۲-۸- روش آماری بیژین

بیژین روشی آماری است که سعی در برآورد پارامترهای ژنتیکی بر اساس قانون احتمالات شرطی دارد. قضیه بیز توسط توماس بیز نوشته شده است و از یک قانون ساده احتمال شرطی استفاده می‌کند. در اکثر موارد علت وقوع پیشامدها از قبل معلوم می‌باشد. مثلاً برای تعیین احتمال شیر بودن یک سکه فرض بر این بود که اگر یک سکه پرتاب شود احتمال شیر بودن آن ۰/۵ می‌باشد زیرا اگر این آزمایش را به دفعات انجام دهیم احتمال آنکه شیر بیاید ۰/۵ است. اما در مواردی ما نمی‌توانیم یک آزمایش را به دفعات انجام دهیم و وقوع پیشامدها برای ما نا معلوم است. فرض کنید $x_1,x_2,\ldots,x_n$ پیشامدهای نا سازگار ما هستند و $p(x_i)$ احتمال وقوع پیشامد x است. همچنین فرض کنید y واقعه‌ای است که همزمان با یکی از وقایع x_i اتفاق می‌افتد. در این صورت $p(x|y)$ احتمال شرطی وقوع x است. اکنون اگر x اتفاق افتاده باشد با استفاده از فرمول بیز می‌توانیم احتمال وقوع $p(y|x)$ را محاسبه کنیم.

$$P(xy) = p(x) \,.\, p(y|x) \Rightarrow p(y|x) = p(xy)/p(x)$$ معادله ۲-۱۰

$$P(xy) = p(y) \,.\, p(x|y) \Rightarrow p(x|y) = p(xy)/p(y) \Rightarrow \quad p(y|x) \,.\, p(x)/p(y)$$

بیژین‌ها احتمال را یک تصویر ذهنی می‌دانند درباره اینکه، چه میزان می‌شود درباره احتمال وقوع یک پیشامد درست فکر کرد در حالی‌که آمارگران روتین، احتمال را به معنی متوسط وقوع طولانی مدت یک اتفاق، وقتی یک آزمایش چندین بار تکرار می‌شود می‌دانند. بنابراین آمارگران روتین با صراحت بین یک متغیر تصادفی که می‌تواند در هر آزمایش تغییر کند و یک پارامتر که همیشه ثابت است، تبعیض قائل شدند. آنها احتمال را برای متغیر تصادفی بیان می‌کنند، نه برای پارامترها اما بیژین‌ها بین این دو هیچ فرقی نمی‌گذارند. آمارگران روتین، پارامترها را به وسیله ماکزیمم درستنمایی برآورد می‌کنند. درستنمایی داده‌ها، $p(y|x)$ است که به صورت تابعی از پارامتر مطرح می‌شود. برآورد حداکثر درستنمایی(ML[1]) عبارت است از یافتن ارزش‌هایی از پارامتر که $p(y|x)$ را ماکزیمم کند و به این مقدار x-hat می‌گویند. اگر یک آزمایش را چندین بار تکرار کنیم مقدار پارامتر ثابت است اما مقادیر متفاوتی از x-hat را به دست می‌آوریم. بنابراین می‌توانیم توزیع x-hat را به شکل $N(x,V^2=\sigma^2/n)$ نشان دهیم که V خطای استاندارد x-hat می‌باشد. بنابراین می‌توانیم x-hat را به این صورت بنویسیم:

$$P(x-2*v<x\text{-}hat<x+2*v)$$ معادله ۱۱-

[1]- Maximum likelihood

بر خلاف آمارگران روتین هدف بیژین در آنالیز آزمایشات ایجاد بیان احتمالی است که نزدیک به مقدار واقعی پارامتر باشد. آمارگر بیژین این کار را از طریق تئوری بیز انجام می‌دهد. تئوری بیز عبارتست از:

$$P(x|y) \propto P(y|x) * P(x)$$ معادله ۱۲-۲

x و y متغییرهایی هستند که به وسیله تابع چگالی احتمال شرح داده می‌شوند. P(x|y) را احتمال پسین می‌گوییم زیرا بعد از انجام آزمایش به دست می‌آیند. P(y|x) درستنمایی است که آمارگران روتین از آن استفاده می‌کردند و P(x) پیش فرض (توزیع پیشین) است، زیرا احتمالی از x است که قبل از انجام آزمایش در دست داریم و به بیژین این امکان را می‌دهد که از دانسته‌های گذشته مقدار x را به دست بیاورد. اگر آمارگر بیژین داده‌ها را آنالیز کند هیچ نظری برای مقدار پارامتر قبل از انجام آزمایش ندارد، وی می‌تواند از داده‌های پیش فرض (پیشین) که مقدار یکسان برای همه مقدارهای x دارد استفاده کند سپس داده‌های پسین به سادگی، متناسب با تابع درستنمایی می‌شوند. توزیع پسین شامل اطلاعات بیشتری علاوه بر میانگین و واریانس خطای پیش‌بینی است و برای یک بیژین، شامل تمام اطلاعات درباره پارامترهایی است که شناخته شده‌اند. یکی از تفاوت‌های بیژین و آمارگران کلاسیک، این است که بیژین‌ها تمام اثرات را به شکل متغیر تصادفی در نظر می‌گیرند و به جای بحث در مورد اینکه یک اثر ثابت است یا تصادفی، بهتر می‌دانند به خصوصیات اثرات برآورد شده رسیدگی شود.

اگر برآورد x-hat آمارگر روتین، برای یک اثر ثابت باشد، برای مثال برآورد حداقل مربعات، E(x-hat|x)=x به این معنی است که اگر شما آزمایش را به دفعات تکرار کنید، به طور متوسط نتیجه‌ی درست را می‌گیرید. این ویژگی آمارگران روتین است که به آن نا اریب می‌گویند. اگر x-hat یک برآورد Blue[2] برای اثرات ثابت باشد، E(x | x-hat)=x-hat به این معنی است که، اگر شما بهترین اثر را تنها مبتنی بر x-hat انتخاب کنید، به طور متوسط اثرات درست، در حد انتظار شما خواهند بود. برآورد حداقل مربعات این ویژگی را ندارد. به عنوان نمونه، بهترین برآورد حداقل مربعات، بیش از مقدار واقعی است.

بدتر از آن این است که برآورد حداقل مربعات، دارای قابلیت اطمینان پایینی است، که بیشتر مشابه برآوردهای بالاتر از ارزش واقعی می‌باشند. در مقایسه با برآورد حداقل مربعات، برآورد Blup[3] نسبت به میانگین گرایش بیشتری دارد

---

[2]- Best Linear Unbiased Estimation
[3]- Best Linear Unbiased Prediction

چون از توزیع پیشین که دارای میانگین صفر است استفاده می‌کند. هر چقدر اطلاعات بیشتری برای یک اثر جمع‌آوری شود درستنمایی بیشتری در داده های پیشین حاکم می‌شود و برآورد اشتباه کمتری دارد.

هر دو روش برآورد، قابل استفاده است. اثر ثابت یا حداقل مربعات، خلاصه خوبی از یک آزمایش است ولی پاسخ‌های بیژین چون خطای واریانس پیش‌بینی کمتری نسبت به پاسخ‌های حداقل مربعات دارند واقعی‌تر هستند. اگر پیش‌فرض‌ها به خوبی انتخاب نشوند این ویژگی مفید برآوردهای بیژین کمتر خود را نشان می‌دهد. با این حال اطلاعات پیشین نسبتاً مبهم، نسبت به نتایج بدون اطلاعات پیشین احتمالات بهتری هستند.

اغلب تعدادی پارامتر، برای توضیح دادن اطلاعات مربوط به مشاهدات لازم است. روش بیژین با استفاده از قضیه بیز، اشتراک توزیع پسین پارامترها را بدست می‌آورد. در یک مثال ساده می‌توانیم از یک نمونه با توزیع نرمال، برای برآورد میانگین و واریانس یک جامعه استفاده کنیم. روش بیژین اشتراک توزیع پسین میانگین و واریانس را به دست می‌آورد.

$$P(\mu,\sigma^2|y) = P(y|\mu,\sigma^2) * P(\mu,\sigma^2)$$ معادله ۲-۱۳

اگر اشتراک میانگین و واریانس شناخته شده باشد، می‌شود به وسیله انتگرال‌گیری از یکی ار متغیرها، توزیع حاشیه ای به دست آورد.

$$P(\sigma^2|y) = \int P(\mu,\sigma^2|y)d\mu$$ معادله ۲-۱۴

توزیع حاشیه‌ای می‌تواند برای استنباط σ^2 استفاده شود. برای مثال، میانگین توزیع پسین حاشیه‌ای، به طور قراردادی، برآورد نا اریب σ^2 است که بر n-1 تقسیم می‌شود. توزیع حاشیه‌ای مشابه توزیع نرمال نیست. با انتگرال‌گیری از μ می‌توانیم خطای تخمین μ را وقتی σ^2 را برآورد می‌کنیم بدست آوریم.

بعضی وقت‌ها مدل‌ها پارامترهایی را دارند که برای ما جالب نیستند، اما چون به ما اطلاعات می‌دهند و باعث بهبود برآورد دیگر پارامترها می‌شوند، باید باقی بمانند. بیژین با این دسته از پارامترها به یک شکل برخورد می‌کند و آنها در مدلی با توزیع پیشین مخصوص خود قرار می‌گیرند، جمع و یکپارچه می‌شوند که توزیع حاشیه‌ای از پارامترهایی که برای ما جالب هستند را به وجود بیاورند. این موضوع یک ویژگی مهم روش بیژین را نشان می‌دهد. آمارگران روتین به صراحت بین پارامترهایی که توزیع را توصیف می‌کنند و درک متغیرهای تصادفی این توزیع‌ها، تفاوت قائل می‌شوند (بلاسکو[۴]، ۲۰۰۱).

4- Blasco

۲-۹ نمونه گیری گیبس

محاسبه توزیع پسین و انتگرال‌گیری بعضی پارامترها مشکل است. اغلب پیدا کردن راهی برای رسیدن به پاسخ دشوار است. بیژین روش‌های مختلفی برای غلبه بر این مشکل پیدا کرده‌اند که از جمله به نمونه‌گیری گیبس می‌توان اشاره نمود (بلاسکو، ۲۰۰۱). نمونه‌گیری گیبس نخستین بار توسط گمن[5]و گمن (۱۹۸۴) ابداع گردید. در اصلاح دام، ونگ[6] و همکاران (۱۹۹۳)، نمونه‌گیری گیبس را برای برآورد مؤلفه‌های واریانس در مدل مولد نر و مدل دام مورد استفاده قرار دادند. همچنین از این روش برای مطالعه مؤلفه‌های کوواریانس در مدل‌های با اثرات مادری (ژنسن[7] و همکاران، ۱۹۹۴)، مدل‌های آستانه‌ای (سورنسن[8] و همکاران،۱۹۹۵) و مدل‌های رگرسیون تصادفی (جامروزیک[9] و همکاران، ۱۹۹۷)، نیز استفاده شده است. اخیراً نیز این روش برای برآورد مؤلفه‌های واریانس و پیش‌بینی ارزش ارثی در مدل‌های خطی-آستانه‌ای به کار گرفته شده است (مرود،۱۹۹۶). نمونه‌گیری گیبس، یک روش انتگرال‌گیری عددی بوده و یکی از انواع روش‌های مونت کارلوی زنجیره مارکوف (MCMC) می‌باشد. در این روش‌ها، نمونه‌هایی از درون توزیع‌های مشخص گرفته شده و به همین دلیل آنها را مونت کارلو می‌نامند. همچنین دلیل اطلاق نام زنجیره مارکوف به این روش‌ها، آن است که هر نمونه وابسته به نمونه قبلی می‌باشد. در نمونه‌گیری گیبس، نمونه‌های تصادفی از توزیع‌های پسین حاشیه‌ای، با استفاده از نمونه‌گیری تکراری از توزیع‌های پسین شرطی، تولید می‌شوند. به عنوان مثال، اگر $Q' = (Q1,Q2)$ بوده و $P(Q1,Q2)$ ، توزیع توأم متغیرهای Q1 و Q2 باشد، آنگاه نمونه‌گیری گیبس شامل نمونه‌گیری از توزیع‌های پسین شرطی کامل متغیرهای Q1 (یعنی $P(Q1|Q2)$) و Q 2 (یعنی $P(Q2|Q1)$) می‌باشد. بنابراین اگر توزیع پسین توأم به طور نسبی معلوم باشد، می‌توان توزیع‌های شرطی را ایجاد نمود. ولی برای توصیف چگالی توأم لازم است که از تئوری بیز استفاده شود. به طور کلی، احتمال وقوع همزمان دو پیشامد ($P(B,Y)$) برابر است با:

$$P(B,Y) = P(B)P(Y|B) = P(Y)P(B|Y)$$ معادله ۲-۱۵

بنابراین:

$$P(B|Y) = P(B)P(Y|B) / P(Y)$$ معادله ۲-۱۶

---

[5]- Geman
[6]- Wang
[7]- Jensen
[8]- Sorensen
[9]- Jamrozik

این معادله (۱۶-۲) نشان می‌دهد که استنباط و نتیجه‌گیری در مورد متغیر B به احتمال پیشین وقوع آن (یعنی$P(B)$) بستگی دارد. اگر مشاهداتی از Y در دسترس باشد، آنگاه از نو شدن این احتمال پیشین، برای محاسبه احتمال پسین یا چگالی B (یعنی $P(B|Y)$) استفاده می‌شود. چون مخرج معادله تابعی از B نیست، لذا معمولاً این معادله به صورت زیر نشان داده می‌شود:

$$P(B|Y) \propto P(B)P(Y|B)$$ معادله ۱۷-۲

بنابراین چگالی پسین B، متناسب با احتمال پیشین وقوع آن ضربدر توزیع شرطی Y به شرط B است. اگر در معادله(۱۷-۲) به جای B، بردار W، که برداری از پارامترها بوده و عبارت از $w' = (w_1, w_2, w_3)$ می‌باشد، در نظر گرفته شود و توزیع پسین توأم نیز به صورت نسبی معلوم فرض گردد (معادله۱۷-۲)، آنگاه احتمال‌های شرطی مورد نظر برای نمونه‌گیری گیبس را می‌توان به صورت $P(W_1|W_2, W_3, Y)$ ، $P(W_2|W_1, W_3, Y)$ و $P(W_3|W_1, W_2, Y)$ ایجاد نمود. اگر مقادیر اولیه به صورت $W_1^{[0]}$ ، $W_2^{[0]}$ و $W_3^{[0]}$ منظور شوند، اجرای نمونه‌گیری گیبس، شامل تکرار حلقه زیر خواهد بود:

۱- مقدار $W_1^{[i+1]}$ از $P\left(W_1 \middle| W_2^{[i]}, W_3^{[i]}, Y\right)$ نمونه‌گیری می‌شود.

۲- مقدار $W_2^{[i+1]}$ از $P\left(W_2 \middle| W_1^{[i+1]}, W_3^{[i]}, Y\right)$ نمونه‌گیری می‌شود.

۳- مقدار $W_3^{[i+1]}$ از $P\left(W_3 \middle| W_2^{[i+1]}, W_3^{[i+1]}, Y\right)$ نمونه‌گیری می‌شود.

۱۰-۲ دورهای سوخته[۱۰]

در نمونه‌گیری گیبس معمولاً نمونه‌های ابتدایی حذف می‌شوند. در دورهای طولانی، مقدار آغازین توزیع که زنجیره ایجاد می کند، چون توزیع در نهایت به همگرایی می رسد ، مؤثر نخواهد بود. اما در دورهای کم ، نخستین مقدار آغازین در نتایج تأثیرگذار می‌باشد. روش نرمال جلوگیری از نتایج متأثر از مقادیر آغازین، حذف کردن تعداد کمی از نمونه‌ها است این دورها را اصطلاحاً دورهای سوخته می‌نامند (گودارد[۱۱] ۲۰۱۰ و مرود،۱۹۹۶).

۱۱-۲ همبستگی[۱۲] بین چرخه‌ها و کاهش تعداد دورها

[10] - Burn - in
[11]- Goddard
[۳] - Autocorrelation

دومین مشکل نمونه‌گیری گیبس این است که زنجیره گیبس آزادانه بین مقادیر دورها حرکت نمی کند. این یک مشکل شایع در روش گیبس است که از طریق همبستگی خودکار بین یک چرخه و چرخه بعد، توصیف می شود. این مسئله، از آنجایی که تنها یک پارامتر در یک زمان تغییر کرده است، تعجب‌آور نیست. ساده‌ترین راه حل، انجام دادن تعداد زیادی سیکل است، به طوری که دورهای یک گروه، میانگین دورهای گروه دیگر باشد. این راه حل ارائه شده می‌تواند تمام راه حل‌های ممکن توسط تمامی آغازگرها را شامل شود. اگر این‌گونه عمل نشود، راه حل‌های ممکن هیچ وقت تست نمی شوند و در نمونه نهایی ظاهر نمی‌شوند. به چنین زنجیره‌هایی ساده شدنی یا تقلیل‌پذیر (کاهش تعداد دورها) می گویند و نمی‌توانند برای برآورد پارامترها استفاده شوند. بنابراین اگر سیکل‌های بیشتری استفاده کنیم و تعداد کمی از نمونه ها را حذف نماییم برآوردهای معتبری از پارمترها به دست خواهد آمد (گودارد۱۹ و بلاسکو، ۲۰۰۱).

۲-۱۲ نمونه‌گیری گیبس در مدل دام تک صفتی

مدل خطی تک صفتی زیر در نظر گرفته شود:

معادله ۲-۱۸

$$y = Xb + Zu + e$$

y = بردار مشاهدات به ابعاد $n \times 1$ n = تعداد مشاهدات

b = بردار اثرات ثابت به ابعاد $p\times 1$ p = تعداد سطوح اثرات ثابت

u = بردار اثرات تصادفی به ابعاد $q \times 1$ q = تعداد سطوح اثرات تصادفی

X = ماتریس ضرائب اثرات ثابت به ابعاد $n \times p$ که رکوردها را به اثرات ثابت مرتبط می‌کند.

Z = ماتریس ضرائب اثرات تصادفی به ابعاد $n \times q$ که رکوردها را به اثرات تصادفی مرتبط می‌کند.

e = بردار اثرات باقیمانده.

توزیع شرطی که داده‌ها(y) را ایجاد می‌کند، عبارت است از:

معادله ۲-۱۹

$$y|b, u, \sigma_e^2 \sim N(Xb + Zu + R\sigma_e^2)$$

۲-۱۲-۱ توزیع‌های پیشین:

برای تکمیل خصوصیات بیزی مدل، توزیع‌های پیشین b، u، σ_e^2، σ_u^2 لازم می‌باشند (ونگ و همکاران، ۱۹۹۳). معمولاً برای b یک توزیع تخت در نظر گرفته می شود. بنابراین:

معادله ۲-۲۰

$$P(b)\sim constant$$

در نظر گرفتن یک توزیع یکنواخت[13] نشان‌دهنده فقدان اطلاعات اولیه راجع به این بردار است. ولی اگر اطلاعاتی در مورد مقدار اولیه b (از نظر حدهای بالا و پایین آن) در اختیار باشد، می‌توان از آن برای تعریف توزیع پسین b استفاده نمود. در این حالت، توزیع پیشین را توزیع پیشین مناسب[14] می نامند. با فرض وجود یک مدل نا محدود، توزیع u از نوع نرمال چند متغیره بوده و عبارت است از:

$$u|A, \sigma_u^2 \sim N(0, A\sigma_u^2)$$ معادله ۲۱-۲

معمولاً برای مؤلفه‌های واریانس از یک توزیع مربع کای معکوس مقیاس شده (χ^2) به عنوان مقادیر اولیه استفاده می‌شود (ونگ و همکاران،۱۹۹۳). بنابراین برای واریانس باقیمانده می‌توان نوشت:

$$P(\sigma_e^2|v_e, s_e^2) \propto (\sigma_e^2)^{-\left(\frac{v_e}{2}+1\right)} \exp\left(\frac{v_e s_e^2}{2\sigma_e^2}\right)$$ معادله ۲۲- ۲

و برای واریانس ژنتیکی افزایشی نیز می‌توان نوشت:

$$P(\sigma_u^2|v_u, s_u^2) \propto (\sigma_u^2)^{-\left(\frac{v_u}{2}+1\right)} \exp\left(\frac{v_u s_u^2}{2\sigma_u^2}\right)$$ معادله ۲۳-۲

در این روابط، (v_e و یا v_u) یک پارامتر درجه اعتماد بوده و(s_e^2 یا s_u^2) را نیز می‌توان به عنوان یک مقدار اولیه مناسب برای مؤلفه واریانس محسوب نمود. در روشی دیگر، می‌توان برای مؤلفه‌های واریانس، توزیع یکنواخت پیشین در نظر گرفت:

$$P(\sigma_j^2) \propto \text{constant}$$ معادله ۲۴-۲

در این رابطه، $\sigma_j^2 = \sigma_e^2$ یا $\sigma_j^2 = \sigma_u^2$ بوده و می‌توان بر اساس اطلاعات قبلی، یک حد بالا برای σ_j^2 در نظر گرفت.

۲-۱۲-۲ توزیع‌های شرطی توأم و کامل:

توزیع پسین توأم پارامترها (b، u، σ_e^2، σ_u^2) متناسب با حاصل‌ضرب تابع درستنمایی و توزیع پیشین توأم می‌باشد. با فرض وجود n مشاهده و m حیوان و با استفاده از معادلات ۱۹-۲ تا ۲۳-۲ توزیع پسین توأم را می‌توان به صورت زیر نوشت:

$$P(b, u, \sigma_u^2, \sigma_e^2|y)$$ معادله ۲۵-۲

$$\propto (\sigma_e^2)^{-\left(\frac{n+v_e}{2}+1\right)} \exp\left[-\frac{(y-Xb-Zu)'(y-Xb-Zu)+v_e s_e^2}{2\sigma_e^2}\right]$$

---

[13] - Uniform distribution
[14] - proper

$$(\sigma_u^2)^{-\left(\frac{m+v_u}{2}+1\right)}\exp\left(\frac{u'A^{-1}+v_u s_u^2}{2\sigma_u^2}\right)$$

با فرض معلوم بودن سایر پارامترها در معادله ۲۵-۲، می‌توان توزیع پسین شرطی کامل هر یک از پارامترها را تعیین نمود. بنابراین برای b می‌توان نوشت:

معادله ۲۶-۲

$$P(b|u,\sigma_u^2,\sigma_e^2,y) \propto \exp\left(\frac{(y-Xb-Zu)'(y-Xb-Zu)}{2\sigma_e^2}\right)$$

توزیع متناظر با رابطه فوق عبارت است از:

معادله ۲۷-۲

$$Xb|u,\sigma_u^2,\sigma_e^2,y \sim N(y-Zu, I\sigma_e^2)$$

$$u,\sigma_u^2,\sigma_e^2,y \sim N(X'(y-Zu), X'X\sigma_e^2)\ X'Xb|$$

بنابراین:

$$b|u,\sigma_u^2,\sigma_e^2,y \sim N(\hat{b},(X'X)^{-1}\sigma_e^2)$$

در این رابطه:

$$\hat{b} = (X'X)^{-1}X'(y-Zu)$$

بنابراین برای i امین سطح b می‌توان نوشت:

معادله ۲۸-۲

$$b_j|b_{-j},u,\sigma_u^2,\sigma_e^2,y \sim N(\hat{b}_j,(X_j'X_j)^{-1}\sigma_e^2)$$

در اینجا رابطه، $b_j = \left(X_j'X_j\right)^{-1} X_j'(y_j - X_{-j}b - Zu)$ می‌باشد که معادل فرمول پاسخ مربوط به jامین اثر ثابت در روش BLUP می‌باشد. همچنین، X_j، ردیف jام ماتریس X بوده و b_{-j}، بردار b که سطح jام آن حذف شده است، می‌باشد.

به همین ترتیب، توزیع jامین اثر تصادفی نیز عبارت است از:

معادله ۲۹-۲

$$u_j|b,u_{-j},\sigma_u^2,\sigma_e^2,y \sim N(\hat{u}_j,\left(z_j'z_j + A_{j,j}^{-1}\alpha\right)^{-1}\sigma_e^2)$$

در این رابطه:

$$\hat{u}_j = z_j'z_j + A_{j,j}^{-1}\alpha)^{-1}z_j'(y - Xb - A_{j,-j}^{-1}\alpha u_{-j})$$

این معادله نیز معادل فرمول پاسخ مربوط به jامین اثر تصادفی در روش BLUP است.

توزیع شرطی کامل واریانس باقیمانده را می‌توان با استفاده از معادله ۲۵-۲ و فقط با در نظر گرفتن عباراتی که شامل σ_e^2 هستند، به صورت χ^2 معکوس مقیاس شده به دست آورد (ونگ و همکاران، ۱۹۹۳). بنابراین برای واریانس باقیمانده می‌توان نوشت:

$$P(\sigma_e^2|b,u,\sigma_u^2,y) \propto (\sigma_e^2)^{-\left(\frac{n+v_e}{2}+1\right)}\exp\left(-\frac{\overleftrightarrow{v}_e\overleftrightarrow{s}_e^2}{2\sigma_e^2}\right)$$

در این رابطه:

$$\overleftrightarrow{s}_e^2 = \frac{(y-Xb-Zu)'(y-Xb-Zu)+v_e s_e^2}{\overleftrightarrow{v}_e} \quad \text{و} \quad \overleftrightarrow{v}_e = n + v_e$$

بنابراین:

$$\sigma_e^2|b,u,\sigma_u^2,y \sim \overleftrightarrow{v}_e\overleftrightarrow{s}_e^2\chi_{\overleftrightarrow{v}_e}^{-2}$$ معادله ۳۰-۲

این نمونه‌گیری از یک توزیع χ^2 با پارامتر مقیاس $(y-Xb-Zu)'(y-Xb-Zu)+v_e s_e^2$ و درجه آزادی برابر $\overleftrightarrow{v}_e$ می‌باشد.

همچنین، توزیع شرطی کامل σ_u^2 نیز به صورت مربع کای معکوس است. بنابراین:

$$P(\sigma_u^2|b,u,\sigma_e^2,y) \propto (\sigma_u^2)^{-\left(\frac{m+v_u}{2}+1\right)}\exp\left(-\frac{\overleftrightarrow{v}_u\overleftrightarrow{s}_u^2}{2\sigma_u^2}\right)$$

در این رابطه:

$$\overleftrightarrow{s}_u^2 = \frac{u'A^{-1}u+v_u s_u^2}{\overleftrightarrow{v}_u} \quad \text{و} \quad \overleftrightarrow{v}_u = m + v_u$$

بنابراین:

$$\sigma_u^2|b,u,\sigma_e^2,y \sim \overleftrightarrow{v}_u\overleftrightarrow{s}_u^2\chi_{\overleftrightarrow{v}_u}^{-2}$$ معادله ۳۱-۲

این رابطه، شامل نمونه‌گیری از یک توزیع χ^2 با پارامتر مقیاس[۱۵] $u'A^{-1}u + v_u s_u^2$ و درجه آزادی برابر $\overleftrightarrow{v}_u$ می‌باشد. بنابراین در نمونه‌گیری گیبس، مقادیر اولیه برای b،u،σ_u^2 و σ_e^2 در نظر گرفته شده و سپس به طور تکراری و به ترتیب از معادله ۲۸-۲ تا ۳۱- ۲، نمونه‌گیری به عمل آمده و در این فرآیند از مقادیر جدید پارامترها که در دور i به دست می‌آیند، در دور $i+1$ استفاده می‌شود. اگر k دور تکرار انجام شده باشد، در این صورت k را طول زنجیره می‌نامند. همان‌گونه که قبلاً توضیح داده شد، معمولاً تعداد j نمونه نخست، به عنوان دور سوخته حذف می‌شوند. این کار برای حصول اطمینان از اینکه نمونه‌های ذخیره شده، تحت تأثیر مقادیر اولیه نبوده و از توزیع پسین حذف شده‌اند، انجام می‌شود. معمولاً اندازه j اختیاری و دلخواه می‌باشد (مرود،۱۹۹۶ و بلاسکو، ۲۰۰۱).

---

[15] - Scale parameter