
بسم الله الرحمن الرحیم

مروری بر پیشنیازهای اساسی در مدل سازی آماری

الهام تبریزی

دانشکده علوم ریاضی و کامپیوتر

آذر ۱۴۰۰

فهرست مطالب

بخش یکم

شناساپذیری مدل و ضرورت آن

مدل آماری پارامتری

- در استنباط‌های آماری و نظریه‌ی تصمیم، یک مجموعه داده، تحقق یا مشاهده‌هایی از یک آزمایش تصادفی تعریف شده روی فضای احتمال (Ω, F, P) است.
- اندازه‌ی احتمال P، جامعه نامیده می‌شود.
- یک جامعه معلوم است اگر و تنها اگر $P(A)$ به ازای هر مجموعه‌ی $A \in F$، مقداری معلوم باشد. در مسایل آماری، جامعه‌ی P حداقل نیمه معلوم است و ما مایل به کشف ویژگی‌هایی از جامعه P بر پایه‌ی نمونه‌ی در دسترس هستیم.
- همواره برای ساده‌تر شدن انجام هرگونه تحلیلی، یک مدل آماری روی جامعه‌ی P اعمال می‌شود.
- هر مدل آماری پارامتری، مجموعه‌ای از فرض‌ها را روی جامعه‌ی P اعمال می‌کند و بیان می‌دارد که جامعه‌ی P به خانواده‌ای پارامتری مثل $\{P_{\boldsymbol{\theta}} \ : \ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ تعلق دارد.
- $Y \sim N(\mu, \sigma^۲)$، $\boldsymbol{\theta}' = (\mu, \sigma^۲)$، $\mathcal{P} = \{N(\mu, \sigma^۲) \mid (\mu, \sigma^۲) \in \boldsymbol{\Theta}\}$.

شناساپذیری

مثال ۱: شناساناپذیری

- $Y \sim N(\mu_1 + \mu_2, \sigma^2)$، $\theta' = (\mu_1, \mu_2, \sigma^2)$
- $Y \sim N(5, 1)$
- $\mu_1 + \mu_2 = 5$ و $\sigma^2 = 1$ $\quad \Longleftarrow \quad$ $(\mu_1, \mu_2, \sigma^2) = ?$

1. شناساپذیری، یک ویژگی در مور توزیع احتمال مشاهده‌ها و یا مدل است نه برآوردگر یا روش برآورد.
2. شناساپذیری، یک ویژگی لازم برای کفایت هر مدل آماری است و اگر مدلی شناسا نباشد، آنگاه هرگونه استنباط در مورد آن پیچیده خواهد شد.

مثال ۱: شناساناپذیری

- $Y \sim N(\mu_1 + \mu_2, \sigma^2)$، $\theta' = (\mu_1, \mu_2, \sigma^2)$
- $Y \sim N(5, 1)$
- $\mu_1 + \mu_2 = 5$ و $\sigma^2 = 1 \quad \Longleftarrow \quad (\mu_1, \mu_2, \sigma^2) = ?$

۱. شناساپذیری، یک ویژگی در مور توزیع احتمال مشاهده‌ها و یا مدل است نه برآوردگر یا روش برآورد.
۲. شناساپذیری، یک ویژگی لازم برای کفایت هر مدل آماری است و اگر مدلی شناسا نباشد، آنگاه هرگونه استنباط در مورد آن پیچیده خواهد شد.

بخش دوم

مطالعات شبیه سازی

❶ مدل ۱:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\beta + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_i + \varepsilon_i, \quad i = 1, \ldots, N,$$

که در آن $u_i \sim N(0, \sigma_u^2)$ و $\varepsilon_i \sim N_2\left(0, \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right)$.

❷ مدل ۲:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\beta + \begin{bmatrix} 1 & 1/5 \\ 1 & -1/5 \end{bmatrix} \mathbf{u}_i + \varepsilon_i, \quad i = 1, \ldots, N,$$

که در آن $\mathbf{u}_i \sim N_2\left(\mathbf{0}, \begin{bmatrix} \sigma_{u1}^2 & 0 \\ 0 & \sigma_{u2}^2 \end{bmatrix}\right)$ و $\varepsilon_i \sim N_2(\mathbf{0}, \sigma^2 \mathbf{I}_2)$.

❸ مدل ۳:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\beta + \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \mathbf{u}_i + \varepsilon_i, \quad i = 1, \ldots, N,$$

که در آن $\mathbf{u}_i \sim N_2(\mathbf{0}, \sigma_u^2 \mathbf{I}_2)$ و $\varepsilon_i \sim N_2(\mathbf{0}, \sigma^2 \mathbf{I}_2)$.

سه مدل شناسا ناپذیر

❶ مدل ۱:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} ۱ & ۰ \\ ۱ & ۱ \end{bmatrix}\beta + \begin{bmatrix} ۱ \\ ۱ \end{bmatrix} u_i + \varepsilon_i, \quad i = ۱, \ldots, N,$$

که در آن $u_i \sim N(۰, \sigma_u^۲)$ و $\varepsilon_i \sim N_۲\left(۰, \sigma^۲ \begin{bmatrix} ۱ & \rho \\ \rho & ۱ \end{bmatrix}\right)$.

❷ مدل ۲:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} ۱ & ۰ \\ ۱ & ۱ \end{bmatrix}\beta + \begin{bmatrix} ۱ & ۱/۵ \\ ۱ & -۱/۵ \end{bmatrix}\mathbf{u}_i + \varepsilon_i, \quad i = ۱, \ldots, N,$$

که در آن $\mathbf{u}_i \sim N_۲\left(\mathbf{0}, \begin{bmatrix} \sigma_{u۱}^۲ & ۰ \\ ۰ & \sigma_{u۲}^۲ \end{bmatrix}\right)$ و $\varepsilon_i \sim N_۲(\mathbf{0}, \sigma^۲\mathbf{I}_۲)$.

❸ مدل ۳:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} ۱ & ۰ \\ ۱ & ۱ \end{bmatrix}\beta + \begin{bmatrix} ۱ & ۱ \\ ۱ & -۱ \end{bmatrix}\mathbf{u}_i + \varepsilon_i, \quad i = ۱, \ldots, N,$$

که در آن $\mathbf{u}_i \sim N_۲(\mathbf{0}, \sigma_u^۲\mathbf{I}_۲)$ و $\varepsilon_i \sim N_۲(\mathbf{0}, \sigma^۲\mathbf{I}_۲)$.

سه مدل شناسا ناپذیر

❶ مدل ۱:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} ۱ & ۰ \\ ۱ & ۱ \end{bmatrix}\beta + \begin{bmatrix} ۱ \\ ۱ \end{bmatrix} u_i + \varepsilon_i, \quad i = ۱, \ldots, N,$$

که در آن $u_i \sim N(۰, \sigma_u^۲)$ و $\varepsilon_i \sim N_۲\left(۰, \sigma^۲ \begin{bmatrix} ۱ & \rho \\ \rho & ۱ \end{bmatrix}\right)$.

❷ مدل ۲:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} ۱ & ۰ \\ ۱ & ۱ \end{bmatrix}\beta + \begin{bmatrix} ۱ & ۱/۵ \\ ۱ & -۱/۵ \end{bmatrix}\mathbf{u}_i + \varepsilon_i, \quad i = ۱, \ldots, N,$$

که در آن $\mathbf{u}_i \sim N_۲\left(\mathbf{0}, \begin{bmatrix} \sigma_{u۱}^۲ & ۰ \\ ۰ & \sigma_{u۲}^۲ \end{bmatrix}\right)$ و $\varepsilon_i \sim N_۲(\mathbf{0}, \sigma^۲\mathbf{I}_۲)$.

❸ مدل ۳:

$$Y_i = X_i\beta + Z_i\mathbf{u}_i + \varepsilon_i = \begin{bmatrix} ۱ & ۰ \\ ۱ & ۱ \end{bmatrix}\beta + \begin{bmatrix} ۱ & ۱ \\ ۱ & -۱ \end{bmatrix}\mathbf{u}_i + \varepsilon_i, \quad i = ۱, \ldots, N,$$

که در آن $\mathbf{u}_i \sim N_۲(\mathbf{0}, \sigma_u^۲\mathbf{I}_۲)$ و $\varepsilon_i \sim N_۲(\mathbf{0}, \sigma^۲\mathbf{I}_۲)$.

مشکل احتمالی ۱:

همگرا نشدن الگوریتم عددی برای محاسبه‌ی برآورد پارامترهای مدل

مشکل احتمالی ۱:

همگرا نشدن الگوریتم عددی برای محاسبه‌ی برآورد پارامترهای مدل

Tabel : تعداد دفعات همگرا شدن الگوریتم عددی طی ۱۰۰۰ بار شبیه‌سازی در نرم‌افزار R (روش REML)

N	مدل ۱	مدل ۲	مدل ۳
۵۰۰	۱۰۰۰	۹۷۳	۱۰۰۰
۱۰۰۰	۱۰۰۰	۹۳۳	۹۹۹

مشکل احتمالی ۱:

همگرا نشدن الگوریتم عددی برای محاسبه‌ی برآورد پارامترهای مدل

Tabel : تعداد دفعات همگرا شدن الگوریتم عددی طی ۱۰۰۰ بار شبیه‌سازی در نرم‌افزار R (روش REML)

N	مدل ۱	مدل ۲	مدل ۳
۵۰۰	۱۰۰۰	۹۷۳	۱۰۰۰
۱۰۰۰	۱۰۰۰	۹۳۳	۹۹۹

Tabel : عداد دفعات همگرا شدن الگوریتم عددی طی ده بار شبیه‌سازی در نرم‌افزار SPSS (روش REML)

N	مدل ۱	مدل ۲	مدل ۳
۵۰۰	۱۰	۰	۰
۱۰۰۰	۱۰	۰	۰

مشکل احتمالی ۲:

یکسان نبودن خروجی نرم‌افزارهای متفاوت برای داده‌های مشابه

مشکل احتمالی ۲:

یکسان نبودن خروجی نرم‌افزارهای متفاوت برای داده‌های مشابه

Tabel : برآورد پارامترهای ماتریس کوواریانس مدل به روش ML، با استفاده از نرم‌افزارهای R و SPSS

مدل	پارامتر	N	SPSS	lme	LME
۱	$\sigma_u^2 = 1$	۵۰۰	۱.۰۶۵	۰.۹۹۵	۰.۹۹۸
		۱۰۰۰	۱.۰۱	۰.۹۸	۰.۹۵۳
	$\rho = 0.5$	۵۰۰	۰.۴۲۷	۰.۵۰۶	۰.۵۰۳
		۱۰۰۰	۰.۳۸۳	۰.۴۱	۰.۴۵۳
	$\sigma^2 = 0.5$	۵۰۰	۰.۴۲۹	۰.۴۹۷	۰.۴۹۴
		۱۰۰۰	۰.۴۴۳	۰.۴۶۳	۰.۵

Tabel : (ادامه).

مدل	پارامتر	N	SPSS	lme	LME
۲	$\sigma_{u1}^2 = 1$	۵۰۰	۰.۴۳۴	۰.۸۹۳	۰.۸۶۹
		۱۰۰۰	۱.۲۳۷	۱.۰۳۶	۰.۹۹
	$\sigma_{u2}^2 = 0.8$	۵۰۰	۰.۶۸۲	۰.۸۸۸	۰.۸۷۵
		۱۰۰۰	۰.۹۷۱	۰.۸۸۲	۰.۸۶۱
	$\sigma^2 = 0.5$	۵۰۰	۱.۳۲	۰.۳۹۳	۰.۴۵۱
		۱۰۰۰	۰.۰۱۴۵	۰.۴۱۵	۰.۵۰۸
۳	$\sigma_u^2 = 1$	۵۰۰	۰.۸۵۲	۱.۱۲	۱.۰۲۲
		۱۰۰۰	۰.۷۹۲	۱.۰۴۳	۰.۹۵۱
	$\sigma^2 = 0.5$	۵۰۰	۰.۸۵۲	۰.۳۱۵	۰.۵۱۱
		۱۰۰۰	۰.۷۹۲	۰.۲۹۲	۰.۴۷۵

Tabel : (ادامه).

مدل	پارامتر	N	SPSS	lme	LME
۲	$\sigma_{u1}^2 = 1$	۵۰۰	۰.۴۳۴	۰.۸۹۳	۰.۸۶۹
		۱۰۰۰	۱.۲۳۷	۱.۰۳۶	۰.۹۹
	$\sigma_{u2}^2 = 0.8$	۵۰۰	۰.۶۸۲	۰.۸۸۸	۰.۸۷۵
		۱۰۰۰	۰.۹۷۱	۰.۸۸۲	۰.۸۶۱
	$\sigma^2 = 0.5$	۵۰۰	۱.۳۲	۰.۳۹۳	۰.۴۵۱
		۱۰۰۰	۰.۰۱۴۵	۰.۴۱۵	۰.۵۰۸
۳	$\sigma_u^2 = 1$	۵۰۰	۰.۸۵۲	۱.۱۲	۱.۰۲۲
		۱۰۰۰	۰.۷۹۲	۱.۰۴۳	۰.۹۵۱
	$\sigma^2 = 0.5$	۵۰۰	۰.۸۵۲	۰.۳۱۵	۰.۵۱۱
		۱۰۰۰	۰.۷۹۲	۰.۲۹۲	۰.۴۷۵

Tabel : (ادامه).

LME	lme	SPSS	N	پارامتر	مدل
۰.۸۶۹	۰.۸۹۳	۰.۴۳۴	۵۰۰	$\sigma_{u1}^2 = 1$	۲
۰.۹۹	۱.۰۳۶	۱.۲۳۷	۱۰۰۰		
۰.۸۷۵	۰.۸۸۸	۰.۶۸۲	۵۰۰	$\sigma_{u2}^2 = 0.8$	
۰.۸۶۱	۰.۸۸۲	۰.۹۷۱	۱۰۰۰		
۰.۴۵۱	۰.۳۹۳	۱.۳۲	۵۰۰	$\sigma^2 = 0.5$	
۰.۵۰۸	۰.۴۱۵	۰.۰۱۴۵	۱۰۰۰		
۱.۰۲۲	۱.۱۲	۰.۸۵۲	۵۰۰	$\sigma_u^2 = 1$	۳
۰.۹۵۱	۱.۰۴۳	۰.۷۹۲	۱۰۰۰		
۰.۵۱۱	۰.۳۱۵	۰.۸۵۲	۵۰۰	$\sigma^2 = 0.5$	
۰.۴۷۵	۰.۲۹۲	۰.۷۹۲	۱۰۰۰		

: Tabel (ادامه).

مدل	پارامتر	N	SPSS	lme	LME
۲	$\sigma^2_{u1} = 1$	۵۰۰	۰.۴۳۴	۰.۸۹۳	۰.۸۶۹
		۱۰۰۰	۱.۲۳۷	۱.۰۳۶	۰.۹۹
	$\sigma^2_{u2} = 0.8$	۵۰۰	۰.۶۸۲	۰.۸۸۸	۰.۸۷۵
		۱۰۰۰	۰.۹۷۱	۰.۸۸۲	۰.۸۶۱
	$\sigma^2 = 0.5$	۵۰۰	۱.۳۲	۰.۳۹۳	۰.۴۵۱
		۱۰۰۰	۰.۰۱۴۵	۰.۴۱۵	۰.۵۰۸
۳	$\sigma^2_{u} = 1$	۵۰۰	۰.۸۵۲	۱.۱۲	۱.۰۲۲
		۱۰۰۰	۰.۷۹۲	۱.۰۴۳	۰.۹۵۱
	$\sigma^2 = 0.5$	۵۰۰	۰.۸۵۲	۰.۳۱۵	۰.۵۱۱
		۱۰۰۰	۰.۷۹۲	۰.۲۹۲	۰.۴۷۵

مشکل احتمالی ۳:

ناتوانی نرم‌افزار در ساختن فواصل اطمینان

مشکل احتمالی ۳:

ناتوانی نرم‌افزار در ساختن فواصل اطمینان

Tabel : تعداد دفعات موفق ساخته شدن فواصل اطمینان طی ۱۰۰۰ بار شبیه‌سازی در نرم‌افزار R

N	مدل ۱	مدل ۲	مدل ۳
۵۰۰	۴۵۲	۵۳۲	۴۴۷
۱۰۰۰	۴۷۳	۶۷۴	۹۶۸

مشکل احتمالی ۴:

گزارش کردن خطای استاندارد صفر و یا خیلی بزرگ برای پارامترهای ماتریس کوواریانس

مشکل احتمالی ۴:

گزارش کردن خطای استاندارد صفر و یا خیلی بزرگ برای پارامترهای ماتریس کوواریانس

Tabel : چندک‌های انحراف استاندارد برای پارامترهای ماتریس کوواریانس به روش ML در نرم‌افزار R طی ۱۰۰۰ بار شبیه‌سازی

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
	$\sigma_u^2 = 1$	۵۰۰	۲.۳	۲.۷	۳.۱	۵.۴	۶.۲	۹	۱۳.۴	۱۹.۳	۴۶.۶	۲۴۴۸.۶
		۱۰۰۰	۱.۸	۲	۲.۲	۳.۷	۴.۳	۶.۲	۸.۹	۱۱.۸	۲۱.۸	۱۳۴.۹
۱	$\rho = 0.5$	۵۰۰	۲.۵	۲.۹	۳.۲	۵.۴	۶.۲	۸.۷	۱۲.۹	۱۹.۴	۴۵.۸	۲۶۴۱
		۱۰۰۰	۱.۸	۲	۲.۳	۳.۸	۴.۳	۶.۲	۸.۸	۱۱.۷	۲۱.۴	۱۳۴.۵
	$\sigma^2 = 0.5$	۵۰۰	۲.۳	۲.۷	۳.۱	۵.۴	۶.۲	۹	۱۳.۴	۱۹.۳	۴۶.۶	۲۴۴۸.۸
		۱۰۰۰	۱.۸	۲	۲.۲	۳.۷	۴.۳	۶.۲	۸.۹	۱۱.۸	۲۱.۸	۱۳۴.۹

مشکل احتمالی ۴:

گزارش کردن خطای استاندارد صفر و یا خیلی بزرگ برای پارامترهای ماتریس کوواریانس

Tabel : چندک‌های انحراف استاندارد برای پارامترهای ماتریس کوواریانس به روش ML در نرم‌افزار R طی ۱۰۰۰ بار شبیه‌سازی

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۱	$\sigma_u^2 = 1$	۵۰۰	۲.۳	۲.۷	۳.۱	۵.۴	۶.۲	۹	۱۳.۴	۱۹.۳	۴۶.۶	۲۴۴۸.۶
		۱۰۰۰	۱.۸	۲	۲.۲	۳.۷	۴.۳	۶.۲	۸.۹	۱۱.۸	۲۱.۸	۱۳۴.۹
	$\rho = 0.5$	۵۰۰	۲.۵	۲.۹	۳.۲	۵.۴	۶.۲	۸.۷	۱۲.۹	۱۹.۴	۴۵.۸	۲۶۴۱
		۱۰۰۰	۱.۸	۲	۲.۳	۳.۸	۴.۳	۶.۲	۸.۸	۱۱.۷	۲۱.۴	۱۳۴.۵
	$\sigma^2 = 0.5$	۵۰۰	۲.۳	۲.۷	۳.۱	۵.۴	۶.۲	۹	۱۳.۴	۱۹.۳	۴۶.۶	۲۴۴۸.۸
		۱۰۰۰	۱.۸	۲	۲.۲	۳.۷	۴.۳	۶.۲	۸.۹	۱۱.۸	۲۱.۸	۱۳۴.۹

Tabel : (ادامه).

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۲	$\sigma^2_{u1} = 1$	۵۰۰	۱.۲	۱.۶	۱.۸	۳.۱	۳.۵	۴.۸	۶.۹	۹.۴	۲۳.۱	۱۴۷.۸
		۱۰۰۰	۳.۴	۲.۴	۲.۴	۲.۴	۲.۴	۲.۴	۲.۴	۲.۴	۲.۴	۲.۴
	$\sigma^2_{u2} = 0.8$	۵۰۰	۰.۵	۰.۷	۰.۸	۱.۳	۱.۵	۲.۱	۳	۴.۲	۱۰.۲	۶۵.۷
		۱۰۰۰	۱	۱	۱	۱	۱	۱	۱	۱	۱	۱
	$\sigma^2 = 0.5$	۵۰۰	۲.۶	۳.۲	۳.۶	۶.۲	۷.۱	۹.۷	۱۳.۹	۱۸.۹	۴۶.۳	۲۹۵.۸
		۱۰۰۰	۴.۸	۴.۸	۴.۸	۴.۸	۴.۸	۴.۸	۴.۸	۴.۸	۴.۸	۴.۸
۳	$\sigma^2_u = 1$	۵۰۰	۰	۰	۱.۹	۳.۳	۳.۶	۴.۹	۸.۷	۱۲.۳	۱۳.۳	۱۴
		۱۰۰۰	۰	۰	۲.۳	۲.۲	۲.۴	۳.۳	۴.۴	۶.۳	۹.۱	۹.۵
	$\sigma^2 = 0.5$	۵۰۰	۰	۰	۳.۸	۶.۶	۷.۳	۹.۸	۱۷.۵	۲۴.۷	۲۶.۷	۲۸.۱
		۱۰۰۰	۰	۰	۲.۶	۴.۵	۴.۸	۶.۶	۸.۸	۱۲.۶	۱۸.۳	۱۹

Tabel : چندک‌های انحراف استاندارد برای پارامترهای کوواریانس به روش ML در نرم‌افزار SPSS طی ۱۰ بار شبیه‌سازی

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۱	$\sigma_u^2 = 1$	۵۰۰	۱۲.۶	۱۲.۹	۱۴.۴	۲۸۴.۵	۲۸۹.۶	۳۰۳.۶	۳۱۴.۱	۳۱۴.۳	۳۱۴.۵	۳۱۴.۶
		۱۰۰۰	۹.۷	۹.۷	۱۰	۲۶.۴	۳۶	۴۴.۲	۸۸.۵	۱۴۷	۱۹۳.۸	۲۰۵.۵
	$\rho = 0.5$	۵۰۰	۱۶.۴	۱۶.۸	۱۸.۵	۳۹۱.۶	۳۹۲.۴	۳۹۳.۴	۳۹۵.۴	۳۹۵.۸	۳۹۶.۲	۳۹۶.۳
		۱۰۰۰	۱۳.۴	۱۳.۴	۱۳.۴	۳۶	۴۹.۱	۵۹.۸	۱۱۵.۷	۱۹۷.۱	۲۶۲.۲	۲۷۸.۵
	$\sigma^2 = 0.5$	۵۰۰	۱۲.۶	۱۲.۹	۱۴.۴	۲۸۴.۲	۲۸۹.۲	۳۰۲.۶	۳۱۴.۱	۳۱۴.۳	۳۱۴.۵	۳۱۴.۶
		۱۰۰۰	۹.۷	۹.۷	۱۰	۲۶.۴	۳۶	۴۴.۲	۸۸.۵	۱۴۷	۱۹۳.۸	۲۰۵.۵
۲	$\sigma_{u1}^2 = 1$	۵۰۰	۰.۰۷	۰.۰۷	۰.۰۷	۰.۱۲	۰.۱۴	۰.۱۵	۰.۱۵	۰.۱۵	۰.۱۵	۰.۱۵
		۱۰۰۰	۰.۰۱	۰.۰۱	۰.۰۲	۰.۰۵	۰.۰۶	۰.۰۱	۲.۴	۰.۰۱	۰.۰۱	۰.۱۸
	$\sigma_{u2}^2 = 0.8$	۵۰۰	۰	۰	۰	۰	۰	۰.۰۵	۰.۰۵۱	۰.۰۵۵	۰.۰۵۹	۰.۰۶
		۱۰۰۰	۰	۰	۰	۰.۰۳	۰.۰۳	۰.۰۴	۰.۰۴	۰.۰۴	۰.۰۴	۰.۰۴
	$\sigma^2 = 0.5$	۵۰۰	۰	۰	۰	۰.۲۴	۰.۲۵	۰.۲۵	۰.۲۶	۰.۲۶	۰.۲۶	۰.۲۷
		۱۰۰۰	۰	۰	۰	۰	۰	۰.۱۷	۰.۱۷	۰.۱۸	۰.۱۸	۰.۱۸

Tabel : چندک‌های انحراف استاندارد برای پارامترهای کوواریانس به روش ML در نرم‌افزار SPSS طی ۱۰ بار شبیه‌سازی

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۱	$\sigma_u^2 = 1$	۵۰۰	۱۲.۶	۱۲.۹	۱۴.۴	۲۸۴.۵	۲۸۹.۶	۳۰۳.۶	۳۱۴.۱	۳۱۴.۳	۳۱۴.۵	۳۱۴.۶
		۱۰۰۰	۹.۷	۹.۷	۱۰	۲۶.۴	۳۶	۴۴.۲	۸۸.۵	۱۴۷	۱۹۳.۸	۲۰۵.۵
	$\rho = 0.5$	۵۰۰	۱۶.۴	۱۶.۸	۱۸.۵	۳۹۱.۶	۳۹۲.۴	۳۹۳.۴	۳۹۵.۴	۳۹۵.۸	۳۹۶.۲	۳۹۶.۳
		۱۰۰۰	۱۳.۴	۱۳.۴	۱۳.۴	۳۶	۴۹.۱	۵۹.۸	۱۱۵.۷	۱۹۷.۱	۲۶۲.۲	۲۷۸.۵
	$\sigma^2 = 0.5$	۵۰۰	۱۲.۶	۱۲.۹	۱۴.۴	۲۸۴.۲	۲۸۹.۲	۳۰۲.۶	۳۱۴.۱	۳۱۴.۳	۳۱۴.۵	۳۱۴.۶
		۱۰۰۰	۹.۷	۹.۷	۱۰	۲۶.۴	۳۶	۴۴.۲	۸۸.۵	۱۴۷	۱۹۳.۸	۲۰۵.۵
۲	$\sigma_{u1}^2 = 1$	۵۰۰	۰.۰۷	۰.۰۷	۰.۰۷	۰.۱۲	۰.۱۴	۰.۱۵	۰.۱۵	۰.۱۵	۰.۱۵	۰.۱۵
		۱۰۰۰	۰.۰۱	۰.۰۱	۰.۰۲	۰.۰۵	۰.۰۶	۰.۰۱	۲.۴	۰.۰۱	۰.۰۱	۰.۱۸
	$\sigma_{u2}^2 = 0.8$	۵۰۰	۰	۰	۰	۰	۰	۰.۰۵	۰.۰۵۱	۰.۰۵۵	۰.۰۵۹	۰.۰۶
		۱۰۰۰	۰	۰	۰	۰.۰۳	۰.۰۳	۰.۰۴	۰.۰۴	۰.۰۴	۰.۰۴	۰.۰۴
	$\sigma^2 = 0.5$	۵۰۰	۰	۰	۰	۰.۲۴	۰.۲۵	۰.۲۵	۰.۲۶	۰.۲۶	۰.۲۶	۰.۲۷
		۱۰۰۰	۰	۰	۰	۰	۰	۰.۱۷	۰.۱۷	۰.۱۸	۰.۱۸	۰.۱۸

Tabel : (ادامه).

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۳	$\sigma_u^2 = 1$	۵۰۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
		۱۰۰۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
	$\sigma^2 = 0.5$	۵۰۰	۰.۰۱	۰.۰۱	۰.۰۱	۰.۰۱۱	۰.۰۱۱	۰.۰۳۱	۰.۱۱۴	۰.۱۱۴	۰.۱۱۴	۰.۱۱۵
		۱۰۰۰	۰.۰۷۴	۰.۰۷۴	۰.۰۷۴	۰.۰۷۷	۰.۰۷۹	۰.۰۸۲	۰.۰۸۲	۰.۰۸۲	۰.۰۸۲	۰.۰۸۳

Tabel : (ادامه).

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۳	$\sigma_u^2 = 1$	۵۰۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
		۱۰۰۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
	$\sigma^2 = 0.5$	۵۰۰	۰.۰۱	۰.۰۱	۰.۰۱	۰.۰۱۱	۰.۰۱۱	۰.۰۳۱	۰.۱۱۴	۰.۱۱۴	۰.۱۱۴	۰.۱۱۵
		۱۰۰۰	۰.۰۷۴	۰.۰۷۴	۰.۰۷۴	۰.۰۷۷	۰.۰۷۹	۰.۰۸۲	۰.۰۸۲	۰.۰۸۲	۰.۰۸۲	۰.۰۸۳

مشکل احتمالی ۵.

گزارش کردن فاصله اطمینان‌های نامعقول برای پارامترهای کوواریانس

مشکل احتمالی ۵.

گزارش کردن فاصله اطمینان‌های نامعقول برای پارامترهای کوواریانس

جدول: چندک‌های طول ۹۵ درصد پارامترهای ماتریس کوواریانس به روش REMLدر نرم‌افزار R طی ۱۰۰۰ بار شبیه‌سازی

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۱	$\sigma_u = ۱$	۵۰۰	۰٫۳	۰٫۳	۰٫۴	۰٫۷	۰٫۷	۱٫۲	۱٫۹	۲٫۷	۸٫۹	۲×۱۰^{۴}
		۱۰۰۰	۰٫۱	۰٫۲	۰٫۲	۰٫۴	۰٫۴	۰٫۷	۱	۱٫۶	۶	$۱٫۸ \times ۱۰^{۳}$
	$\rho = ۰٫۵$	۵۰۰	۰٫۶	۰٫۷	۰٫۸	۱٫۳	۱٫۴	۱٫۷	۱٫۹	۱٫۹	۱٫۹	۲×۱۰^{۴}
		۱۰۰۰	۰٫۵	۰٫۵	۰٫۶	۱	۱٫۲	۱٫۵	۱٫۸	۱٫۹	۱٫۹	۲
	$\sigma = \sqrt{۰٫۵}$	۵۰۰	۰٫۴	۰٫۵	۰٫۶	۱٫۱	۱٫۲	۲	۳٫۷	۷	۵۶٫۳	$۵٫۱ \times ۱۰^{۹}$
		۱۰۰۰	۰٫۳	۰٫۳	۰٫۳	۰٫۶	۰٫۷	۱٫۲	۲٫۲	۴	۶۴٫۹	$۵٫۴ \times ۱۰^{۳۸}$
۲	$\sigma_{u۱} = ۱$	۵۰۰	۰٫۱	۰٫۱	۰٫۲	۰٫۳	۰٫۴	۰٫۵	۰٫۸	۱٫۲	۳٫۴	$۲٫۵ \times ۱۰^{۳}$
		۱۰۰۰	۰٫۰۸	۰٫۱	۰٫۱	۰٫۱	۰٫۱	۰٫۲	۰٫۳	۰٫۵	۱٫۲	۴٫۴
	$\sigma_{u۲} = \sqrt{۰٫۸}$	۵۰۰	۰٫۱	۰٫۱	۰٫۱	۰٫۲	۰٫۳	۰٫۰۴	۰٫۵۱	۰٫۸	۱٫۲	$۱٫۷ \times ۱۰^{۲}$
		۱۰۰۰	۰٫۰۸	۰٫۰۹	۰٫۰۹	۰٫۱	۰٫۱	۰٫۱	۰٫۲	۰٫۲	۰٫۷	۲
	$\sigma = \sqrt{۰٫۵}$	۵۰۰	۰٫۶	۰٫۷	۰٫۸	۱٫۴	۱٫۶	۲٫۷	۵	۱۱٫۹	۴۰۳٫۸	$۹٫۴ \times ۱۰^{۱۸}$
		۱۰۰۰	۰٫۳	۰٫۳	۰٫۳	۰٫۵	۰٫۶	۰٫۹	۱٫۳	۲٫۱	۱۶٫۷	۴۸۴۱۰۱

مشکل احتمالی ۵.

گزارش کردن فاصله اطمینان‌های نامعقول برای پارامترهای کوواریانس

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۱	$\sigma_u = 1$	۵۰۰	۰٫۳	۰٫۳	۰٫۴	۰٫۷	۰٫۷	۱٫۲	۱٫۹	۲٫۷	۸٫۹	۲ × ۱۰^۴
		۱۰۰۰	۰٫۱	۰٫۲	۰٫۲	۰٫۴	۰٫۴	۰٫۷	۱	۱٫۶	۶	۱٫۸ × ۱۰^۳
	$\rho = 0.5$	۵۰۰	۰٫۶	۰٫۷	۰٫۸	۱٫۳	۱٫۴	۱٫۷	۱٫۹	۱٫۹	۱٫۹	۲ × ۱۰^۴
		۱۰۰۰	۰٫۵	۰٫۵	۰٫۶	۱	۱٫۲	۱٫۵	۱٫۸	۱٫۹	۱٫۹	۲
	$\sigma = \sqrt{0.5}$	۵۰۰	۰٫۴	۰٫۵	۰٫۶	۱٫۱	۱٫۲	۲	۳٫۷	۷	۵۶٫۳	۵٫۱ × ۱۰^۹
		۱۰۰۰	۰٫۳	۰٫۳	۰٫۳	۰٫۶	۰٫۷	۱٫۲	۲٫۲	۴	۶۴٫۹	۵٫۴ × ۱۰^۳۸
۲	$\sigma_{u1} = 1$	۵۰۰	۰٫۱	۰٫۱	۰٫۲	۰٫۳	۰٫۴	۰٫۵	۰٫۸	۱٫۲	۳٫۴	۲٫۵ × ۱۰^۳
		۱۰۰۰	۰٫۰۸	۰٫۱	۰٫۱	۰٫۱	۰٫۱	۰٫۲	۰٫۳	۰٫۵	۱٫۲	۴٫۴
	$\sigma_{u2} = \sqrt{0.8}$	۵۰۰	۰٫۱	۰٫۱	۰٫۱	۰٫۲	۰٫۳	۰٫۰۴	۰٫۵۱	۰٫۸	۱٫۲	۱٫۷ × ۱۰^۲
		۱۰۰۰	۰٫۰۸	۰٫۰۹	۰٫۰۹	۰٫۱	۰٫۱	۰٫۱	۰٫۲	۰٫۲	۰٫۷	۲
	$\sigma = \sqrt{0.5}$	۵۰۰	۰٫۶	۰٫۷	۰٫۸	۱٫۴	۱٫۶	۲٫۷	۵	۱۱٫۹	۴۰۳٫۸	۹٫۴ × ۱۰^۱۸
		۱۰۰۰	۰٫۳	۰٫۳	۰٫۳	۰٫۵	۰٫۶	۰٫۹	۱٫۳	۲٫۱	۱۶٫۷	۴۸۴۱۰۱

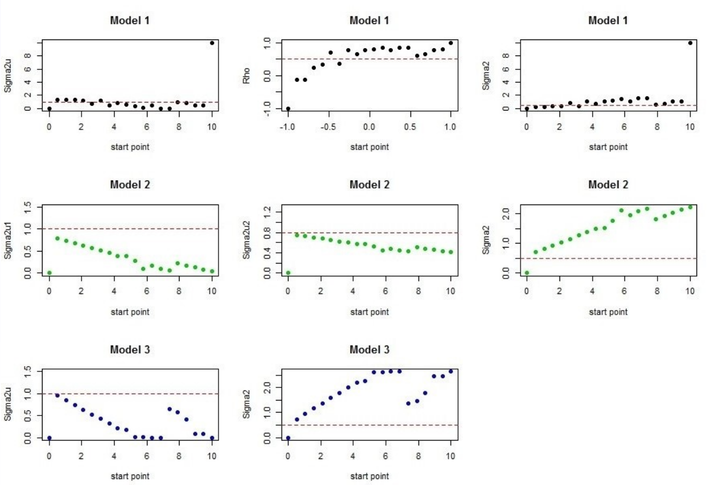
Tabel : (ادامه).

مدل	پارامتر	N	۰٪	۱٪	۵٪	۵۰٪	۶۰٪	۸۰٪	۹۰٪	۹۵٪	۹۹٪	۱۰۰٪
۳	$\sigma_u =$ ۱	۵۰۰	۰٫۲	۰٫۲	۰٫۲	۰٫۴	۰٫۵	۰٫۷	۱٫۱	۱٫۶	۴٫۶	۱٫۳ × ۱۰^۴
		۱۰۰۰	۰٫۰۹	۰٫۱	۰٫۱	۰٫۱	۰٫۱	۰٫۲	۰٫۲	۰٫۳	۰٫۵	۳٫۴
	$\sigma = \sqrt{۰٫۵}$	۵۰۰	۰٫۹	۱	۱٫۲	۲٫۶	۳٫۲	۶٫۷	۲۱٫۲	۹۹٫۹	۲۸۱۶۵٫۳	۶٫۱ × ۱۰^۲۸
		۱۰۰۰	۰٫۴	۰٫۴	۰٫۴	۰٫۵	۰٫۶	۰٫۸	۱٫۱	۱٫۵	۳٫۹	۵۴۷۲۰۲

مشکل احتمالی ۶.

تغییر برآورد پارامترهای شناساناپذیر با عوض کردن مقادیر اولیّه

تغییر برآورد پارامترهای شناساناپذیر با عوض کردن مقادیر اولیّه



تعریف شناساپذیری

تعریف ۱

شناساپذیری بر اساس خانواده‌ای از توزیع‌های آماری: مدل آماری (۱)، تعریف شده به واسطه‌ی خانواده‌ای از توزیع‌ها بر پایه‌ی بردار تصادفی $\mathbf{Y}$ که تحت بردار پارامتری $\boldsymbol{\theta}$، پارامتری شده است را در نظر بگیرید:

$$P = \{P_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \boldsymbol{\Theta}\} \tag{۱}$$

که در آن $P_{\boldsymbol{\theta}}$ توزیع بردار تصادفی $\mathbf{Y}$ و پارامتری شده به‌واسطه‌ی بردار پارامتری $\boldsymbol{\theta}$ و $\boldsymbol{\Theta}$ اشاره به فضای پارامتر دارد. این مدل روی فضای پارامتری $\boldsymbol{\Theta}$، شناساپذیر است

اگر به ازای تمام مقادیر y موجود در تکیه‌گاه $P^y_{\boldsymbol{\theta}} = P^y_{\boldsymbol{\theta}^*}$، آنگاه $\boldsymbol{\theta} = \boldsymbol{\theta}^*$.

تعریف شناساپذیری

تعریف ۱

شناساپذیری بر اساس خانواده‌ای از توزیع‌های آماری: مدل آماری (۱)، تعریف شده به واسطه‌ی خانواده‌ای از توزیع‌ها بر پایه‌ی بردار تصادفی $\mathbf{Y}$ که تحت بردار پارامتری $\boldsymbol{\theta}$، پارامتری شده است را در نظر بگیرید:

$$P = \{P_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \boldsymbol{\Theta}\} \qquad (۱)$$

که در آن $P_{\boldsymbol{\theta}}$ توزیع بردار تصادفی $\mathbf{Y}$ و پارامتری شده به‌واسطه‌ی بردار پارامتری $\boldsymbol{\theta}$ و $\boldsymbol{\Theta}$ اشاره به فضای پارامتر دارد. این مدل روی فضای پارامتری $\boldsymbol{\Theta}$، شناساپذیر است

اگر به ازای تمام مقادیر y موجود در تکیه‌گاه $P^y_{\boldsymbol{\theta}} = P^y_{\boldsymbol{\theta}^*}$، آنگاه $\boldsymbol{\theta} = \boldsymbol{\theta}^*$.

$$\boldsymbol{\theta} \to P^{\boldsymbol{\theta}}$$

For all $y \in S_y$???

۱ :$Y \sim INFBE(\alpha, \lambda, \mu, \phi)$

$f^{INFBE}(y; \alpha, \lambda, \mu, \phi) = f^{INFBE}(y; \alpha^*, \lambda^*, \mu^*, \phi^*)$ for all $y \in [۰, ۱]$:

$$\begin{cases} \alpha(۱ - \lambda), & y = ۰ \\ \alpha\lambda, & y = ۱ \\ (۱ - \alpha) f^{Beta}(y; \mu, \phi), & y \in (۰, ۱) \end{cases} = \begin{cases} \alpha^*(۱ - \lambda^*), & y = ۰ \\ \alpha^*\lambda^*, & y = ۱ \\ (۱ - \alpha^*) f^{Beta}(y; \mu^*, \phi^*), & y \in (۰, ۱) \end{cases}$$

۲ :$i = ۱, ..., I$ ،$Y_i \overset{iid}{\sim} INFBE(\alpha, \lambda, \mu, \phi)$

$$\prod_{i=۱}^{I} f^{INFBE}(y_i; \alpha, \lambda, \mu, \phi) = \prod_{i=۱}^{I} f^{INFBE}(y_i; \alpha^*, \lambda^*, \mu^*, \phi^*) \text{ for all } y_i \in [۰, ۱].$$

For all $y \in \mathcal{S}_y$???

۱. $Y \sim INFBE(\alpha, \lambda, \mu, \phi)$:

$f^{INFBE}(y; \alpha, \lambda, \mu, \phi) = f^{INFBE}(y; \alpha^*, \lambda^*, \mu^*, \phi^*)$ for all $y \in [۰, ۱]$:

$$\begin{cases} \alpha(۱ - \lambda), & y = ۰ \\ \alpha\lambda, & y = ۱ \\ (۱ - \alpha) f^{Beta}(y; \mu, \phi), & y \in (۰, ۱) \end{cases} = \begin{cases} \alpha^*(۱ - \lambda^*), & y = ۰ \\ \alpha^*\lambda^*, & y = ۱ \\ (۱ - \alpha^*) f^{Beta}(y; \mu^*, \phi^*), & y \in (۰, ۱) \end{cases}$$

۲. $Y_i \overset{iid}{\sim} INFBE(\alpha, \lambda, \mu, \phi)$، $i = ۱, ..., I$:

$\prod_{i=۱}^{I} f^{INFBE}(y_i; \alpha, \lambda, \mu, \phi) = \prod_{i=۱}^{I} f^{INFBE}(y_i; \alpha^*, \lambda^*, \mu^*, \phi^*)$ for all $y_i \in [۰, ۱]$.

مدل رگرسیون خطی چندگانه با متغیرهای مستقل ثابت

- $Y_i \mid \mathbf{x}_i \overset{id}{\sim} N(\mathbf{x}_i'\boldsymbol{\beta}, \sigma^2)$ به ازای $i = 1, \ldots, I$:

$$\prod_{i=1}^{I} f(y_i \mid \mathbf{x}_i; \boldsymbol{\beta}, \sigma) = \prod_{i=1}^{I} f(y_i \mid \mathbf{x}_i; \boldsymbol{\beta}^*, \sigma^*) \quad \text{for all } y_i \in \mathcal{S}_y$$

$$\frac{1}{\sigma^I} exp\{-(\mathbf{y}-\mathbf{X}\beta)'(\mathbf{y}-\mathbf{X}\beta)/2\sigma^2\} = \frac{1}{\sigma^{*I}} exp\{-(\mathbf{y}-\mathbf{X}\beta^*)'(\mathbf{y}-\mathbf{X}\beta^*)/2\sigma^{*2}\},$$

تساوی بالا برقرار است اگر و تنها اگر $\mathbf{X}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta}^*$ و $\sigma = \sigma^*$.

- $\mathbf{X}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta}^*$ دلالت بر $\boldsymbol{\beta} = \boldsymbol{\beta}^* \iff \mathbf{X}$ ماتریس طرحی پررتبه باشد.

ساده‌سازی تعریف شناساپذیری

مدل عام

$$Y = f(\beta) + \varepsilon \qquad (۲)$$

ساده‌سازی

تعریف ۲

شناساپذیری در حالت کلی مدل‌های عام: بردار پارامتری β شناساپذیر است اگر برای هر بردار $\beta_۱$ ، $\beta_۲$ و تابع برداری دلخواه $f(\cdot)$، $f(\beta_۱) = f(\beta_۲)$ دلالت بر $\beta_۱ = \beta_۲$ بکند. اگر β شناساپذیر باشد، آن‌گاه می‌توانیم بگوییم پارامتریدن $f(\beta)$ (مدل) شناساپذیر است.

ساده‌سازی تعریف شناساپذیری

مدل عام

$$Y = f(\beta) + \varepsilon \qquad (۲)$$

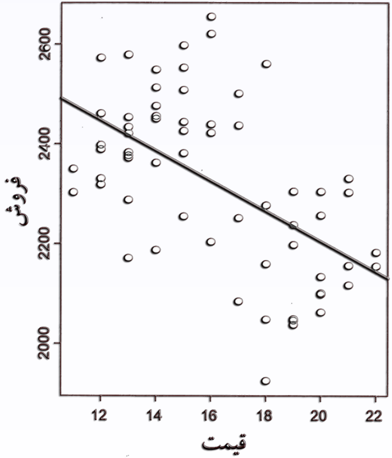
ساده‌سازی

تعریف ۲

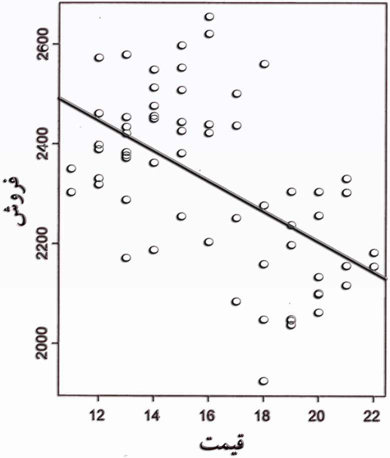
شناساپذیری در حالت کلی مدل‌های عام: بردار پارامتری β شناساپذیر است اگر برای هر بردار $\beta_۱$ ، $\beta_۲$ و تابع برداری دلخواه $f(\cdot)$، $f(\beta_۱) = f(\beta_۲)$ دلالت بر $\beta_۱ = \beta_۲$ بکند. اگر β شناساپذیر باشد، آنگاه می‌توانیم بگوییم پارامتریدن $f(\beta)$ (مدل) شناساپذیر است.

بخش سوم

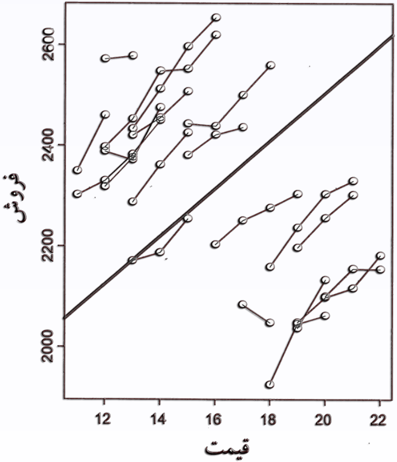
چند قضیه در شناساپذیری مدل‌های آمیخته



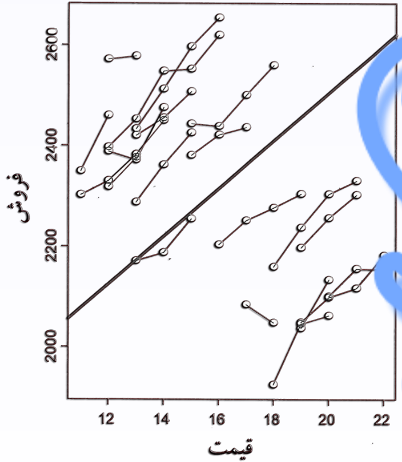
شکل ۱: نمودار پراکنش قیمت محصول و تعداد فروش آن.



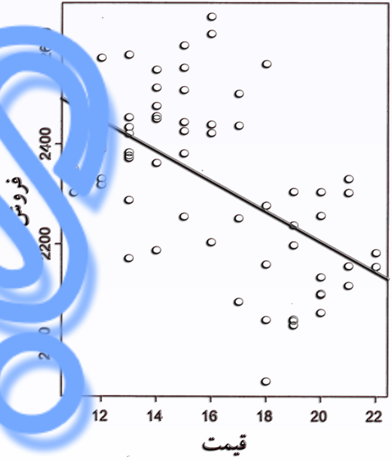
شکل ۱: نمودار پراکنش قیمت محصول و تعداد فروش آن.



شکل ۲: نمودار پراکنش قیمت و تعداد فروش محصول که داده‌های مربوط به هر فروشگاه به هم وصل شده‌اند.



شکل ۲: نمودار پراکنش قیمت و تعداد فروش محصول که داده‌های مربوط به هر فروشگاه به هم وصل شده‌اند.



شکل ۱: نمودار پراکنش قیمت محصول و تعداد فروش آن.

مدل آمیخته خطی

$$\mathbf{Y}_i = \mathbf{X}_i\gamma + \mathbf{W}_i\mathbf{b}_i + \varepsilon_i, \qquad i = ۱, \ldots, I.$$

مدل آمیخته خطی

$$\mathbf{Y}_i = \mathbf{X}_i\gamma + \mathbf{W}_i\mathbf{b}_i + \varepsilon_i, \qquad i = ۱, \ldots, I.$$

مدل خطی تعمیم‌یافته

$$h(E(Y_i)) = \mathbf{x}_i'\gamma, \qquad i = ۱, \ldots, I.$$

مدل آمیخته خطی

$$\mathbf{Y}_i = \mathbf{X}_i\gamma + \mathbf{W}_i\mathbf{b}_i + \varepsilon_i, \qquad i = ۱, \ldots, I.$$

مدل خطی تعمیم‌یافته

$$h(E(Y_i)) = \mathbf{x}_i'\gamma, \qquad i = ۱, \ldots, I.$$

مدل آمیخته‌ی خطی تعمیم‌یافته

$$\mathbf{h}(E(\mathbf{Y}_i \mid \mathbf{b}_i)) = \mathbf{X}_i\gamma + \mathbf{W}_i\mathbf{b}_i, \qquad i = ۱, \ldots, I.$$

قضیه ۳

رابطه‌ی شناساپذیری پارامترهای مدل تکی و توأم: *مدل توأم (۳) شناساپذیر است اگر و تنها اگر حداقل یکی از مدل‌های تکی آن شناساپذیر باشد.*

قضیه ۴

به ازای هر بردار پارامتری $(\theta_{\mathbf{u}}, \theta_{\varepsilon})$ ، بردار پارامتری $(\theta_{\mathbf{u}}, \theta_{\varepsilon}) \neq (\theta^*_{\mathbf{u}}, \theta^*_{\varepsilon})$ چنان وجود دارد که $Z[\Sigma^*_{\mathbf{u}} - \Sigma_{\mathbf{u}}]Z' = \Sigma^*_{\varepsilon} - \Sigma^*_{\varepsilon}$ برقرار است اگر و تنها اگر به ازای هر بردار پارامتری θ_{ε}، $\theta^*_{\varepsilon} \neq \theta_{\varepsilon}$ چنان وجود داشته باشد که $Z'\Sigma^*_{\varepsilon}Z \neq Z'\Sigma_{\varepsilon}Z$،
$H_Z[\Sigma_{\varepsilon} - \Sigma^*_{\varepsilon}] = \Sigma_{\varepsilon} - \Sigma^*_{\varepsilon}$ و

$$\Sigma^*_{\mathbf{u}} = \Sigma_{\mathbf{u}} + (Z'Z)^{-1}Z'[\Sigma_{\varepsilon} - \Sigma^*_{\varepsilon}]Z(Z'Z)^{-1}, \quad (۳)$$

که در آن $H_Z = Z(Z'Z)^{-1}Z'$.

قضیه ۳

رابطه‌ی شناساپذیری پارامترهای مدل تکی و توأم: *مدل توأم (۳) شناساپذیر است اگر و تنها اگر حداقل یکی از مدل‌های تکی آن شناساپذیر باشد.*

قضیه ۴

به ازای هر بردار پارامتری $(\boldsymbol{\theta}_{\mathbf{u}}, \boldsymbol{\theta}_{\varepsilon})$، بردار پارامتری $(\boldsymbol{\theta}_{\mathbf{u}}, \boldsymbol{\theta}_{\varepsilon}) \neq (\boldsymbol{\theta}^*_{\mathbf{u}}, \boldsymbol{\theta}^*_{\varepsilon})$ چنان وجود دارد که $Z[\boldsymbol{\Sigma}^*_{\mathbf{u}} - \boldsymbol{\Sigma}_{\mathbf{u}}]Z' = \boldsymbol{\Sigma}^*_{\varepsilon} - \boldsymbol{\Sigma}^*_{\varepsilon}$ برقرار است اگر و تنها اگر به ازای هر بردار پارامتری $\boldsymbol{\theta}_{\varepsilon}$، $\boldsymbol{\theta}_{\varepsilon} \neq \boldsymbol{\theta}^*_{\varepsilon}$ چنان وجود داشته باشد که $Z'\boldsymbol{\Sigma}^*_{\varepsilon}Z \neq Z'\boldsymbol{\Sigma}_{\varepsilon}Z$،

$\boldsymbol{\Sigma}^*_{\varepsilon} - \boldsymbol{\Sigma}_{\varepsilon} = H_Z[\boldsymbol{\Sigma}_{\varepsilon} - \boldsymbol{\Sigma}^*_{\varepsilon}]$ و

$$\boldsymbol{\Sigma}^*_{\mathbf{u}} = \boldsymbol{\Sigma}_{\mathbf{u}} + (Z'Z)^{-۱}Z'[\boldsymbol{\Sigma}_{\varepsilon} - \boldsymbol{\Sigma}^*_{\varepsilon}]Z(Z'Z)^{-۱}, \qquad (۳)$$

که در آن $H_Z = Z(Z'Z)^{-۱}Z'$.

قضیه ۵

مدل (۳) را مد نظر قرار دهید. فرض کنید $k > n$. اگر به ازای هر دو بردار پارامتری $\boldsymbol{\theta}_\varepsilon^*$ و $\boldsymbol{\theta}_\varepsilon$ عضو $\boldsymbol{\Theta}_\varepsilon$ که $\boldsymbol{\theta}_\varepsilon \neq \boldsymbol{\theta}_\varepsilon^*$، $\text{rank}\ (\boldsymbol{\Sigma}_\varepsilon - \boldsymbol{\Sigma}_\varepsilon^*) > k$، آنگاه مدل (۳) شناساپذیر است.

نتیجه ۶

فرض کنید ماتریس‌های کوواریانس $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_\varepsilon$ به ترتیب دارای ساختارهای UN و CS باشند. به علاوه ماتریس طرح Z در شرط‌های $\mathbf{1}'\mathbf{Z} \neq \mathbf{0}$ و $\text{rank}\ Z = k$ که در آن $۱ \leq k < (n - ۱)$، صدق کند. مدل (۳) شناساناپذیر است اگر و تنها اگر $H_Z\mathbf{1} = \mathbf{1}$.

قضیه ۵

مدل (۳) را مد نظر قرار دهید. فرض کنید k $n >$. اگر به ازای هر دو بردار پارامتری $\boldsymbol{\theta}_\varepsilon^*$ و $\boldsymbol{\theta}_\varepsilon$ عضو $\boldsymbol{\Theta}_\varepsilon$ که $\boldsymbol{\theta}_\varepsilon \neq \boldsymbol{\theta}_\varepsilon^*$، $\operatorname{rank}\ (\boldsymbol{\Sigma}_\varepsilon - \boldsymbol{\Sigma}_\varepsilon^*) > k$، آن‌گاه مدل (۳) شناساپذیر است.

نتیجه ۶

فرض کنید ماتریس‌های کوواریانس $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_\varepsilon$ به ترتیب دارای ساختارهای UN و CS باشند. به علاوه ماتریس طرح Z در شرط‌های $\mathbf{1}'\mathbf{Z} \neq \mathbf{0}$ و $\operatorname{rank} Z = k$ که در آن $(n - ۱) > k \leq ۱$، صدق کند. مدل (۳) شناساناپذیر است اگر و تنها اگر $H_Z\mathbf{1} = \mathbf{1}$.

قضیه ۵

مدل (۳) را مد نظر قرار دهید. فرض کنید $k > n$. اگر به ازای هر دو بردار پارامتری $\boldsymbol{\theta}_\varepsilon^*$ و $\boldsymbol{\theta}_\varepsilon$ عضو $\boldsymbol{\Theta}_\varepsilon$ که $\boldsymbol{\theta}_\varepsilon \neq \boldsymbol{\theta}_\varepsilon^*$، $\operatorname{rank}\ (\boldsymbol{\Sigma}_\varepsilon - \boldsymbol{\Sigma}_\varepsilon^*) > k$، آنگاه مدل (۳) شناساپذیر است.

نتیجه ۶

فرض کنید ماتریس‌های کوواریانس $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_\varepsilon$ به ترتیب دارای ساختارهای UN و CS باشند. به علاوه ماتریس طرح Z در شرط‌های $\mathbf{1}'\mathbf{Z} \neq \mathbf{0}$ و $\operatorname{rank} Z = k$ که در آن $(n - ۱) > k \leq ۱$، صدق کند. مدل (۳) شناساناپذیر است اگر و تنها اگر $H_Z\mathbf{1} = \mathbf{1}$.

نتیجه ۷

فرض کنید $k > n$. اگر $\Sigma_\varepsilon = \sigma^۲ R$ که در آن $\sigma^۲ > ۰$ و R یک ماتریس کوواریانس معلوم یا به عبارتی ساختار Σ_ε، MK باشد، آن‌گاه مدل (۳) شناساپذیر است.

نتیجه ۸

فرض کنید $k > n - ۱$. اگر ماتریس کوواریانس Σ_ε، ساختار $MA(۱)$ داشته باشد، آن‌گاه مدل (۳) شناساپذیر است.

نتیجه ۷

فرض کنید $n > k$. اگر $\Sigma_\varepsilon = \sigma^۲ R$ که در آن $\sigma^۲ > ۰$ و R یک ماتریس کوواریانس معلوم یا به عبارتی ساختار Σ_ε، MK باشد، آن‌گاه مدل (۳) شناساپذیر است.

نتیجه ۸

فرض کنید $n - ۱ > k$. اگر ماتریس کوواریانس Σ_ε، ساختار $MA(۱)$ داشته باشد، آن‌گاه مدل (۳) شناساپذیر است.

قضیه ۹

فرض کنید ماتریس Z به صورت بالا باشد. مدل (۳) تحت هر یک از شرایط زیر، شناساپذیر نخواهد بود:

آ) $z = -v$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} به ترتیب ساختارهای VC و MI داشته باشند.

ب) $z = -v$، $z^{۲} = ۱$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} هر دو ساختار MI داشته باشند.

پ) $zv = -۱$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} به ترتیب ساختارهای MI و VC داشته باشند.

قضیه ۹

فرض کنید ماتریس Z به صورت بالا باشد. مدل (۳) تحت هر یک از شرایط زیر، شناساپذیر نخواهد بود:

آ) $z=-v$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} به ترتیب ساختارهای VC و MI داشته باشند.

ب) $z=-v$، $z^{۲}=۱$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} هر دو ساختار MI داشته باشند.

پ) $zv=-۱$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} به ترتیب ساختارهای MI و VC داشته باشند.

قضیه ۹

فرض کنید ماتریس Z به صورت بالا باشد. مدل (۳) تحت هر یک از شرایط زیر، شناساپذیر نخواهد بود:

آ) $z = -v$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} به ترتیب ساختارهای VC و MI داشته باشند.

ب) $z = -v$، $z^{۲} = ۱$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} هر دو ساختار MI داشته باشند.

پ) $zv = -۱$، $\Sigma_{\mathbf{u}}$ و Σ_{ε} به ترتیب ساختارهای MI و VC داشته باشند.

قضیه ۹

فرض کنید ماتریس Z به صورت بالا باشد. مدل (۳) تحت هر یک از شرایط زیر، شناساپذیر نخواهد بود:

آ) $z = -v$، $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_{\varepsilon}$ به ترتیب ساختارهای VC و MI داشته باشند.

ب) $z = -v$، $z^۲ = ۱$، $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_{\varepsilon}$ هر دو ساختار MI داشته باشند.

پ) $zv = -۱$، $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_{\varepsilon}$ به ترتیب ساختارهای MI و VC داشته باشند.

قضیه ۹

فرض کنید ماتریس Z به صورت بالا باشد. مدل (۳) تحت هر یک از شرایط زیر، شناساپذیر نخواهد بود:

آ) $z = -v$، $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_{\varepsilon}$ به ترتیب ساختارهای VC و MI داشته باشند.

ب) $z = -v$، $z^{۲} = ۱$، $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_{\varepsilon}$ هر دو ساختار MI داشته باشند.

پ) $zv = -۱$، $\boldsymbol{\Sigma}_{\mathbf{u}}$ و $\boldsymbol{\Sigma}_{\varepsilon}$ به ترتیب ساختارهای MI و VC داشته باشند.

قضیه ۱۰

مدل (۳)، تحت هر یک از شرایط زیر شناساپذیر است:

آ) $\Sigma_{\mathbf{u}}$ ماتریسی معلوم باشد.

ب) Σ_{ε} ماتریسی معلوم باشد.

پ) $K = Z\Sigma_{\mathbf{u}}Z'\Sigma_{\varepsilon}^{-1}$ که در آن K، ماتریسی معلوم بوده و $I + K$ پررتبه‌ی ستونی است.

قضیه ۱۰

مدل (۳)، تحت هر یک از شرایط زیر شناساپذیر است:

آ) $\boldsymbol{\Sigma}_{\mathbf{u}}$ ماتریسی معلوم باشد.

ب) Σ_{ε} ماتریسی معلوم باشد.

پ) $K = Z\Sigma_{\mathbf{u}}Z'\Sigma_{\varepsilon}^{-1}$ که در آن K، ماتریسی معلوم بوده و $I + K$ پررتبه‌ی ستونی است.

قضیه ۱۰

مدل (۳)، تحت هر یک از شرایط زیر شناساپذیر است:

آ) $\Sigma_{\mathbf{u}}$ ماتریسی معلوم باشد.

ب) Σ_{ε} ماتریسی معلوم باشد.

پ) $K = Z\Sigma_{\mathbf{u}}Z'\Sigma_{\varepsilon}^{-1}$ که درآن K، ماتریسی معلوم بوده و $I + K$ پررتبه‌ی ستونی است.

قضیه ۱۰

مدل (۳)، تحت هر یک از شرایط زیر شناساپذیر است:

آ) $\boldsymbol{\Sigma}_{\mathbf{u}}$ ماتریسی معلوم باشد.

ب) $\boldsymbol{\Sigma}_{\varepsilon}$ ماتریسی معلوم باشد.

پ) $K = Z\boldsymbol{\Sigma}_{\mathbf{u}}Z'\boldsymbol{\Sigma}_{\varepsilon}^{-1}$ که در آن K، ماتریسی معلوم بوده و $I + K$ پررتبه‌ی ستونی است.

مثال: فروشگاه‌ها!

مدل ۱

$$Y_{ij} = \alpha + \beta x_{ij} + u_i + \varepsilon_{ij}, \qquad u_i \sim N(۰, \sigma_u^{۲}), \quad \varepsilon_{ij} \sim N(۰, \sigma^{۲})$$

آیا این مدل توأم مطلوب است؟

مدل ۲

$$Y_i \overset{r}{=} \mathbf{1}_{n_i} u_i + \varepsilon_i$$

ساختار Σ_{ε_i}	توضیح‌ها	شناساپذیر	شناساناپذیر

مثال: فروشگاه‌ها!

مدل ۱

$$Y_{ij} = \alpha + \beta x_{ij} + u_i + \varepsilon_{ij}, \qquad u_i \sim N(۰, \sigma_u^۲), \quad \varepsilon_{ij} \sim N(۰, \sigma^۲)$$

آیا این مدل توأم مطلوب است؟

مدل ۲

$$Y_i \stackrel{r}{=} \mathbf{1}_{n_i} u_i + \varepsilon_i$$

ساختار Σ_{ε_i}	توضیح‌ها	شناساپذیر	شناساناپذیر

مثال: فروشگاه‌ها!

مدل ۱

$$Y_{ij} = \alpha + \beta x_{ij} + u_i + \varepsilon_{ij}, \qquad u_i \sim N(۰, \sigma_u^۲), \quad \varepsilon_{ij} \sim N(۰, \sigma^۲)$$

آیا این مدل توأم مطلوب است؟

مدل ۲

$$Y_i \overset{r}{=} \mathbf{1}_{n_i} u_i + \varepsilon_i$$

ساختار Σ_{ε_i}	توضیح‌ها	شناساپذیر	شناساناپذیر

مثال: فروشگاه‌ها!

مدل ۱

$$Y_{ij} = \alpha + \beta x_{ij} + u_i + \varepsilon_{ij}, \qquad u_i \sim N(۰, \sigma_u^۲), \quad \varepsilon_{ij} \sim N(۰, \sigma^۲)$$

آیا این مدل توأم مطلوب است؟

مدل ۲

$$Y_i \overset{r}{=} \mathbf{1}_{n_i} u_i + \varepsilon_i$$

ساختار Σ_{ε_i}	توضیح‌ها	شناساپذیر	شناساناپذیر

مثال: فروشگاه‌ها!

مدل ۱

$$Y_{ij} = \alpha + \beta x_{ij} + u_i + \varepsilon_{ij}, \qquad u_i \sim N(۰, \sigma_u^۲), \quad \varepsilon_{ij} \sim N(۰, \sigma^۲)$$

آیا این مدل توأم مطلوب است؟

مدل ۲

$$Y_i \overset{r}{=} \mathbf{1}_{n_i} u_i + \varepsilon_i$$

ساختار Σ_{ε_i}	توضیح‌ها	شناساپذیر	شناساناپذیر
MI	نتیجه ۱۱	✓	

مثال: فروشگاه‌ها!

مدل ۱

$$Y_{ij} = \alpha + \beta x_{ij} + u_i + \varepsilon_{ij}, \qquad u_i \sim N(۰, \sigma_u^۲), \quad \varepsilon_{ij} \sim N(۰, \sigma^۲)$$

آیا این مدل توأم مطلوب است؟

مدل ۲

$$Y_i \overset{r}{=} \mathbf{1}_{n_i} u_i + \varepsilon_i$$

ساختار Σ_{ε_i}	توضیح‌ها	شناساپذیر	شناساناپذیر
MI	نتیجه ۱۱	$\checkmark$	
CS	نتیجه ۱۰		$\times$

مثال: فروشگاه‌ها!

مدل ۱

$$Y_{ij} = \alpha + \beta x_{ij} + u_i + \varepsilon_{ij}, \qquad u_i \sim N(۰, \sigma_u^۲), \quad \varepsilon_{ij} \sim N(۰, \sigma^۲)$$

آیا این مدل توأم مطلوب است؟

مدل ۲

$$Y_i \overset{r}{=} \mathbf{1}_{n_i} u_i + \varepsilon_i$$

ساختار Σ_{ε_i}	توضیح‌ها	شناساپذیر	شناساناپذیر
MI	نتیجه ۱۱	$\checkmark$	
CS	نتیجه ۱۰		$\times$
UN	بیش‌پارامتری شده		$\times$

بخش چهارم

نتیجه‌گیری و پیشنهادها

پیشنهادها:

1. تحلیل بیزی مدل‌های بیان شده؛
2. تحلیل مدل‌های توأم اندازه‌های طولی پیوسته با تکیه‌گاه بازه‌ای و برآمدهای زمان تا رخداد؛
3. رگرسیون سری زمانی برای مدل‌سازی پاسخ‌هایی با تکیه‌گاه بازه‌ای؛
4. ارائه‌ی شرایط لازم و کافی برای شناساپذیری مدل‌های بحث شده؛
5. تحلیل مدل‌های بیان شده برای داده‌های سانسور شده؛
6. از آنجایی که ممکن است تعداد نقاط آماسیدگی، تصادفی باشد، مطالعه در مورد تشخیص نقاط آماسیدگی و آزمون‌های مربوط به آن از اهداف ما خواهد بود.

نتیجه‌گیری:
شناساپذیری یکی از ویژگی‌های لازم برای کفایت یک مدل آماری است. وقتی مدلی شناساپذیر نباشد، با هیچ اندازه‌ای از نمونه، نمی‌توان پارامتر حقیقی مدل را تعیین کرد. از سوی دیگر، معمولاً نرم‌افزارهای آماری، بعد از برازش مدل شناساناپذیر در آن‌ها، اشاره‌ای به این مسأله نکرده و خروجی‌های غیرموثق ارائه می‌دهند. بنابراین یافتن راهی برای بررسی شناساپذیری مدل قبل از برازش آن، خالی از فایده نخواهد بود. در این ارائه مروری بر مفهوم این پیشنیاز اساسی دهر مدل پارامتری آماری شد.

مراجع

Demidenko, E. (۲۰۰۴).
Mixed Models: Theory and Applications. John Wiley & Sons.

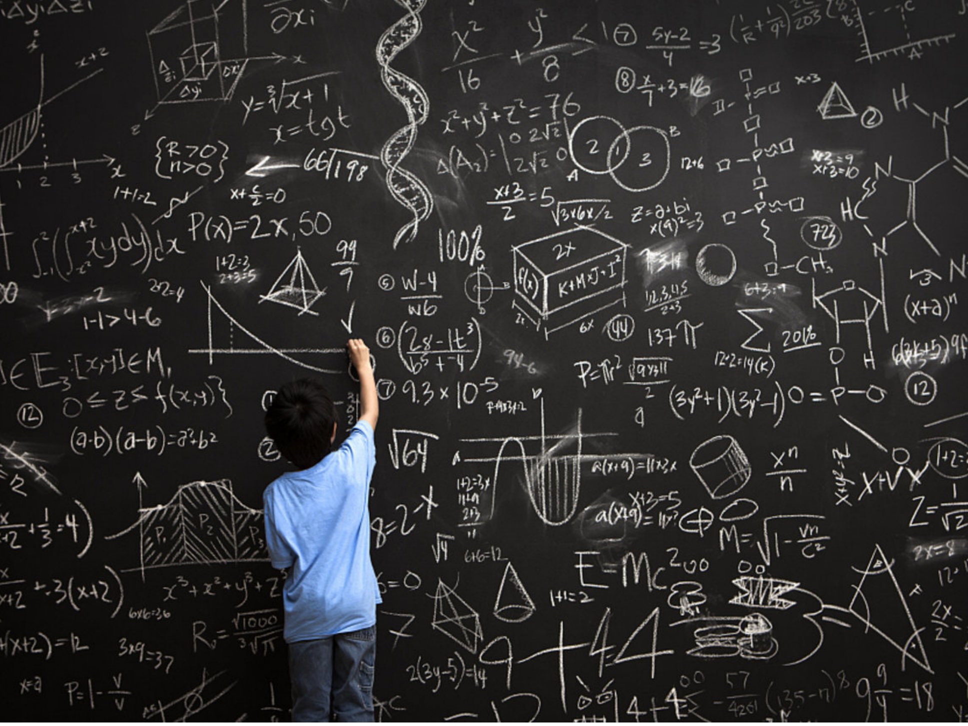
Shao, J. (۲۰۰۳).
Mathematical Statistics (Second ed.). Springer.

Wang, W. (۲۰۱۳).
Identifiability of linear mixed effects models. Electronic Journal of Statistics, **۷**, ۲۴۴ – ۲۶۳.

Tabrizi, E., Samani, E. B., and Ganjali, M. (۲۰۱۹).
A note on the identifiability of latent variable models for mixed longitudinal data. Statistics and Probability Letters.

Tabrizi, E., Samani, E. B., and Ganjali, M. (۲۰۲۰).
Identifiability of parameters in longitudinal correlated poisson and inflated beta regression model with non-ignorable missing mechanism. Statistics.

از توجه و صبر شما سپاسگزارم



$\frac{x^2}{16}+\frac{y^2}{8}$
$x^2+y^2+z^2=76$
$P(x)=2x,50$
$\frac{x+3}{2}=5$
100%
$(a-b)(a-b)=a^2+b^2$
9.3×10^{-5}
$\sqrt{64}$
$E=Mc^2$
1+1=2
6x6=36
3x9=37