

تحلیل بیزی با استفاده از رایانه

پرویز نصیری

دانشگاه پیام نور

چکیده. تحلیل بیز با انتخاب یک توزیع پیشین که معمولاً از دانش قبلی ما درباره‌ی پارامتر به دست می‌آید، شروع می‌شود. توزیع مناسب روی فضای پارامتر انتخاب و بر اساس آن توزیع پسین به صورت قطعی یا تقریبی برآورد می‌گردد. در این مقاله روشی تقریبی مبتنی بر رایانه معرفی می‌گردد که در آن نمونه‌ای به حجم بزرگ از توزیع پیشین تولید و با استفاده از آن و تابع درستنمایی و با تکرار آن، نمونه‌ای از توزیع پسین برای برآورد پارامتر(های) توزیع پسین تولید می‌گردد. گاهی با توجه به فرم تابع چگالی توأم دو یا چند متغیره از روش تقریبی نمونه‌گیری گیبس نیز استفاده می‌شود. در این مقاله این روش‌های تقریبی مرور و با چندین مثال به کار گرفته می‌شوند.

۱- مقدمه

اسمیت و گلفند [۴] دو روش برای تولید نمونه‌ی تصادفی از توزیع پسین ارائه دادند. در این روش‌ها نمونه‌های بزرگی از توزیع پیشین $\pi(\theta)$ با استفاده از رایانه تولید و با استفاده از آن‌ها و تابع درستنمایی نمونه‌ای از توزیع پسین تقریبی، برای توزیع واقعی پسین $\pi(\theta|x)$ ، تولید می‌شود. نمونه‌گیری گیبس یکی دیگر از شیوه‌هایی است که با استفاده از رایانه برای کامل کردن تحلیل توزیع پسین در یک موقعیت خاص مورد استفاده قرار می‌گیرد.

در این مقاله در بخش اول روش‌های تولید نمونه‌ی تصادفی از توزیع پسین بحث و در انتها مثال‌هایی جهت معرفی بهتر روش‌ها ارائه خواهد شد. در بخش دوم روش

واژگان کلیدی: توزیع پیشین؛ توزیع پسین؛ تحلیل بیز؛ نمونه‌گیری گیبس.

دریافت: ۱۳۸۷/۸/۲۷، پذیرش: ۱۳۸۸/۷/۵

نمونه‌گیری گیبس معرفی و با استفاده از آن برآورد تابع چگالی توزیع پسین ارائه می‌شود.

۲- تولید نمونه از توزیع پسین

در این بخش دو روش برای تولید نمونه از توزیع پسین ارائه می‌شود و در انتها نمونه‌هایی از توزیع مفروض تولید خواهد شد.

۲-۱- روش اول

فرض کنید f یک تابع انتگرال‌پذیر غیرمنفی باشد، به طوری که

$$f(x) \propto h(x)$$

که در آن $h(x)$ یک تابع چگالی احتمال است (یعنی $\int_{-\infty}^{\infty} h(x) dx = 1$). تابع چگالی

احتمال $g(x)$ را با شرایط زیر در نظر می‌گیریم.

الف- تکیه‌گاه $g(x)$ شامل تکیه‌گاه $f(x)$ باشد.

ب- تولید مشاهدات از $g(x)$ با استفاده از رایانه آسان باشد.

پ- یک عدد معلوم M وجود داشته باشد به طوری که $f(x) \leq M$.

برای تولید مشاهدات، الگوریتم زیر را در نظر می‌گیریم.

۱- یک مشاهده‌ی x از $g(x)$ و مشاهده‌ی u مستقل از x از توزیع یکنواخت

روی بازه‌ی $(0, 1)$ تولید می‌کنیم.

۲- اگر $u \leq \frac{f(x)}{Mg(x)}$ باشد، مشاهده‌ی x به‌عنوان یک مشاهده انتخاب و ذخیره

می‌شود. در غیر این صورت زوج (x, u) نادیده گرفته شده و مجدداً از

مرحله‌ی ۱ تکرار می‌شود. بعد از تکرار مراحل ۱ و ۲ می‌توان نمونه‌ای به حجم

n تولید نمود. با استفاده از استدلال زیر می‌توان نشان داد که اگر Y یک

مشاهده‌ی انتخاب شده باشد تابع چگالی آن $h(\cdot)$ خواهد بود [۳].

$$\begin{aligned}
 P(Y \leq y) &= P\left[X \leq y \mid u \leq \frac{f(x)}{Mg(x)}\right] \\
 &= \frac{P\left[X \leq y, u \leq \frac{f(x)}{Mg(x)}\right]}{P\left[u \leq \frac{f(x)}{Mg(x)}\right]} \\
 &= \frac{\int_{-\infty}^y P\left[u \leq \frac{f(x)}{Mg(x)}\right] g(x) dx}{\int_{-\infty}^{\infty} P\left[u \leq \frac{f(x)}{Mg(x)}\right] g(x) dx}
 \end{aligned}$$

حال از آن جا که متغیر تصادفی u یکنواخت است، داریم

$$\begin{aligned}
 P(Y \leq y) &= \frac{\int_{-\infty}^y \frac{f(x)}{Mg(x)} g(x) dx}{\int_{-\infty}^{\infty} \frac{f(x)}{Mg(x)} g(x) dx} \\
 &= \frac{\int_{-\infty}^y f(x) dx}{\int_{-\infty}^{\infty} f(x) dx} = \int_{-\infty}^y h(x) dx
 \end{aligned}$$

بنا بر این Y دارای تابع چگالی احتمال $h(x)$ است.

برای تحلیل بیزی با توزیع پیشین $\pi(\theta)$ و تابع درستنمایی $L(x|\theta)$ توابع $f(\theta)$ و $g(\theta)$ را به صورت زیر در نظر می گیریم.

$$f(\theta) = L(x|\theta) \pi(\theta), \quad g(\theta) = \pi(\theta)$$

اگر $\hat{\theta}$ یک برآوردگر ماکسیمم درستنمایی باشد،

$$f(\theta) = L(x|\theta) \pi(\theta) \leq L(x|\hat{\theta}) g(\theta).$$

بنا بر این با اختیار $M = L(x|\hat{\theta})$ ،

$$\frac{f(\theta)}{Mg(\theta)} = \frac{f(x|\theta)\pi(\theta)}{f(x|\hat{\theta})\pi(\theta)} = \frac{f(x|\theta)}{f(x|\hat{\theta})}$$

برای استفاده از الگوریتم بالا یک مشاهده از $\pi(\theta)$ تولید و مقدار $\frac{f(x|\theta)}{f(x|\hat{\theta})}$ را محاسبه کرده و آن را با یک مشاهده‌ی مستقل از توزیع یکنواخت روی $(0,1)$ مقایسه و نسبت به قبول یا رد زوج (θ, u) تصمیم‌گیری می‌کنیم. با تکرار مراحل اخیر می‌توان یک نمونه به حجم دلخواه از توزیع پسین تولید کرد.

مثال: فرض کنید متغیر تصادفی X دارای تابع چگالی برنولی با پارامتر θ باشد و توزیع پیشین $0 < \theta < 1$ ، $\pi(\theta) = 2\theta$ در نظر گرفته شود. بنا بر این

$$f(x|\theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1, \quad 0 < \theta < 1$$

اگر x_1, x_2, \dots, x_n نمونه‌ی مشاهده شده باشد

$$L(x|\theta) = \theta^{n\bar{x}} (1-\theta)^{n-n\bar{x}}$$

$$L(x|\hat{\theta}) = (\hat{\theta})^{n\bar{x}} (1-\hat{\theta})^{n-n\bar{x}}, \quad \hat{\theta} = \bar{X}$$

یک مشاهده از $\pi(\theta) = 2\theta$ با استفاده از تابع توزیع $F(\theta) = \theta^x$ تولید می‌کنیم. از آنجا که $0 \leq F(\theta) \leq 1$ است

$$F(\theta) = \theta^x = u \Rightarrow \theta = \sqrt{x}.$$

با فرض این که عدد تولیدشده‌ی بین صفر و یک برابر با 0.4 باشد، نتیجه می‌شود $\theta = \sqrt{0.4} = 0.2$ در این صورت

$$f(x|\theta) = f(x|0.2) = (0.2)^x (1-0.2)^{1-x}, \quad x = 0, 1$$

یک نمونه‌ی بیست تایی از $f(x|0.2)$ می‌توان به صورت زیر تولید نمود.

○ ○ ۱ ۱ ۱ ○ ○ ۱ ○ ○ ○ ○ ۱ ۱ ○ ○ ○ ○ ○ ۱

$$\hat{\theta} = \bar{X} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{7}{20} = 0.35$$

$$L(x|0.2) = \theta^{n\bar{x}} (1-\theta)^{n-n\bar{x}} = (0.2)^{7 \cdot (0.35)} (1-0.2)^{20-7 \cdot (0.35)} \\ = (0.2)^7 (0.8)^{13}$$

$$L(x|\hat{\theta}) = L(x|0.35) = (0.35)^y (1-0.35)^{12}$$

$$\frac{L(x|\theta)}{L(x|\hat{\theta})} = \left(\frac{0.2}{0.35}\right)^y \left(\frac{0.8}{0.65}\right)^{12} = 0.2958$$

حال یک مشاهده به تصادف از توزیع یکنواخت روی بازه‌ی $(0,1)$ تولید می‌کنیم. فرض کنید $u = 0.1670$ مشاهده شده باشد. چون $u < \frac{L(x|\theta)}{L(x|\hat{\theta})} = 0.2958$ است زوج $(\theta, u) = (0.2, 0.1670)$ انتخاب می‌شود بنا بر این $\theta = 0.2$ به‌عنوان نمونه از توزیع پسین انتخاب می‌شود.

حال فرض کنید برای $\theta = 0.2$ نمونه‌ی مشاهده شده به حجم ۲۰ از $f(x|0.2)$ به‌صورت زیر باشد.

○ ۱ ○ ۱ ○ ○ ۱ ○ ○ ۱ ○ ○ ۱ ○ ○ ○ ۱ ○ ○

$$\hat{\theta} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{6}{20} = 0.3$$

$$L(x|0.2) = \theta^{n\bar{x}} (1-\theta)^{n-n\bar{x}} = (0.2)^6 (1-0.2)^{14} = (0.2)^6 (0.8)^{14}$$

$$L(x|\hat{\theta}) = \hat{\theta}^{n\bar{x}} (1-\hat{\theta})^{n-n\bar{x}} = (0.3)^6 (1-0.3)^{14} = (0.3)^6 (0.7)^{14}$$

$$\frac{L(x|0.2)}{L(x|\hat{\theta}=0.3)} = \frac{(0.2)^6 (0.8)^{14}}{(0.3)^6 (0.7)^{14}} = \left(\frac{2}{3}\right)^6 \left(\frac{8}{7}\right)^{14} = 0.5693$$

اگر مشاهده‌ی نمونه‌ی تصادفی از بازه‌ی $(0,1)$ برابر با 0.5696 باشد زوج $(0.2, 0.5693)$ انتخاب نمی‌شود.

از این روش وقتی استفاده می‌شود که θ به‌روش ماکسیمم درست‌نمایی قابل برآورد باشد در غیر این صورت از روش دوم که در ذیل بحث خواهد شد، استفاده می‌شود.

۲-۲- روش دوم

در این روش یک نمونه به حجم n (بزرگ) از توزیع پیشین تولید می‌شود. اگر $\theta_1, \theta_2, \dots, \theta_n$ نمونه‌ی تولید شده باشد، متغیر W_i به صورت زیر تعریف می‌شود.

$$W_i = \frac{L(x|\theta_i)}{\sum_{i=1}^n L(x|\theta_i)} = \frac{\prod_{j=1}^n f(x_j|\theta_i)}{\sum_{i=1}^n \prod_{j=1}^n f(x_j|\theta_i)}$$

اگر θ^* یک متغیر تصادفی گسسته باشد و مقادیر $\theta_1, \theta_2, \dots, \theta_n$ را به ترتیب با احتمالات W_1, W_2, \dots, W_n اختیار کند آنگاه

θ_0	θ_1	θ_2	...	θ_n
$P(\theta^* = \theta_0)$	W_1	W_2	...	W_n

9

$$\lim_{n \rightarrow \infty} P(\theta^* \leq \theta_0) = \frac{\int_{-\infty}^{\theta_0} L(x|\theta) \pi(\theta) d\theta}{\int_{-\infty}^{\infty} L(x|\theta) \pi(\theta) d\theta}$$

برای n بزرگ، توزیع θ^* یک توزیع تقریبی برای توزیع پسین است.
اثبات:

$$\begin{aligned} (1) \quad P(\theta^* \leq \theta_0) &= \sum_{\theta_i \leq \theta_0} \frac{L(x|\theta_i)}{\sum L(x|\theta_i)} \\ &= \frac{\frac{1}{n} \sum_{i=1}^n L(x|\theta_i) I_{[\theta_i \leq \theta_0]}^{(\theta_i)}}{\frac{1}{n} \sum_{i=1}^n L(x|\theta_i)} \end{aligned}$$

چون θ_i ها از هم مستقل و هم توزیع‌اند (iid) با استفاده از قانون قوی اعداد بزرگ برای صورت و منخرج کسر اخیر داریم

$$\lim_{n \rightarrow \infty} P(\theta^* \leq \theta_0) = \frac{\int_{-\infty}^{\theta_0} L(x|\theta) \pi(\theta) d\theta}{\int_{-\infty}^{\infty} L(x|\theta) \pi(\theta) d\theta}$$

با تولید مشاهدات از توزیع θ^* ، یک نمونه از توزیع پسین به دست می‌آید.
 مثال: برای مثال قبلی، یک نمونه به حجم $n=10$ از $0 < \theta < 1$ ، $\pi(\theta) = 2\theta$ تولید می‌کنیم. فرض کنید نمونه‌ی تولید شده به صورت زیر باشد.

۰/۲ ۰/۱ ۰/۰۵ ۰/۴ ۰/۲۵ ۰/۶ ۰/۷ ۰/۳ ۰/۸ ۰/۹

تابع توزیع θ^* با توجه به رابطه‌ی (۱) برابر است با

$$G(\theta^*) = P(\theta^* \leq \theta_0) = \frac{\sum_{i=1}^n L(x_i|\theta_i) I_{[\theta_i \leq \theta_0]}^{(\theta_i)}}{\sum_{i=1}^n L(x_i|\theta_i)}$$

به ازای مقادیر مختلف θ نمونه‌ای به حجم ۱۰ از $f(x|\theta)$ تولید می‌شود (جدول ۱).

جدول ۱- نمونه‌های به حجم ده به ازای مقادیر مختلف θ از $f(x|\theta)$

X_i	θ_i									
	۰/۰۵	۰/۱	۰/۲	۰/۲۵	۰/۳	۰/۴	۰/۶	۰/۷	۰/۸	۰/۹
x_1	۰	۰	۰	۰	۰	۰	۱	۱	۱	۱
x_2	۰	۰	۰	۰	۱	۱	۰	۱	۱	۱
x_3	۰	۰	۱	۰	۰	۰	۱	۰	۱	۱
x_4	۰	۰	۰	۰	۰	۰	۰	۱	۱	۱
x_5	۰	۱	۰	۱	۱	۰	۱	۱	۰	۱
x_6	۰	۰	۰	۰	۱	۰	۱	۱	۱	۱
x_7	۰	۰	۱	۱	۰	۰	۰	۱	۱	۱
x_8	۰	۰	۰	۰	۰	۱	۱	۱	۱	۱
x_9	۰	۰	۰	۱	۰	۰	۱	۱	۱	۱
x_{10}	۱	۱	۰	۰	۰	۱	۰	۱	۱	۰
$\sum X_i$	۱	۲	۲	۳	۳	۳	۶	۹	۹	۹

$$L(x|\theta_i) = \prod_{j=1}^n f(x_j|\theta_i) = \theta_i^{\sum_{j=1}^n x_j} (1-\theta_i)^{n-\sum_{j=1}^n x_j}$$

$$= \theta_i^{\sum_{j=1}^n x_j} (1-\theta_i)^{1-\sum_{j=1}^n x_j}, \quad i = 1, 2, \dots, 10.$$

$$L(x|\theta_1) = L(x|\theta_1 = 0.5) = (0.5)^1 (0.5)^9 = 0.315$$

$$L(x|\theta_2) = L(x|\theta_2 = 0.1) = (0.1)^7 (0.9)^3 = 0.043$$

$$L(x|\theta_3) = L(x|\theta_3 = 0.2) = (0.2)^7 (0.8)^3 = 0.067$$

$$L(x|\theta_4) = L(x|\theta_4 = 0.25) = (0.25)^7 (0.75)^3 = 0.208$$

$$L(x|\theta_5) = L(x|\theta_5 = 0.3) = (0.3)^7 (0.7)^3 = 0.022$$

$$L(x|\theta_6) = L(x|\theta_6 = 0.4) = (0.4)^7 (0.6)^3 = 0.017$$

$$L(x|\theta_7) = L(x|\theta_7 = 0.6) = (0.6)^7 (0.4)^3 = 0.011$$

$$L(x|\theta_8) = L(x|\theta_8 = 0.7) = (0.7)^9 (0.3)^1 = 0.012$$

$$L(x|\theta_9) = L(x|\theta_9 = 0.8) = (0.8)^9 (0.2)^1 = 0.268$$

$$L(x|\theta_{10}) = L(x|\theta_{10} = 0.9) = (0.9)^9 (0.1)^1 = 0.38$$

θ_i	0.5	0.1	0.2	0.25	0.3	0.4	0.6	0.7	0.8	0.9
$L(x \theta_i)$	0.315	0.043	0.067	0.208	0.022	0.017	0.011	0.012	0.268	0.38

$$\sum_{i=1}^n L(x|\theta_i) = 0.1451 \quad P(\theta^* \leq \theta_0) = \frac{\sum_{i=1}^n L(x|\theta_i) I_{[\theta_i \leq \theta_0]}^{(\theta_i)}}{\sum_{i=1}^n L(x|\theta_i)}$$

θ_0	0.5	0.1	0.2	0.25	0.3	0.4	0.6	0.7	0.8	0.9
$P(\theta^* \leq \theta_0)$	0.2171	0.2467	0.2929	0.4363	0.45	0.4618	0.4693	0.5520	0.7367	1

با توجه به رابطه‌ی تابع توزیع و تابع چگالی احتمال برای متغیر گسسته می‌توان نوشت

$$g(\theta^*) = G(\theta^*) - G(\theta^* -)$$

بنا بر این تابع چگالی پسین $g(\theta(\underline{X}))$ برابر است با

θ^*	۰/۰۵	۰/۱	۰/۲	۰/۲۵	۰/۳	۰/۴	۰/۶	۰/۷	۰/۸	۰/۹
$g(\theta^*)$	۰/۲۱۷۱	۰/۰۲۹۶	۰/۰۴۶۲	۰/۱۴۳۴	۰/۰۱۳۷	۰/۰۱۱۸	۰/۰۰۷۵	۰/۰۸۲۷	۰/۱۸۴۷	۰/۲۶۳۳

مثال: اگر متغیر تصادفی X دارای تابع چگالی احتمال دوجمله‌ای با پارامتر θ

$$f(x|\theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \quad x = 0, 1, 2, \dots, n$$

و تابع چگالی پیشین بتا با پارامترهای $\alpha = 3$ و $\beta = 2$ باشد،

$$\pi(\theta) = \frac{\pi(\alpha + \beta)}{\pi(\alpha)\pi(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} = 3\theta^2 (1-\theta), \quad 0 < \theta < 1$$

با تولید نمونه‌ای به حجم ۷ از $\pi(\theta)$ و نمونه‌ای به حجم ۱۰ از $f(x|\theta)$ تابع چگالی احتمال پسین را به دست آورید.

نمونه‌ی تولید شده از $\pi(\theta)$ عبارت است از

$$0/751574 \quad 0/618027 \quad 0/529509 \quad 0/892439 \quad 0/390308 \quad 0/904114 \quad 0/951667$$

به ازای هر θ تولید شده یک نمونه به حجم ۱۰ از $f(x|\theta)$ تولید شده است (جدول ۲).

با توجه به نمونه‌های تولید شده $L(x|\theta_i) = \prod_{j=1}^n f(x_j|\theta_i)$ را محاسبه می‌کنیم

$$L(x|\theta_i) = \prod_{j=1}^n f(x_j|\theta_i) = \left\{ \prod_{j=1}^n \binom{n}{x_j} \theta_i^{x_j} (1-\theta_i)^{n-x_j} \right\} \\ = \left\{ \prod_{j=1}^n \binom{n}{x_j} \right\} \theta_i^{\sum x_j} (1-\theta_i)^{\sum (n-x_j)}$$

$$= \left\{ \prod_{j=1}^{10} \binom{10}{x_j} \right\} \theta_i^{\sum x_j} (1-\theta_i)^{\sum (10-x_j)}$$

$$= \left\{ \prod_{j=1}^{10} \binom{10}{x_j} \right\} \theta_i^{\sum x_j} (1-\theta_i)^{100-\sum x_j}$$

جدول ۲- نمونه‌های به حجم ده به ازای مقادیر مختلف θ از $f(x|\theta)$

X_i	θ_i						
	۰/۷۵۱۵۷۴	۰/۶۱۸۰۲۷	۰/۵۲۹۵۰۹	۰/۱۹۲۴۳۹	۰/۳۹۰۳۰۸	۰/۹۰۴۱۱۴	۰/۹۵۱۶۶۷
x_1	۷	۴	۵	۸	۴	۱۰	۱۰
x_2	۹	۶	۵	۱۰	۳	۹	۸
x_3	۵	۷	۷	۱۰	۲	۱۰	۸
x_4	۹	۹	۴	۹	۶	۸	۹
x_5	۵	۴	۵	۱۰	۴	۸	۱۰
x_6	۷	۶	۷	۹	۴	۹	۹
x_7	۹	۸	۵	۹	۶	۱۰	۱۰
x_8	۷	۵	۴	۸	۴	۹	۱۰
x_9	۷	۷	۵	۹	۳	۷	۱۰
x_{10}	۶	۶	۵	۹	۲	۱۰	۱۰
$\sum_{i=1}^{10} x_i$	۷۱	۶۲	۵۲	۹۱	۳۸	۹۰	۹۴

$$L(x|\theta_1) = \left\{ \prod_{j=1}^{10} \binom{10}{x_j} \right\} \theta_1^{\sum x_j} (1-\theta_1)^{100-\sum x_j}$$

$$= [120 \times 10 \times 252 \times 10 \times 252 \times 120 \times 10 \times 120 \times 120 \times 210] \theta_1^{91} (1-\theta_1)^{100-91}$$

$$= 2/765319 \times 10^{-18} (0/751574)^{91} (1-0/751574)^{99}$$

$$= 1/248676 \times 10^{-8}$$

$$L(x|\theta_2) = 9/633301 \times 10^{-9} \qquad L(x|\theta_3) = 1/364839 \times 10^{-7}$$

$$L(x|\theta_4) = 0/012416 \qquad L(x|\theta_5) = 3/535389 \times 10^{-8}$$

$$L(x | \theta_x) = 1/833351 \times 10^{-6} \quad L(x | \theta_y) = 2/451762 \times 10^{-5}$$

$$G(\theta^*) = \frac{\sum_{i=1}^y L(x | \theta_i) I_{(\theta_i \leq \theta_0)}^{(\theta_i)}}{\sum_{i=1}^y L(x | \theta_i)}$$

θ	۰/۷۵۱۵۷۴	۰/۶۱۸۰۲۷	۰/۵۲۹۵۰۹	۰/۸۹۲۴۳۹	۰/۳۹۰۳۰۸	۰/۹۰۴۱۱۴	۰/۹۵۱۶۶۷
$G(\theta)$	$1/0.04 \times 10^{-7}$	$1/775 \times 10^{-6}$	$1/274 \times 10^{-5}$	۰/۹۹۷۰	۰/۹۹۷۵	۰/۹۹۸۰	۱

θ	۰/۷۵۱۵۷۴	۰/۶۱۸۰۲۷	۰/۵۲۹۵۰۹	۰/۸۹۲۴۳۹	۰/۳۹۰۳۰۸	۰/۹۰۴۱۱۴	۰/۹۵۱۶۶۷
$g(\theta)$	$1/0.04 \times 10^{-7}$	$1/67562 \times 10^{-6}$	$1/0.965 \times 10^{-5}$	۰/۹۹۶۹	۰/۰۰۱	۰/۰۰۰۵	۰/۰۰۲

۳- نمونه‌گیری گیبس و استفاده از آن در تحلیل بیزی

پیشرفت و تکامل روز به روز رایانه و دسترسی آسان به رایانه‌های با سرعت بالا، باعث شده است که روش‌های الگوریتمی، نظیر نمونه‌گیری گیبس، ابزارهای مفید و ضروری برای حل مسائل آماری باشند.

نمونه‌گیری گیبس یک روش الگوریتمی است که ایده‌ی اولیه‌ی آن توسط متروپولیس و همکاران در سال ۱۹۵۳ ارائه گردید و سپس با مقاله‌ای که گیمان و گیمان در مورد مدل‌های پردازش تصویر ارائه دادند، وارد مرحله‌ی جدیدی شد [۱]. در این بخش ابتدا نمونه‌گیری گیبس را شرح می‌دهیم و با استفاده از آن برای برآورد پارامتر (پارامترها)، توزیع آن را برآورد می‌کنیم.

۳-۱- نمونه‌گیری گیبس

فرض کنید متغیرهای تصادفی X و Y دارای تابع چگالی توأم $f(x, y)$ بوده، به طوری که $(x, y) \in \Omega = \{(x, y) | x \in A, y \in B\}$ و توابع چگالی شرطی آن‌ها $f(x|y)$ و $f(y|x)$ باشند. هدف به دست آوردن تابع چگالی حاشیه‌ای X (یا Y)

است. محاسبه‌ی انتگرال روی فضای نمونه‌ی متغیر Y ممکن است به راحتی امکان پذیر نباشد یا به عبارت دیگر محاسبه‌ی انتگرال زیر آسان نباشد.

$$f(x) = \int_{-\infty}^{\infty} f(x|y) f(y) dy = \int_{-\infty}^{\infty} f(x, y) dy$$

با توجه به معلوم بودن فرم توابع چگالی شرطی X به شرط $Y = y$ و Y به شرط $X = x$ برای تولید یک نمونه به حجم n (بزرگ) الگوریتم زیر را اجرا می‌کنیم.

- ۱- نمونه‌ی اختیاری (x^0, y^0) را از فضای نمونه‌ی Ω انتخاب می‌کنیم.
- ۲- با استفاده از $f(y|x)$ و x^0 ، مشاهده‌ی y^1 را از $f(y|x^0)$ تولید می‌کنیم.
- ۳- با استفاده از (x^0, y^1) و (x^0, y^0) ، مشاهده‌ی x^1 را از $f(x|y^1)$ تولید می‌کنیم.
- ۴- با جایگزینی (x^1, y^1) به جای (x^0, y^0) مراحل ۱ تا ۳ را m مرتبه برای به دست آوردن (x^m, y^m) تکرار می‌کنیم.
- ۵- برای تولید نمونه‌ی (x_i^m, y_i^m) ، $i = 1, 2, \dots, n$ ، مرتبه‌ی مراحل ۱ تا ۵ را تکرار می‌کنیم.

فرایند اخیر تولید نمونه را، نمونه‌گیری گیبس گویند.

۲-۳- برآورد تابع چگالی حاشیه‌ای X

اگر $f(x^*)$ تابع چگالی حاشیه‌ای X در نقطه‌ی x^* باشد، با توجه به روابط بین توابع چگالی حاشیه‌ای و شرطی می‌توان نوشت

$$\begin{aligned} f(x^*) &= \int f(x^*|y) f(y) dy \\ &= E[f(X^*|y)]. \end{aligned}$$

براساس مقادیر مشاهده شده‌ی Y_i^m برآوردگر $f(x^*)$ برابر است

با

$$f(x^*) = \frac{1}{n} \sum_{i=1}^n f(x^* | y_i^m)$$

$f(x^*)$ را براورگر تابع چگالی به‌روش نمونه‌گیری گیبس گویند. به دنبال آن $\widehat{E(X)}$ برابر است با

$$\widehat{E(X)} = \frac{1}{n} \sum_{i=1}^n E(X | Y_i^m)$$

روابط اخیر را می‌توان برای تابع چگالی حاشیه‌ای متغیر Y نیز نوشت. برای استفاده از تکنیک نمونه‌گیری گیبس در تحلیل بیزی، فرض کنید تابع چگالی متغیر تصادفی X به پارامتر θ بستگی داشته باشد به‌طوری که θ دارای تابع چگالی و توزیع پسین $\pi(\theta | X)$ فرم صریح از متغیرها نباشد. برای تولید نمونه‌ی تصادفی به‌روش نمونه‌گیری گیبس، فرض کنید متغیر تصادفی Y با شرایط زیر در دسترس باشد.

الف- فرم توزیع شرطی Y به شرط θ و X معلوم و صریح باشد، به‌طوری که بتوان نمونه‌ای از $g(Y | \theta, X)$ را تولید کرد.

ب- فرم توزیع شرطی θ به شرط Y و X معلوم و صریح باشد، به‌طوری که بتوان نمونه‌ای از $g(\theta | X, Y)$ تولید کرد.

با توجه به بند الف و ب توزیع پسین $\pi(\theta | X)$ از توزیع حاشیه‌ای $\pi(\theta, Y | X)$ به دست می‌آید و با استفاده از توابع چگالی

$$1) \pi(\theta | X, Y) \quad 2) \pi(Y | X, \theta)$$

نمونه‌ی تصادفی از $\pi(\theta | X)$ با استفاده از روش نمونه‌گیری گیبس تولید می‌شود.

مثال: فرض کنید متغیر تصادفی y دارای تابع چگالی احتمال دو جمله‌ای با پارامترهای x و θ باشد.

$$f(y | x, \theta) = \binom{x}{y} \theta^y (1-\theta)^{x-y}, \quad y = 0, 1, 2, \dots, x; \quad 0 < \theta < 1$$

برای $x=10$ و $\theta=0.2$ نمونه‌ی زیر از این توزیع تولید شده است.

۱ ۲ ۴ ۴ ۲ ۱ ۲ ۱ ۳ ۱ ۱: نمونه‌ی تولید شده

حال فرض کنید متغیر تصادفی θ دارای تابع چگالی احتمال بتا با پارامترهای x و y باشد؛

$$\pi(\theta|x, y) = \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)} \theta^{x-1}(1-\theta)^{y-1}, \quad 0 < \theta < 1$$

نمونه‌ای به حجم ۱۰ از این توزیع برای $x=10$ و $y=2$ به صورت زیر تولید شده است

نمونه‌ی تولیدشده: ۰/۳۸۹۳۸۰ ۰/۸۹۲۱۶۴ ۰/۹۱۰۶۲۹ ۰/۷۷۸۱۵۵ ۰/۸۰۸۴۹۹ ۰/۸۷۵۰۴۵
* * * *

علامت * نمایانگر آن است که نمونه‌ی تولیدشده رؤیت نشده است. با استفاده از روش نمونه‌گیری گیبس آن‌ها را تولید کنید.

برای تولید نمونه از $\pi(\theta|x, y)$ برای θ ‌های مختلف از $f(y|\theta, 10)$ چهار نمونه به صورت زیر تولید می‌کنیم

θ	۰/۱۸	۰/۱۹	۰/۲	۰/۲۱
y تولید شده	۲	۳	۲	۵

به ازای هر مقدار y تولید شده در جدول اخیر نمونه‌ای از $\pi(\theta|y, 10)$ به صورت زیر تولید می‌کنیم.

x	۱۰	۱۰	۱۰	۱۰
y	۲	۳	۲	۵
θ تولید شده	۰/۷۰۲۳۳۱	۰/۹۳۲۷۷۳	۰/۹۶۹۳۶۰	۰/۶۸۴۸۵۷

۴- نتیجه‌گیری

تکامل روز افزون رایانه و دسترسی آسان به آن، کاربرد تحلیل بیز در تحلیل داده‌های آماری و روش‌های تولید نمونه تصادفی از توزیع پسین در مواقعی که توزیع‌های حاشیه‌ای یا شرطی فرم صریحی ندارند، این موارد در این مقاله مورد بحث قرار گرفته و با ارائه‌ی مثال‌ها ضمن تولید نمونه از توزیع پسین، برآورد پارامترها به دست آمده است. در ادامه برای حل مسائل آماری روش‌های نمونه‌گیری گیبس که ابزار مفید و ضروری است ارائه

شده است. انتظار می‌رود از کاربرد روش‌ها در آمار کاربردی در مراکز مهم از جمله مرکز آمار مورد استفاده قرار گیرد.

سپاس‌گزاری

بدین وسیله از داوران مقاله که نظرات‌شان موجب ارتقاء کیفی مقاله گردید سپاس‌گزاری می‌شود.

مرجع‌ها

- [۱] زال زاده، سعید (۱۳۸۱). نمونه‌گیری گیبس، اندیشه‌ی آماری، سال هفتم، شماره‌ی اول.
- [2] Gaman, S.; Gaman D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intelligence*, **6**, 721-741.
- [3] Metropolis, N.; Rosenbluth A. W.; Rosenbluth M. N.; Teller A. H.; Teller E. (1953). Equations of State calculations by fant computing methods. *J. Chem. Phys.* **21**, 1087-1091.
- [4] Ripley, B. (1986). *Stochastic Simulation*, New York: John Wiley.
- [5] Smith, A. F. M.; Gelfand A. C. (1992). Bayesian Statistics Without Tears: A Sampling Re sampling Perspective, *The Amer. Statist.*, **46**, 84-88.

پرویز نصیری

تهران، دانشگاه پیام‌نور.
پیام‌نگار:

