فرض کنید $Y_1, \dots, Y_n$ متغیرهای تصادفی باشند که هر کدام به طور مستقل دارای توزیع آمیخته ۳ بعدی از نرمال هستند. یعنی

$$Y_1, \dots, Y_n \overset{iid}{\sim} \sum_{i=1}^{3} \pi_i N(\mu_i, \sigma_i^2)$$

تمامی پارامترها مجهول؛

برای بدست آوردن برآوردیابی پارامترهای ~~مجهول~~ توزیع، روش های مختلفی وجود دارد.

در اینجا می خواهیم روش برآوردیابی گیبس که بیشتر مورد استفاده قرار می گیرد را هم اشاره کنیم.

در اینجا برای هر متغیر تصادفی Y_i، یک دونه متغیر تصادفی پنهانی (latent Variable) W_i وجود داره بطوریکه با داشتن این متغیر تصادفی، توزیع متغیر تصادفی Y_i به طور کامل معلوم میشود. در توزیع ما این متغیر تصادفی پنهان وجود ندارد ولی ما این متغیر را به مجموعمون اضافه می کنیم که بتونیم راحت تر شبیه سازی گیبس و برآورد پارامترهامون را بدست بیاوریم.

منظور از اینکه با داشتن متغیر تصادفی W_i، توزیع Y_i به طور دقیق مشخص می شود یعنی اینکه

$$Y_i \mid W_i = j \sim N(\mu_j, \sigma_j^2)$$

نشان داده شده که اضافه کردن این متغیر تصادفی پنهانی به مدل تأثیری در همگرایی برآورد پارامترهای مدل ندارد.

خب هدف ما چیه؟!

هدف پیدا کردن برآورد پارامترهای توزیع آمیخته نرمال به روش بیزی هست.

پس باید ابتدا برای پارامترهای توزیع، توزیع پیشین مناسب در نظر بگیریم.

نکته ای که قبل از ارائه توزیع های پیشین باید گفت این است که در الگوریتم گیبس مهمترین مسأله، بدست آوردن Full Condition ها هست.

پس روند کار ما بدین صورت هست که با استفاده از توزیع های پیشین برای پارامترها و با کمک متغیر پنهانی که تعریف می کنیم می خواهیم ابتدا توزیع پسین را بدست آورده و سپس Full Condition ها را بدست بیاوریم.

خب همان طور که گفته شد باید برای پارامترهای مدل توزیع پیشین درنظر بگیریم:

توزیع پیشین مناسبی که برای میانگین هر توزیع وجود دارد همان پیشین نرمال و توزیع مناسب برای واریانس توزیع نرمال ، توزیع معکوس گاما هست.

امّا برای پارامترهای (π_1, π_2, π_3) توزیع توأم دریخله یا دریکله را درنظر می گیریم.

توزیع دریکله ، توزیع بتای چندمتغیره هست !!

بنابراین داریم:

(1) $(\pi_1, \pi_2, \pi_3) \sim Dr(\alpha_1, \alpha_2, \alpha_3)$

$\Rightarrow \pi(\pi_1, \pi_2, \pi_3) \propto \pi_1^{\alpha_1 - 1}\ \pi_2^{\alpha_2 - 1}\ \pi_3^{\alpha_3 - 1}$

(2) $\forall\ j = 1, 2, 3 \qquad \mu_j | \sigma_j^2 \sim N(\lambda_j, \sigma_j^2 / \tau_j) \Rightarrow$

دقت شود اینجا Conditional Conjuga
درنظر گرفتیم

(3) $\forall j = 1, 2, 3 \qquad \sigma_j^2 \sim IG(a_j, b_j)$

نکته: خب ما گفتیم متغیر تصادفی پنهانی به نحوی عمل می کنه که با دانستن این متغیر پنهانی ، توزیع ما مشخص میشود؛ $Y_i | W_i = j \sim N(\mu_j, \sigma_j^2)$

پس $P(W_i = j) = \pi_j$.

درنتیجه

$W_i \sim M(1, \pi_1, \pi_2, \pi_3)$

که بهتره بگیم

$W_i | \underline{\pi} \sim M(1, \pi_1, \pi_2, \pi_3)$

(دقت شود ما این متغیر تصادفی پنهانی را به عنوان یک پارامتر درنظر می گیریم.)

حالا نوبت بدست آوردن توزیع پسین است :

فرض کنید ابتدا متغیرهای پنهانی $(w_1, \dots, w_N)$ را به توزیع اضافه نکنیم
در این صورت :

$$\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \pi_1, \pi_2, \pi_3, \mu_3, \sigma_3^2 \mid \underset{\sim}{y}) \propto L(\mu_1, \dots, \pi_3 \mid \underset{\sim}{y}) \prod_{j=1}^{3} \pi(\mu_j \mid \sigma_j^2)\; \pi(\pi_1, \pi_2, \pi_3)$$

$$\Rightarrow \pi(\mu_1, \dots, \pi_3 \mid \underset{\sim}{y}) \propto \prod_{i=1}^{N} \sum_{j=1}^{3} \pi_j f_j(y_i) \cdot \dots\dots$$

خب همان طور که مشاهده می کنید این قسمت به مشکل برمی خوریم و دیگه نمی توانیم
جلوتر برویم و ساده کنیم (مخصوصا اگر N زیاد شود)

راه حل: راه حل این مشکل اضافه کردن همین پارامتر پنهان به مدل است.

در این صورت توزیع پسین به شکل زیر تغییر و اصلاح می شود:

$$\pi(\mu_1, \dots, \pi_3, w_1, \dots, w_N \mid \underset{\sim}{y}) \propto L(\mu_1, \dots, \pi_3 \mid \underset{\sim}{w}, \underset{\sim}{y})\; \pi(\underset{\sim}{w} \mid \pi_1, \pi_2, \pi_3)$$

حالا در نظر بگیرید n_1 تا از مشاهدات از توزیع نرمال با میانگین μ_1 و واریانس σ_1^2
n_2 تا از نرمال دوم و n_3 تا از نرمال سوم باشند بطوریکه $n_1+n_2+n_3=N$

در نتیجه توزیع پسین ما به صورت زیر خواهد بود :

(در ادامه خواهیم گفت که در طول شبیه سازی مقدار n_1 ، n_2 و n_3 تغییر می کند ولی
در عمل ما نمی دانیم که y_i دقیقاً از چه توزیعی است.)

$$\pi(\mu_1,\ldots,\pi_3 \mid \underset{\sim}{W}, \underset{\sim}{Y}) \propto \left(\frac{1}{\sigma_1^2}\right)^{\frac{n_1}{r}} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{N} I_{W_i=1}(y_i-\mu_1)^2\right\} \times \left(\frac{1}{\sigma_2^2}\right)^{\frac{n_r}{2}} \exp\left\{-\frac{1}{2\sigma_2^2}\sum_{i=1}^{N} I_{W_i=2}(y_i-\mu_2)^2\right.$$

$$\times \left(\frac{1}{\sigma_3^2}\right)^{+\frac{n_r}{2}} \exp\left\{-\frac{1}{2\sigma_3^2}\sum_{i=1}^{N} I_{W_i=3}(y_i-\mu_3)^2\right\} \times \prod_{j=1}^{3} \left(\frac{1}{\sigma_j^2}\right)^{+\frac{1}{2}} \exp\left\{-\frac{\tau_j}{2\sigma_j^2}(\mu_j-\lambda_j)^2\right\} \left(\frac{1}{\sigma_j^2}\right)^{a_j+1} e^{-\frac{b_j}{\sigma_j^2}}$$

$$\times\ \pi_1^{\alpha_1-1}\ \pi_2^{\alpha_2-1}\ \pi_3^{\alpha_3-1}$$ [redacted] $$\times\, p(W_1,\ldots,W_N) =$$

$$\left(\frac{1}{\sigma_1^2}\right)^{\frac{n_1}{r}+a_1+\frac{1}{r}+1} \exp\left\{-\frac{1}{2\sigma_1^2}\sum_{i=1}^{N} I_{W_i=1}(y_i-\mu_1)^2 - \frac{\tau_1}{2\sigma_1^2}(\mu_1-\lambda_1)^2 - \frac{b_1}{\sigma_1^2}\right\} \times$$

$$\left(\frac{1}{\sigma_2^2}\right)^{\frac{n_r}{2}+a_r+\frac{1}{r}+1} \exp\left\{-\frac{1}{2\sigma_2^2}\sum_{i=1}^{N} I_{W_i=2}(y_i-\mu_2)^2 - \frac{\tau_2}{2\sigma_2^2}(\mu_2-\lambda_2)^2 - \frac{b_r}{\sigma_2^2}\right\} \times$$

$$\left(\frac{1}{\sigma_3^2}\right)^{\frac{n_r}{2}+a_r+\frac{1}{2}+1} \exp\left\{-\frac{1}{2\sigma_3^2}\sum_{i=1}^{N} I_{W_i=3}(y_i-\mu_3)^2 - \frac{\tau_2}{2\sigma_2^2}(\mu_3-\lambda_3)^2 - \frac{b_3}{\sigma_3^2}\right\}$$

$$\times\ \pi_1^{\alpha_1-1}\ \pi_2^{\alpha_r-1}\ \pi_3^{\alpha_3-1} \times \pi_1^{n_1}\ \pi_2^{n_r}\ \pi_3^{n_r}$$

حال برای اجرای الگوریتم گیبس نیاز به بدست آوردن Full Condition ها داریم:

پس باید توزیع های زیر را بدست بیاوریم:

$$\forall j=1,2,3 \Rightarrow \begin{cases} \mu_j \mid \text{تمامی پارامترها} \text{ و } \underset{\sim}{y} \text{ و } \underset{\sim}{w} \\ \sigma_j^2 \mid \text{تمامی پارامترها} \text{ و } \underset{\sim}{y} \text{ و } \underset{\sim}{w} \end{cases}$$

$$(\pi_1, \pi_2, \pi_3) \mid \text{تمامی پارامترها} \text{ و } \underset{\sim}{y} \text{ و } \underset{\sim}{w}$$

درمثال، با استفاده از توزیع پیشین، Full Condition های ما بصورت زیر خواهد شد:

$$(\pi_1, \pi_2, \pi_3) \mid \underset{\sim}{y}, \underset{\sim}{w}, \underset{\sim}{\mu}, \underset{\sim}{\sigma}^2 \sim Dr(\alpha_1+n_1, \alpha_2+n_2, \alpha_3+n_3)$$

$$\mu_j \mid \sigma_j^2, \underset{\sim}{y}, \underset{\sim}{w} \sim N\left(\frac{\lambda_j \tau_j + S_j^y}{\tau_j + n_j}, \frac{\sigma_j^2}{\tau_j + n_j}\right)$$

$$\text{s.t} \quad S_j^y = \sum_{i=1}^{N} I_{w_i=j}\, y_i$$

$$\sigma_j^2 \mid \mu_j, \underset{\sim}{y}, \underset{\sim}{w} \sim IG\left(\frac{n_j}{2} + \frac{1}{2} + a_j, b_j + \frac{1}{2} S_j^v + \frac{1}{2}\tau_j (\mu_j - \lambda_j)^2\right)$$

همچنین

$$p(w_i = j) \propto \frac{\pi_j}{\sigma_j} \exp\left(-(x_i - \mu_j)^2 / 2\sigma_j^2\right)$$

?!

حال برای بدست آوردن برآورد بیزی پارامترهای روش الگوریتم گیبس به شکل زیر عمل می‌کنیم:

نحوهٔ اجرای الگوریتم گیبس:

~~تولید~~ ① ابتدا یک مقدار اولیه برای بردارهای π و μ و σ^2 در نظر می‌گیریم و داریم $\pi^{(0)}$, $\mu^{(0)}$, $\sigma^{2(0)}$

مرحلهٔ t ام الگوریتم گیبس به صورت زیر خواهد بود:

① تولید w_i برای $i=1,\dots,n$ از $j=1,2,3$ بوسیله توزیع چند جمله ای با احتمال های

مرحلهٔ t ام (t) — مرحلهٔ $t-1$ $(t-1)$

$$P(w_i=j)^{(t)} = \frac{\pi_j^{(t-1)}}{\sigma_j^{(t-1)}} \exp\left\{-\left(y_i-\mu_j^{(t-1)}\right)^2 / 2\sigma_j^{2(t-1)}\right\}$$

② محاسبهٔ $n_j^{(t)}=\sum_{i=1}^{N} I_{w_i=j}$ & $S_j^{y(t)}$

③ تولید $(\pi_1,\pi_2,\pi_3)^{(t)}$ from $Dr(\alpha_1+n_1,\alpha_2+n_2,\alpha_3+n_3)$

④ تولید $\mu_j^{(t)}$

$$\mu_j^{(t)} \text{ from } N\left(\frac{\lambda_j\tau_j+S_j^{y(t)}}{\tau_j+n_j^{(t)}},\ \frac{\sigma_j^{2(t-1)}}{\tau_j+n_j^{(t)}}\right)$$

⑤ سپس محاسبه $S_j^{v(t)}$ که برابر است با $\sum_{i=1}^{N} I_{w_i=j}\left(y_i-\mu_j^{(t)}\right)^2$

۵) تولید $\sigma_j^{2\,(t)}$

$$\sigma_j^{2\,(t)} \text{ from } IG\left(a_j+\frac{n_j+1}{2},\ b_j+\frac{1}{2}\tau_j(\mu_j^{(t)}-\lambda_j)^2+\frac{1}{2}S_j^{2\,(t)}\right)$$

* دقت شود اون زمان که ابتدا در توزیع پسین تعریف کردیم رو نگاریم و برای روشن شدن مطالب و بدست آوردن Full Condition ها گفته شد.

توضیح مختصر در مورد الگوریتم گیبس :

الگوریتم نمونه گیری گیبز، یکی از منعطف ترین روش های طرح نمونه از توزیع های احتمال است که برای اولین بار توسط گیبز در فیزیک آماری به کار برده شده.

نمونه گیری گیبز حالت خاصی از الگوریتم متروپولیس - هستینگ است که در اینجا مقادیر پیشنهاد شده همواره مورد پذیرش قرار می گیرد.

فرض کنید بردار پارامتر و تصادفی ما به صورت $\Theta = (\theta_1, \dots, \theta_p)$ در دست است و علاوه بر آن فرض کنید می توان از چگالی های شرطی یک متغیره $f_1, \dots, f_p$ که آنها را شرطی کامل یا full Conditionals گویند شبیه سازی کرد.

در این صورت الگوریتم گیبس به صورت زیر عمل می کند:

(۱) با مقدار آغازین $\Theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_p^{(0)})$ شروع می کنیم.

(۲) در مرحله $(n+1)$ ام با داشتن مقادیر $\Theta^{(n)} = (\theta_1^{(n)}, \dots, \theta_p^{(n)})$ مقادیر زیر را از چگالی های شرطی $f_1, \dots, f_p$ تولید می کنیم:

$$\theta_1^{(n+1)} \sim f_1(\theta_1 \mid \theta_2^{(n)}, \dots, \theta_p^{(n)})$$

$$\theta_2^{(n+1)} \sim f_1(\theta_2 \mid \theta_1^{(n+1)}, \theta_3^{(n)}, \dots, \theta_p^{(n)})$$

$$\vdots$$

$$\theta_p^{(n+1)} \sim f_1(\theta_{p+1} \mid \theta_1^{(n+1)}, \dots, \theta_{p-1}^{(n+1)})$$

۳- قرار می دهیم $n = n+1$ و مرحله ۲ را تکرار می کنیم.