همان ضریب همبستگی بین دو متغیر مستقل و وابسته است.

ضرایب غیراستاندارد شده مسیر همانند ضرایب غیراستاندارد شده رگرسیون خطی تفسیر می‌شوند. مثلاً اگر ضریب مسیر مستقیم و غیر استاندارد شده‌ی برابر 0.8 باشد یعنی تغییر به اندازه یک واحد در متغیر مستقل بطور متوسط باعث 0.8 واحد تغییر در متغیر وابسته می‌شود.

آزمون ضرایب در مدل‌بندی معادلات ساختاری

- برای آزمون معنی‌داری ضرایب مسیر مستقیم و غیر مستقیم از آماره آزمون Wald و مبتنی بر T score تصمیم‌گیری می‌شود.
- در مواردی با حجم نمونه بزرگ (مثلاً درجه آزادی بیش از ۲۵) آزمون Z و آزمون T یکسان هستند.
- همچنین قادر خواهیم بود که غیر از p -value برای ضرایب مسیر، فاصله اطمینان جانبی آنها را نیز محاسبه کنیم.
- براساس معنی‌داری و اندازه T -score و پیشنهادات نرم افزار، می‌توان مدل را اصلاح و بعضی مسیرها حذف و بعضی مسیرهای جدید و معنی‌داری که تاثیر بسزایی روی شاخص‌های نیکویی برازش و شاخص R -square مدل دارد را به مدل مفهومی/دیاگرام اضافه نمود.

مدلیابی معادلات ساختاری با برازش بیزی

✓ در بیشتر مواقع محقق دانش قبلی در مورد خیلی از پارامترها و مسیرها دارد که می‌تواند در قالب توزیع‌های پیشین وارد برازش بیزی کند که این انعطاف برازش بیزی است. در مواردی که اطلاعات دقیقی از پارامترها در اختیار نیست می‌توان از توزیع پیشین کم و یا ناآگاهی بخش در برازش بیزی استفاده نمود.

✓ در برازش بیزی برای مدلیابی معادلات ساختاری، برآورد پارامترها مجهول مسیرها، مقایسه مدل‌ها و شاخص نیکویی برازش متفاوت از برازش کلاسیک است.

✓ در برازش بیزی برای معادلات ساختاری از توزیع‌های پیشین مزدوج (conjugate) برای پارامترها و از الگوریتم نمونه‌گیری گیبس با یک چرخه‌ی تکراری در روش زنجیره مارکف مونت کارلویی (MCMC) استفاده می‌شود.

توزیع‌های پیشین در معادلات ساختاری

✓ در بیشتر مواقع از توزیع‌های گامای معکوس برای ساختار واریانس، توزیع بتا برای پارامتر همبستگی بین متغیرهای پنهان استفاده می‌شود و به عنوان پیشین‌های مزدوج در بیشتر داده‌ها محسوب می‌شوند.

✓ برای بارهای عاملی در هر متغیر پنهان نیز از توزیع احتمال نرمال با میانگین صفر و واریانس بزرگ استفاده می‌شود.

✓ برای پارامتر ضرایب مسیر مجهول نیز از توزیع احتمال نرمال با میانگین صفر (تحت فرض صفر) و واریانس‌های بزرگ استفاده می‌شود.

مزایای برآزش بیزی در مقایسه با کلاسیک در معادلات ساختاری

- ✓ قابلیت استفاده از تجربیات پیشین محقق و نتایج مطالعات پیشین در مدل‌بندی معادلات ساختاری
- ✓ عدم نیاز به حل انتگرال‌های پیچیده حاصل از توزیع توأم در برآزش حداکثر درست‌نمایی
- ✓ قابلیت کاربرد در مواردی که حجم نمونه به نسبت برآزش کلاسیک کم است
- ✓ سازگاری برآوردگرهای بیزی
- ✓ عدم حساسیت به مشاهدات دورافتاده و تخطی از فرضیه نرمالیتی
- ✓

شاخص‌های مقایسه و ارزیابی نیکویی برازش بیزی در معادلات ساختاری

✓ رایج‌ترین و شاید مناسب‌ترین شاخص در مقایسه مدل‌ها با برازش بیزی بیز فاکتور است. این شاخص همانند لگاریتم درستنمایی در برازش کلاسیک دارای ویژگی‌های بسیار خوبی است اما محاسبه آن برخلاف لگاریتم درستنمایی مشکل و تنها از طریق زنجیره مارکف مونت کارلو امکان پذیر است.

✓ یکی دیگر از شاخص‌های مقایسه مدل در برازش بیزی شاخص DIC است.

شاخص‌های مقایسه و ارزیابی نیکویی برازش بیزی در معادلات ساختاری

✓ از شاخص **posterior predictive p-value** برای بررسی مطلق نیکویی برازش بیزی استفاده می‌شود. این شاخص کاملاً مشابه شاخص **hierarchical chi-square p-value** در برازش کلاسیک به عنوان شاخص مطلق برازش قلمداد می‌شود.

✓ از تحلیل باقیمانده‌ها نیز می‌توان در جهت بررسی مدل و ارزیابی حساسیت مدل نسبت به مشاهدات دورافتاده استفاده نمود.

شاخص Bayes factor

✓ این شاخص برای مقایسه دو مدل تحت دو فرضیه است. این شاخص برابر است با نسبتی از حاصلضرب توزیع پسین برای مدل تحت فرض صفر در توزیع پیشین برای مدل جایگزین

$$B_{10} = \frac{p(\mathbf{Y}|M_1)}{p(\mathbf{Y}|M_0)}$$

تفسير شاخص Bayes factor

B_{10}	$2 \log B_{10}$	Evidence against $H_0(M_0)$
<1	<0	Negative (supports $H_0(M_0)$)
1 to 3	0 to 2	Not worth more than a bare mention
3 to 20	2 to 6	Positive (supports $H_1(M_1)$)
20 to 150	6 to 10	Strong
>150	>10	Decisive

Reproduced by permission of *Psychometrika* from Lee, S. Y. and Song, X. Y. (2003a). Model comparison of nonlinear structural equation models with fixed covariates. *Psychometrika*, **68**, 27–47.

شاخص PP p-value

این شاخص یک احتمال است و هرگاه نزدیک 0.5 باشد حاکی از آن است آماره آزمون تقریباً نزدیک مرکز (میانه) توزیع پسین پیشگو تحت مدل برازش یافته است.
به عبارت دیگر هرگاه این شاخص نزدیک 0.5 باشد یعنی مدل برازش یافته تقریباً برازش قابل قبولی داشته است و انحراف از 0.5 برازش مطلق مدل فرضی را زیر سوال می برد.

برازش بیزی معادلات ساختاری با استفاده از نرم افزارهای رایج

✓ برای برازش بیزی مدلیابی معادلات ساختاری می توان از نرم افزار آماری WinBUGS استفاده نمود
(<http://www.mrc-bsu.cam.ac.uk/bugs/>)

✓ همچنین برای برازش بیزی در این نوع از مدلبندی ها می توان از بسته آماری blavaan در نرم افزار R استفاده کرد (<https://www.jstatsoft.org/article/view/v085i04>).

Bayesian SEM: R codes

```
#Bayesian structural equation models
library(MIIVsem)
library(blavaan)
data(bollen1989a)
head(bollen1989a)
#View(bollen1989a)

## The industrialization and Political Democracy Example
## Bollen (1989), page 332
model <- '
# latent variable definitions
  ind60 =~ x1 + x2 + x3
  dem60 =~ y1 + a*y2 + b*y3 + c*y4
  dem65 =~ y5 + a*y6 + b*y7 + c*y8

# regressions
  dem60 ~ ind60
  dem65 ~ ind60 + dem60

# residual correlations
  y1 ~~ y5
  y2 ~~ y4 + y6
  y3 ~~ y7
  y4 ~~ y8
  y6 ~~ y8

## unique priors for mv intercepts; parallel chains
fit <- bsem(model, data=PoliticalDemocracy,
            dp=dpriors(nu="normal(5,10)"))
summary(fit)
```

- Lee SY, Song XY. Basic and advanced Bayesian structural equation modeling: With applications in the medical and behavioral sciences. John Wiley & Sons; 2012 Jul 5.
- Palomo J, Dunson DB, Bollen K. Bayesian structural equation modeling. In: Handbook of latent variable and related models 2007 Jan 1 (pp. 163-188). North-Holland.
- Kumar S, Kumar S. Structure equation modeling basic assumptions and concepts: A novices guide. Asian Journal of Management Sciences. 2015;3(07):25-8.

