استنباط بیزی مقیّد

جواد اطمینان
گروه آمار، دانشگاه بیرجند

مسائل استنباط آماری مقیّد در عرصه‌ای گسترده از کاربردها نمایان می‌شوند، هرچند قسمت ناچیزی از نوشتارهای آماری در بر گیرندهٔ دیدگاه بیزی این موضوع است. از دیدگاه فنی، محاسبه‌های بیزی اغلب مسئلهٔ غامضی در زمینهٔ نامقیّد است و برای انواع مسائل مقیّد مورد نظر ما می‌تواند تقریباً ناممکن باشد. در این مقاله ابتدا به معرفی روش نمونه‌گیر گیبز به عنوان ابزاری برای محاسبه‌های بیزی (نوعاً توزیع پسین حاشیه‌ای و امید آن) و سپس به بررسی ساختار کلی مسئله می‌پردازیم. در ادامه دو مدل خاص، شامل پارامترهای چند جمله‌ای مقیّد و پارامترهای مدل خطّی مقیّد را برای مثال به همراه تحلیل دو مجموعه از داده‌های آنها ارائه شده است.

واژه‌های کلیدی: استنباط بیزی، انتگرال‌گیری مونت کارلویی، پارامترهای مقیّد، نمونه‌گیر گیبز.

۱ مقدمه

مسائل استنباط آماری مقیّد در عرصه‌ای گسترده از کاربردها، از قبیل عیارسنجی زیستی، برآورد تابعها در بیمهٔ آمار، داده‌های رسته‌ای ترتیبی، رویه‌های پاسخ، آزمون نمو قابلیّت اعتماد و مدلهای مؤلّفهٔ واریانس نمایان می‌شوند. اگرچه نوشتارهای آماری، هم از نظر کاربردی و هم از جنبهٔ نظری، دربارهٔ این گونه مسائل بسیار زیادند، ولی قسمت ناچیزی از آنها در بر گیرندهٔ دیدگاه بیزی این موضوع است. از دیدگاه فنی، انتگرال عددی (چندگانه) ابزار بنیادی برای محاسبه‌های بیزی در مدلهای نوعی واقعی است که اغلب مسئلهٔ غامضی در زمینهٔ نامقیّد است و برای انواع مسائل مقیّد مورد نظر ما می‌تواند تقریباً ناممکن باشد. خوشبختانه نمونه‌گیر گیبز می‌تواند به صورت رایج و بدون پیچیدگی برای انجام محاسبه‌های بیزی به طور کلی و برای پارامتر مقیّد مورد نظر ما به طور ویژه به کار گرفته شود. گلفاند و همکاران (۱۹۹۲) به کارامدی استفاده از نمونه‌گیری گیبز در تحلیل بیزی مسائل پارامترهای

مقیّد و داده‌های بریده شده پرداخته‌اند.

۲ نمونه‌گیر گیبز

به عنوان یک نماد قراردادی، چگالیها را با قرار دادن متغیر(ها) بین دو کروشه نشان می‌دهیم. $[U,V]$، $[U|V]$ و $[U]$ به ترتیب نشانگر چگالی توأم U و V، چگالی شرطی U به شرط V و چگالی حاشیه‌ای U هستند. تعیین توزیع حاشیه‌ای توسط انتگرال‌گیری را به صورت زیر نشان می‌دهیم:

$$[U] = \int [U|V][V].$$

برای گردایه‌ای از متغیرهای تصادفی $\{U_1,\dots,U_k\}$، چگالیهای شرطی کامل را به صورت،

$$[U_s|U_r, r\neq s], \quad s=1,\dots,k$$

و چگالیهای حاشیه‌ای را به صورت $[U_s]$، $s=1,\dots,k$ می‌توان نشان داد.

نمونه‌گیری گیبز یک الگوی مارکوفی است. برای اجرای نمونه‌گیر گیبز، در ابتدا یک مجموعهٔ اختیاری از مقادیر $U_1^{(0)}$،...،$U_k^{(0)}$ در نظر گرفته می‌شود و سپس $U_1^{(1)}$ از $[U_1|U_2^{(0)},\dots,U_k^{(0)}]$، $U_2^{(1)}$ از $[U_2|U_1^{(1)},U_3^{(0)},\dots,U_k^{(0)}]$ و سرانجام $U_k^{(1)}$ از $[U_k|U_1^{(1)},\dots,U_{k-1}^{(1)}]$ استخراج می‌شود تا یک تکرار از این الگو کامل شود. در این الگو هر متغیر در همان ترتیب طبیعی خودش قرار می‌گیرد و در هر چرخه لازم است k متغیر تصادفی تولید شود. بعد از انجام t بار از این گونه تکرارها، نمونه $(U_1^{(t)},\dots,U_k^{(t)})$ حاصل می‌شود. گِمَن و گِمَن (۱۹۸۴)، نشان دادند که تحت شرایطی معتدل، نتایج زیر برقرارند:

همگرایی: $\left(U_1^{(t)},\dots,U_k^{(t)}\right) \xrightarrow{d} (U_1,\dots,U_k)\sim[U_1,\dots,U_k]$، هرگاه $t\to\infty$.
و لذا برای هر s، $U_s^{(t)}\xrightarrow{d}U_s\sim[U_s]$، هرگاه $t\to\infty$.

آهنگ همگرایی: در صورتی که از نُرم سوپریمم به جای نُرم L_1 استفاده شود، تحت ترتیب طبیعی، چگالی توأم $(U_1^{(t)},\dots,U_k^{(t)})$ به چگالی توأم واقعی با یک آهنگ هندسی بر حسب t همگراست.

قضیه ارگودیک: برای هر تابع اندازه‌پذیر T از U_1،...،U_k که امیدش موجود باشد، اگر $t\to\infty$، آنگاه:

$$\frac{1}{t}\sum_{l=1}^{t} T(U_1^{(l)},\dots,U_k^{(l)})\xrightarrow{a.s.}E\left(T(U_1,\dots,U_k)\right).$$

بنابراین برای t به اندازهٔ کافی بزرگ، مثلاً $U_s^{(t)}$ را می‌توان به عنوان نمونه‌ای از U_s در نظر گرفت. حال اگر این فرایند را m بار به طور مستقل از هم تکرار کنیم، k تایی‌های،

$$(U_{1j}^{(t)}, \ldots, U_{kj}^{(t)}), \quad j = 1, \ldots, m$$

نتیجه می‌شوند. برآورد هسته‌ای چگالی برای $[U_s]$ می‌تواند بر اساس $U_{sj}^{(t)}$ به دست آید (سیلورمن، ۱۹۸۶) و اگر m به اندازهٔ کافی بزرگ باشد، نتیجه رضایت‌بخش است. برآورد بهتری از چگالی تحت دامنه‌ای وسیع از تابع‌های زیان با استفاده از الگوریتم رائو ـ بلک‌ولی (گلفاند و اسمیت، ۱۹۹۰) به صورت زیر است:

$$[\hat{U}_s] = \frac{1}{m}\sum_{j=1}^{m}[U_s|U_{rj}^{(t)}, r \neq s]. \qquad (1)$$

به طور مشابه، $E\left(h(U_s)\right)$ را می‌توان از طریق یک برآورد نمونه‌ای بر اساس $U_{sj}^{(t)}$ یا از طریق یک نسخهٔ رائو ـ بلک‌ولی مشابه (۱) بر اساس $E\left(h(U_s)|U_r, r \neq s\right)$ برآورد کرد.

حال $W(U_1, \ldots, U_k)$ را تابعی از U_iها فرض کنید. هر k تایی $(U_{1j}^{(t)}, \ldots, U_{kj}^{(t)})$ یک مشاهدهٔ $W_j^{(t)} \equiv W(U_{1j}^{(t)}, \ldots, U_{kj}^{(t)})$ با توزیع حاشیه‌ای تقریباً برابر $[W]$ ایجاد می‌کند، وقتی که $t \to \infty$. از این رو یک برآورد هسته‌ای چگالی برای $[W]$ با استفاده از $W_j^{(t)}$ها می‌تواند حاصل شود.

۳ ساختار کلی مسئله

برای بردار داده‌های $\boldsymbol{Y}$، یک مدل پارامتری شامل بردار پارامتر k بعدی $\boldsymbol{\theta}$ را با این قید که $\boldsymbol{\theta}$ متعلّق به زیر مجموعه‌ای از $\mathcal{R}^k$ مانند $\mathcal{S}_Y^k$ باشد در نظر بگیرید. توجه کنید که تکیه‌گاه مثبت توزیع $\boldsymbol{Y}$ ممکن است بستگی به $\boldsymbol{\theta}$ داشته باشد. مدل بیزی مقیّد (پیشین × درستنمایی) به صورت،

$$\begin{cases} [\boldsymbol{Y}|\boldsymbol{\theta}][\boldsymbol{\theta}|\boldsymbol{\lambda}], & (\boldsymbol{y}, \boldsymbol{\theta}) \in \mathcal{S}, \\ 0, & (\boldsymbol{y}, \boldsymbol{\theta}) \notin \mathcal{S}, \end{cases}$$

داده می‌شود که در آن $\mathcal{S} = \left\{(\boldsymbol{y}, \boldsymbol{\theta}) : \boldsymbol{\theta} \in \mathcal{S}_Y^k\right\}$. به طور کلی $[\boldsymbol{Y}|\boldsymbol{\theta}]$ و $[\boldsymbol{\theta}|\boldsymbol{\lambda}]$ به عنوان توابعی از $\boldsymbol{\theta}$، در صورتی که قیدها نادیده گرفته شده باشند دارای صورتهای تابعی داده شده‌ای هستند.

گزاره ۱ فرض کنید $[\boldsymbol{\theta}|\boldsymbol{\lambda}]$ توزیع پیشینی و $[\boldsymbol{\theta}|\boldsymbol{Y}]$ توزیع پسینی نامقید برای $\boldsymbol{\theta}$ باشند. هرگاه $\boldsymbol{\theta}$ محدود به مجموعهٔ $\mathcal{S}_Y^k$ باشد، آنگاه:

الف)
$$[\boldsymbol{\theta}|\boldsymbol{Y}]^* = \frac{[\boldsymbol{\theta}|\boldsymbol{Y}]}{\int_{\mathcal{S}_Y^k}[\boldsymbol{\theta}|\boldsymbol{Y}]},$$

که در آن $[\boldsymbol{\theta}|\boldsymbol{Y}]^*$ توزیع پسینی مقید $\boldsymbol{\theta}$ است. در واقع، توزیع پسینی مقید $\boldsymbol{\theta}$ به زبان ساده همان توزیع پسینی نامقید است که به طرز مناسب نرمالیده شده است.

ب)
$$[\boldsymbol{Y}]^* = \frac{\int_{\mathcal{S}_Y^k}[\boldsymbol{\theta}|\boldsymbol{Y}]}{\int_{\mathcal{S}^k}[\boldsymbol{\theta}]}[\boldsymbol{Y}],$$

که در آن $[\boldsymbol{Y}]^*$ توزیع حاشیه‌ای مقید $\boldsymbol{Y}$ است.

حال $\mathcal{S}_i^k(\theta_j, j \neq i)$ را مقطعی از $\mathcal{S}_Y^k$ در نظر بگیرید که به صورت قیدهایی بر روی θ_i در یک مجموعه مقادیر تعیین شدهٔ θ_j، $j \neq i$، تعریف می‌شود که می‌تواند بستگی به $\boldsymbol{Y}$ نیز داشته باشد. در حالتی که مؤلّفه‌ها اسکالر هستند، $\mathcal{S}_i^k(\theta_j, j \neq i)$ زیر مجموعه‌ای از $\mathcal{R}^1$ و نوعاً یک بازه یا یک گردایه از بازه‌ها است. مستقیماً از قسمت (الف) گزارهٔ ۱.۲، نتیجه می‌شود که توزیع شرطی پسینی کامل θ_i عبارت است از:

$$[\theta_i|\boldsymbol{Y}, \lambda, \theta_j, j \neq i] \propto [\boldsymbol{Y}|\boldsymbol{\theta}][\boldsymbol{\theta}|\boldsymbol{\lambda}], \qquad \theta_i \in \mathcal{S}_i^k(\theta_j, j \neq i), \qquad (۲)$$

فرض کنید تابع توزیع شرطی کامل θ_i فراهم باشد و آن را با F_i نشان دهیم و مقطع مربوط بازه‌ای به صورت $[a, b]$ باشد. در صورتی که متغیر تصادفی U، دارای توزیع یکنواخت $(۰, ۱)$ باشد،

$$\theta_i = F_i^{-1}[F_i(a) + U(F_i(b) - F_i(a))],$$

یک برآمد از توزیع شرطی کامل مقیّد است و با این شیوه می‌توان نمونه‌ای تصادفی تولید کرد. برای حالتی که مقطع، گردایه‌ای از بازه‌های مجزا از هم به صورت $\cup_{l=1}^r[a_l, b_l]$ است، این روش را به سادگی می‌توان بسط داد. در این حالت، با احتمال،

$$[\sum_{l=1}^r(F_i(b_l) - F_i(a_l))]^{-1}[F_i(b_j) - F_i(a_j)],$$

$J = j$ را انتخاب می‌کنیم و سپس θ_i را از بازه $[a_j, b_j]$ به صورت،

$$\theta_i = F_i^{-1}[F_i(a_j) + U(F_i(b_j) - F_i(a_j))],$$

تولید می‌کنیم.

سرانجام به جنبه‌ای دیگر از به کارگیری نمونه‌گیر گیبز اشاره می‌کنیم که در محاسبهٔ یک مدل سلسله مراتبی ظاهر می‌شود که در آن برای ابرپارامتر $\boldsymbol{\lambda}$، که پیش از این معلوم فرض کردیم، پیشین $[\boldsymbol{\lambda}]$ در نظر گرفته می‌شود. توزیع شرطی کامل θ_i که در (۲) داده شده، تغییر نمی‌کند امّا توزیع شرطی کامل $\boldsymbol{\lambda}$ به $\boldsymbol{Y}$ بستگی ندارد و به صورت،

$$[\boldsymbol{\lambda}|\boldsymbol{Y},\boldsymbol{\theta}] \propto [\boldsymbol{\theta}|\boldsymbol{\lambda}][\boldsymbol{\lambda}]c(\boldsymbol{\lambda})\,, \qquad (۳)$$

است که در آن $c(\boldsymbol{\lambda}) = (\int_{\mathcal{S}}[\boldsymbol{Y}|\boldsymbol{\theta}][\boldsymbol{\theta}|\boldsymbol{\lambda}])^{-۱}$. اگر θ مقیّد به $\boldsymbol{Y}$ نباشد $c(\boldsymbol{\lambda})$ دارای صورت ساده‌تر $(\int_{\mathcal{S}^k}[\boldsymbol{\theta}|\boldsymbol{\lambda}])^{-۱}$ است.

۴ مدلها: مثالهای ویژه

در این بخش به ارائهٔ صورتهای صریح نمونه‌گیر گیبز برای دو مثال مختلف می‌پردازیم که در بخش ۳ ساختار کلی آنها مورد بحث قرار گرفت. این بحث برای نشان دادن سادگی شگفت‌آور نمونه‌گیر گیبز در حل مشکل محاسبات بیزی در زمینهٔ پارامتر مقیّد است در صورتی که به طور مناسب طراحی شده باشد.

۴-۱ پارامترهای مرتب در توزیع چندجمله‌ای

سِدرانسک و همکاران (۱۹۸۵)، مسئلهٔ برآورد بیزی پارامترهای جامعهٔ متناهی را وقتی که متغیر تصادفی X یکی از مقادیر مجموعهٔ متناهی $\{b_۱,\ldots,b_k\}$ را با احتمالهای $p_۱$، …، p_k می‌گیرد مورد بررسی قرار دادند. یک کاربرد خاص این مدل، درآمد خانوار است که b_j می‌تواند بیانگر مقدار میانی درآمد برای طبقهٔ jام (نشانه) باشد.

فرض کنید که طبقه‌ها به ترتیب افزایشی منظم شده‌اند. انتظار داریم که با افزایش j، p_jها تا طبقه‌ای، مثلاً t $(۱ \leq t \leq k)$، افزایشی و سپس کاهشی باشند. نوعاً t نامعلوم است. در چنین حالتی، کمّیّت مورد علاقهٔ ابتدائی می‌تواند میانگین جامعهٔ متناهی، یعنی $\sum_{j=۱}^{k} b_j p_j$، باشد هرچند که می‌توان هر تابعی از $\boldsymbol{p} = (p_۱,\ldots,p_k)'$ را مورد نظر قرار داد. یک مدل بیزی برای این گونه مسائل به این صورت است که Y_j تعداد مشاهدات طبقهٔ jام تعریف شود،

$\sum_{j=1}^{k} Y_j = n$. در این صورت $\boldsymbol{Y}$ به شرط $\boldsymbol{p}$ دارای توزیع چند جمله‌ای است:

$$[\boldsymbol{Y}|\boldsymbol{p}] = Mu(n; p_1, \ldots, p_k)\,. \tag{۴}$$

با پیروی از سِدرانسک و همکاران (۱۹۸۵)، به شرط معلوم بودن t، برای $\boldsymbol{p}$ پیشین را به صورت،

$$c(\beta_1, \ldots, \beta_k; t) \prod_{j=1}^{k} p_j^{\beta_j - 1}\,, \quad \boldsymbol{p} \in \mathcal{S}^k, \tag{۵}$$

اختیار می‌کنیم که در آن،

$$\mathcal{S}^k = \{(p_1, \ldots, p_k):\ p_1 \leq p_2 \leq \ldots \leq p_t \geq p_{t+1} \geq \ldots \geq p_k,\ 0 \leq p_j \leq 1, \sum_{j=1}^{k} p_j = 1\}$$

و $c(\beta_1, \ldots, \beta_k; t)$ ثابت نرمالیده است.
به منظور به کارگیری نمونه‌گیر گیبز، توزیع شرطی کامل p_iها، یعنی $[p_i|\boldsymbol{Y}, t, p_j,\ j \neq i]$، $i = 1, \ldots, k-1$، (p_k به عنوان تابعی از بقیهٔ p_iها) را نیاز داریم. واضح است که این یک توزیع بتای مقیاسیده به بازهٔ $[0, 1 - \sum_{j=1, j\neq i}^{k-1} p_j]$ است که بر طبق قید مشخص شده توسط t، به طور مناسب مقیّد شده است.

حال فرض کنید t نامعلوم است. برای آن پیشینی گسسته به صورت،

$$\Pr(t = j) = \pi_j\,,\ j = 1, \ldots, k,$$

در نظر می‌گیریم. توجّه کنید که $[t|\boldsymbol{Y}, \boldsymbol{p}]$ یک توزیع تباهیده است. بنابراین نمونه‌گیر گیبز به طور مستقیم نمی‌تواند به کار گرفته شود، زیرا یک شرط برای همگرایی آن این است که انتقال از یک مقدار t به هر مقدار دیگر آن ممکن باشد. خوشبختانه می‌توان پسین حاشیه‌ای t را مستقیماً تعیین کرد و دارای صورت زیر است:

$$\Pr(t = j|\boldsymbol{y}) = \frac{c(\beta_1, \ldots, \beta_k; j)\pi_j / c(\beta_1 + y_1, \ldots, \beta_k + y_k; j)}{\sum_{i=1}^{k} c(\beta_1, \ldots, \beta_k; i)\pi_i / c(\beta_1 + y_1, \ldots, \beta_k + y_k; i)}\,, \tag{۶}$$

با استفاده از انتگرال‌گیری مونت کارلویی به روش نمونه‌گیری نقاط مهم، می‌توان تعداد $2k$ مقدار ثابت که در (۶) ظاهر می‌شوند، را تعیین کرد. بنابراین پسین حاشیه‌ای p_iها را بر اساس

p_i	۰٫۰۳	۰٫۰۷	۰٫۱۰	۰٫۲۵	۰٫۳۰	۰٫۱۲	۰٫۰۸	۰٫۰۵
y_i	۱	۴	۱	۱۲	۱۳	۴	۴	۱

جدولِ ۱: جامعهٔ چند جمله‌ای و داده‌های تولید شده

t	۱	۲	۳	۴	۵	۶	۷	۸
$\Pr(t\|y)$	۰٫۰۰۰	۰٫۰۰۰	۰٫۰۰۱	۰٫۳۵۳	۰٫۶۳۵	۰٫۰۱۰	۰٫۰۰۱	۰٫۰۰۰

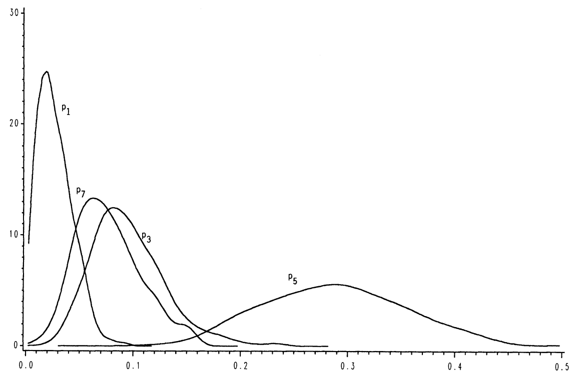
جدولِ ۲: توزیع پسین حاشیه‌ای t

رابطه،

$$[p_i|\boldsymbol{Y}] = \sum_{t=1}^{k} [p_i|\boldsymbol{Y}, t][t|\boldsymbol{Y}], \qquad i = 1, \dots, k-1, \tag{۷}$$

می‌توان برآورد کرد. ابتدا به طور تصادفی t را بر طبق $[t|\boldsymbol{Y}]$ انتخاب کرده و سپس از $[p_i|\boldsymbol{Y}, t]$ نمونه‌گیری می‌شود. به عنوان مثالی از مسئلهٔ مورد بحث، جدول ۱، یک مدل چندجمله‌ای با $k = 8$ خانه و نتایج نمونهٔ تصادفی ۴۰تایی استخراج شده از این مدل را نشان می‌دهد. فرض کنید که p_i برای $i = 1, \dots, t$ افزایشی و بعد از آن کاهشی هستند ولی p_iها و t نامعلوم هستند. پیشین را دیریکلهٔ یکنواخت تعمیم‌یافته، $\alpha_i = 1$ ، $i = 1, \dots, 8$ ، در نظر گرفته و ثابتهای $c(1, \dots, 1; t)$ و $c(1 + y_1, \dots, 1 + y_8; t)$ برای $t = 1, \dots, 8$، را به کمک انتگرال‌گیری مونت کارلویی و به روش نمونه‌گیری نقاط مهم حساب می‌شود و با به کارگیری (۶)، پسین حاشیه‌ای $[t|\boldsymbol{Y}]$ به دست می‌آید که در جدول ۲ آورده شده است. توجّه کنید که با وجود فقط ۴۰ مشاهده از یک جدول دارای ۸ خانه و یک پیشین یکنواخت (تخت)، $[t|\boldsymbol{Y}]$ تقریباً تمام جرم خود را بر روی نقاط ۴ و ۵ قرار داده است.

با توجّه به اینکه $[t|\boldsymbol{Y}]$ پیش از این به دست آمده است، برای نمونه گرفتن از $[p_i|\boldsymbol{Y}]$ پیشنهاد می‌شود که ابتدا t را بر طبق $[t|\boldsymbol{Y}]$ انتخاب کرده و سپس p_i از $[p_i|\boldsymbol{Y}, t]$ نمونه‌گیری شود. ازدیدگاه نظری قادر به دست آوردن برآورد هسته‌ای چگالی برای هر $[p_i|\boldsymbol{Y}, t]$، $t = 1, \dots, k$، هستیم و در نتیجه بر اساس (۷) برآوردی از چگالی حاصل می‌شود که ترکیبی متناهی از آنهاست. در عمل صرفاً t بر طبق $[t|\boldsymbol{Y}]$ به طور تصادفی انتخاب شده و سپس یک انتخاب با شانس برابر با آن از p_{ij}^t، $j = 1, \dots, m$، انجام می‌گیرد. این شیوهٔ بازنمونه‌گیری، مشاهده‌ای را نتیجه می‌دهد که تقریباً بر طبق $[p_i|\boldsymbol{Y}]$ توزیع شده است. با تکرار این فرآیند به تعداد زیاد (۱۰۰۰ بار برای ایجاد جدول ۳) می‌توانیم برآورد هسته‌ای چگالی $[p_i|\boldsymbol{Y}]$ را محاسبه کنیم.



شکلِ ۱: مدل چندجمله‌ای با پارامترهای مرتب، توزیعهای پسین حاشیه‌ای برای احتمال خانه‌های انتخاب شده

i	۱	۲	۳	۴	۵	۶	۷	۸
مد	۰٫۰۱۹	۰٫۰۶۱	۰٫۰۸۲	۰٫۲۴۶	۰٫۲۸۹	۰٫۰۹۶	۰٫۰۶۳	۰٫۰۱۹

جدولِ ۳: مد پسین حاشیه‌ای p_i

با توجّه به هدف بیان شده در ابتدای این زیربخش، برآوردهای هسته‌ای چگالی برای $p_۱$، $p_۳$، $p_۵$ و $p_۷$ در شکل۱ رسم شده‌اند. توجّه کنید که این پسینها منعکس کنندهٔ قیدهای ترتیب و دارای مدهای نزدیک به مقادیر واقعی متناظرشان هستند. مجموعهٔ کامل مدهای پسینی در جدول ۳ آورده شده است.

۴-۲ پارامترهای مرتب در مدل خطّی

در این بخش قصد داریم با در نظر گرفتن تحلیل یک طرح دوطرفه ساده، توان بالقوهٔ نمونه‌گیر گیبز در تحلیل بیزی پارامترهای مقیّد در مدلهای خطّی نرمال کلی را نشان دهیم. به کارگیری میانگینهای نرمال بدون ساختار خطّی، توسط گلفاند و اسمیت (۱۹۹۱)، ارائه شده و به کارگیری شیبهای مرتب در یک مدل رگرسیونی نقطهٔ تغییر توسط کارلین و همکاران (۱۹۹۲)، داده شده است. با بسط ذیل، گسترش به دیگر مسائل واضح خواهد بود.

مدل را به صورت،

$$Y_{ij} = \alpha_i + \beta_j + \epsilon_{ij}\,,\ i = ۱,\ldots,I,\ \ j = ۱,\ldots,J, \qquad (۸)$$

j/i	۱	۲	۳	۴	۵
۱	۰/۹۸۲	۱/۹۰۲	۳/۷۹۷	−۱/۵۳۱	۰/۵۷۰
۲	−۱/۴۱۷	۱/۳۵۶	۱/۲۸۷	−۳/۶۲۹	−۳/۴۱۳
۳	−۱/۶۰۱	۴/۷۱۳	۰/۸۱۴	۰/۸۳۴	−۲/۰۸۲
۴	−۴/۹۱۲	−۴/۵۴۱	−۴/۷۶۸	−۹/۰۵۱	−۲/۷۴۴

جدولِ ۴: داده‌های طرحِ دو طرفه

در نظر بگیرید که در آن ϵ_{ij}ها مستقل از هم و دارای توزیع $N(0,\sigma^2)$ هستند و دانش پیشین در مورد محدودیت پارامترهای خطّی این است که α_iها کاهشی و β_jها تا سطح نامعلوم t افزایشی و سپس کاهشی هستند. جدول ۴ مجموعه‌ای از داده‌های $\boldsymbol{y}$ را نشان می‌دهد که برطبق مدل (۸) با $\alpha_1 = 2$ ، $\alpha_2 = 1$ ، $\alpha_3 = 0$ ، $\alpha_4 = -2$ ، $\beta_1 = -1$ ، $\beta_2 = 0$ ، $\beta_3 = 2$ ، $\beta_4 = -1$ ، $\beta_5 = -2$ و $\sigma^2 = 3$ تولید شده‌اند. لذا برای هر ستون، امید خانه‌ها کاهشی و برای هر سطر امید تا خانه وسط افزایشی و سپس کاهشی است. داده‌ها نسبتاً شلوغ و اغلب مغایر با این امیدها هستند. نتیجه تحلیل به روش کمترین توانهای دوّم عادی ([1]OLS) با نادیده گرفتن قیدهای ترتیبی، به طور وحشتناک گمراه کننده است،

$$\hat{\alpha}_1 = 1*064,\ \hat{\alpha}_2 = -1*163,\ \hat{\alpha}_3 = 0*536,\ \hat{\alpha}_4 = -5*203,$$

$$\hat{\beta}_1 = -1*737,\ \hat{\beta}_2 = 0*758,\ \hat{\beta}_3 = 0*283,\ \hat{\beta}_4 = -3*344,\ \hat{\beta}_5 = -1*917,$$

$$\hat{\sigma}^2 = 3*59\,,$$

این تحلیل منجر به برآوردهایی برای α_i و β_j می‌شود که در قیدها صدق نمی‌کنند و اغلب دور از مقادیر واقعی هستند. یک روش کلاسیک بهتر، راه حل کمترین توانهای دوّم مقیّد (رگرسیون همنوا) است. تحلیل بیزی بر اساس استفاده از نمونه‌گیر گیبز در این حالت به آسانی انجام می‌شود. برای سادگی، پیشین نرمال مزدوج را برای α_i و β_j و گامای وارون را برای σ^2 در نظر می‌گیریم. یعنی با چشم‌پوشی از قیدها، α_iها مستقل از هم و دارای توزیع مشترک $N(0,\sigma^2_\alpha)$ و β_jها مستقل از هم و دارای توزیع مشترک $N(0,\sigma^2_\beta)$ هستند. (برای راحتی مرکز پیشینها را صفر قرار می‌دهیم و این انتخاب یعنی اینکه مرکز α_iها و β_jها تقریباً در صفر است.) σ^2 مستقل از α_iها و β_jها و دارای توزیع $IG(a,b)$ است. با قرار دادن

[1]Ordinary Least Squares

$\sigma_\alpha^2 = 5$، $\sigma_\beta^2 = 5$، $a = 0$ و $b = 1$، این پیشینها را مبهم‌تر می‌سازیم. توزیعهای حاشیه‌ای کامل با استفاده از صورت کلی σ_α^2، σ_β^2، a، b و وارد کردن قیدهای ترتیبی در مدل، عبارتند از:

$$[\alpha_i|\boldsymbol{Y}, \boldsymbol{\beta}, \sigma^2, \alpha_r; r \neq i] = N\left(\frac{5\sigma_\alpha^2(\bar{y}_{i\circ} - \beta_\circ)}{5\sigma_\alpha^2 + \sigma^2}, \frac{\sigma_\alpha^2\sigma^2}{5\sigma_\alpha^2 + \sigma^2}\right),\ i = 1, 2, 3, 4,$$

که به بازهٔ $(\alpha_{i-1}, \alpha_{i+1})$ مقیّد است، $\alpha_\circ \equiv -\infty$، $\alpha_5 \equiv \infty$، $\bar{y}_{i\circ} = \frac{1}{5}\sum_{j=1}^{5} y_{ij}$ و $\beta_\circ = \frac{1}{5}\sum_{j=1}^{5} \beta_j$،

$$[\beta_j|\boldsymbol{Y}, \boldsymbol{\alpha}, \sigma^2, \beta_s; s \neq j] = N\left(\frac{4\sigma_\beta^2(\bar{y}_{\circ j} - \alpha_\circ)}{4\sigma_\beta^2 + \sigma^2}, \frac{\sigma_\beta^2\sigma^2}{4\sigma_\beta^2 + \sigma^2}\right),\ j = 1, 2, 3, 4, 5,$$

که برای $j = 1, 2$ به بازهٔ $(\beta_{j-1}, \beta_{j+1})$، برای $j = 4, 5$ به بازهٔ $(\beta_{j-1}, \beta_{j+1})$ و برای $j = 3$ به بازهٔ $[\max(\beta_2, \beta_4), \infty)$ مقیّد است، $\beta_\circ \equiv -\infty$، $\beta_6 \equiv \infty$، $\bar{y}_{\circ j} = \frac{1}{4}\sum_{i=1}^{4} y_{ij}$، $\alpha_\circ = \frac{1}{4}\sum_{i=1}^{4} \alpha_i$ و بالاخره،

$$[\sigma^2|\boldsymbol{Y}, \boldsymbol{\alpha}, \boldsymbol{\beta}] = IG\Big(a + 10, b + \frac{1}{2}\sum_i\sum_j (y_{ij} - \alpha_i - \beta_j)^2\Big).$$

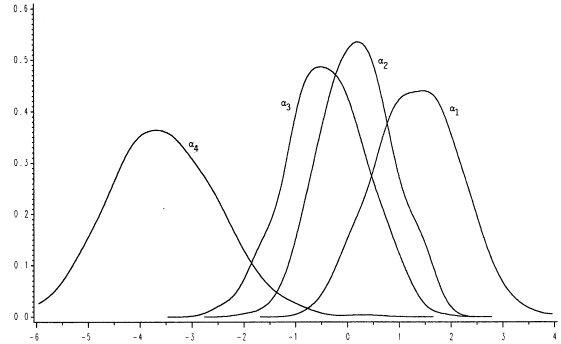
شکل ۲ توزیعهای پسینی حاشیه‌ای برای α_iها را به عنوان خروجی اجرای نمونه‌گیر گیبز نشان می‌دهد. متأثر بودن آنها از قیدهای ترتیبی کاملاً مشهود است. نمودار پسین برای β_jها و σ^2 را به طور مشابه می‌توان ترسیم کرد. مدهای پسینی حاشیه‌ای عبارتند از:

$$\alpha_1^* = 1.480,\ \alpha_2^* = 0.197,\ \alpha_3^* = -0.507,\ \alpha_4^* = -3.684,$$

$$\beta_1^* = -1.039,\ \beta_2^* = 0.635,\ \beta_3^* = 1.261,\ \beta_4^* = -1.149,\ \beta_5^* = -1.790,$$

$$\sigma^{*2} = 3.975,$$

که در محدودیتها صدق می‌کنند و به طور کلی نزدیکتر به مقادیر واقعی نسبت به برآوردهای OLS هستند.



شکلِ ۲: طرح دو طرفه با پارامترهای مرتب، توزیعهای پسین حاشیه‌ای برای اثرات اوّلین عامل.

۵ نتیجه‌گیری

هدف ما معرفی و توصیف استنباط بیزی مقیّد و نشان دادن این مطلب بوده است که چگونه استفاده از نمونه‌گیر گیبز در تحلیل بیزی میدان وسیعی از مسائل پارامترهای مقید به سادگی و به صورت سرراست می‌تواند انجام گیرد. این روش از انتگرال‌گیریهای عددی، اغلب غیر ممکن، بر روی مجموعه‌های با بعدهای بالا جلوگیری می‌کند که توسط قیدهای پیچیده تعریف می‌شوند. برای انجام این روش تنها نیاز به نمونه‌گیری از توزیعهای کامل شرطی تک متغیره داریم که محدود به یک زیر مجموعهٔ ساده از $\mathcal{R}^1$ شده‌اند. همچنین در دو مثال نشان دادیم که چگونه استنباط می‌توان به مقدار زیادی قوی‌تر شود در صورتی که قیدها موجود، در مدل‌بندی مسئله وارد شوند.

مراجع

[1] Carlin, B. P., Gelfand, A. E. and Smith, A. F. M. (1992). Hierarchical Bayesian Analysis of Change Point Problems. *Applied Statistics*, **41**, 389-405.

[2] Gelfand, A. E. and Smith, A. F. M. (1990). Sampling Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, **85**, 398-409.

[3] Gelfand, A. E. and Smith, A. F. M. (1991). Gibbs Sampling for Marginal Posterior Expectations. *Communications in Statistics B*, **20**, 1747-1766.

[4] Gelfand, A. E., Smith, A. F. M. and Lee, T. M. (1992). Bayesian Analysis of Constrained Parameter and Truncated Data Problems Using Gibbs Sampling. *Journal of the American Statistical Association*, **87**, 523-532.

[5] Geman, D. and Geman, S. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Transacions on Pattern Analysis and Machine Intelligence*, **6**, 721-741.

[6] Sedransk, J., Monahan, J. and Chiu, H. Y. (1985). Bayesian Estimation of Finite Population Parameters in Categorical Data Models Incorporating Order Restrictions. *Journal of the Royal Statistical Society*, **Ser.B, 47**, 519-527.

[7] Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.