Bölüm 3

Bayesyen Regresyon Modelinde Stokastik Kısıt Altında Parametrelerin Tahmini

Berrin Gültay[1]

Özet

Genel lineer regresyon modelinde parametrelerin tahmin edilebilmesi için stokastik düzgün önbilgi (önsel) kullanımı son yıllarda önemli bir çözümleme tekniği olarak karşılaşılmaktadır. Bu yöntemi en etkili ve en anlamlı şekilde uygulamanın yolu ise Bayesyen yaklaşımını kullanarak mümkündür. Bayesyen regresyon analizinin karakteristiğini yansıtan özellik, analizde ön bilgiye yer verilmesidir. Bayesyen yaklaşımda, denemeler yapılmadan önce parametreye ilişkin sahip olunan ön bilgi ön olasılık yoğunluk fonksiyonu sayesinde analize dahil edilir. Bu çalışmada da Bayesyen regresyon modelindeki parametreleri tahmin etmek için stokastik ön bilgi olması durumunda kullanılabilecek teorik çıkarımlar ele alınmıştır.

1. Stokastik Kısıt Altında Genel Lineer Regresyon Modelinde Parametrelerin Tahmini

$$y = X\beta + u \tag{1.1}$$

şeklinde genel lineer regresyon modelimiz olsun. Burada y, $T \times 1$ tipinde yanıt değişkenlerin vektörü, X, $T \times p$ tipinde stokastik olmayan açıklayıcı değişkenlerin gözlenen matrisi, β, $p \times 1$ tipinde bilinmeyen regresyon katsayılarının vektörü ve u, $T \times 1$ tipinde sıfır ortalamalı, $\sigma^2 I$ varyans-kovaryans matrisli rasgele hataların vektörüdür, yani $u \sim N(0, \sigma^2 I)$.

Theil ve Goldberger (1961), $R\beta = r$ lineer kısıtlarını stokastik formda kullanarak mixed regresyon modelini geliştirmişlerdir. r, $m \le p$ olmak üzere rassal değişkenin $m \times 1$ boyutlu bir vektör olan lineer kısıtlar,

$$r = R\beta + v$$

[1] Dr. Öğr. Üyesi, Çanakkale Onsekiz Mart Üniversitesi, Fen Fakültesi, İstatistik Bölümü, berringultay@comu.edu.tr, Orcid: 0000-0001-5889-7461

https://doi.org/10.58830/ozgur.pub165.c678

(1.2) olarak yazılabilir. Burada R, $m \times p$ tipinde kısıtların belirlendiği matris, v sıfır ortalamalı, $\sigma_v^2\Omega$ kovaryans matrisli, normal dağılımlı hataların vektörüdür. Yani $v \sim N\left(0, \sigma_v^2\Omega\right)$' dır. Ayrıca u ve v bağımsızdır. (1.2) ile verilen kısıt altında (1.1) ile verilen modelin parametrelerini tahmin edebilmek için modeli y ve X gözlemlerine (1.2)'deki bilgi eklenerek,

$$\begin{bmatrix} y \\ r \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix} \beta + \begin{bmatrix} u \\ v \end{bmatrix} \tag{1.3}$$

şeklinde ya da,

$$y_M = X_M \beta + u_M \tag{1.4}$$

matris formunda yazılabilir. Burada M mixed olduğunu göstermektedir. (1.3) modeli (1.2) önbilgisi ile (1.1)'deki örneklem bilgisini birleştirmektedir. $u_M \sim N\left(0, \Omega_M\right)$'dir. Ω_M, $P:(m+T)\times(m+T)$ dönüşüm matrisi olmak üzere $P'\Omega_M P = I$, $\Omega_M^{-1} = PP'$ olacak şekilde,

$$\Omega_M = \begin{bmatrix} \sigma^2 I_T & 0 \\ 0 & \sigma_v^2\Omega \end{bmatrix} \quad \text{ve} \quad \Omega_M^{-1} = \begin{bmatrix} \frac{1}{\sigma^2} I_T & 0 \\ 0 & \frac{1}{\sigma_v^2}\Omega^{-1} \end{bmatrix} \tag{1.5}$$

formunda pd (pozitif tanımlı) bir matristir. Böylece $P'u_M \sim N\left(0, I\right)$ olur ve

$$P'y_M = P'X_M \beta + P'u_M$$

(1.6)

şeklindeki dönüştürülmüş modele En Küçük Kareler (EKK) tahmin yöntemi uygulanabilir. Bunun için,

$$\left(P'y_M - P'X_M\beta\right)'\left(P'y_M - P'X'_M\beta\right) = \left(y_M - X_M\beta\right)'\Omega_M^{-1}\left(y_M - X_M\beta\right)$$
$$= \frac{\left(y - X\beta\right)'\left(y - X\beta\right)}{\sigma^2} + \frac{1}{\sigma_v^2}\left(r - R\beta\right)'\Omega^{-1}\left(r - R\beta\right)$$

fonksiyonu β'ya göre minimize edilmelidir. Buradan,

$$\begin{aligned} b_M &= \left[\left(P'X_M\right)'\left(P'X_M\right)\right]^{-1}\left(P'X_M\right)'P'y_M \\ &= \left(X'_M\Omega_M^{-1}X_M\right)^{-1}X'_M\Omega_M^{-1}y_M \end{aligned} \tag{1.7}$$

şeklinde bilinen bir Ω_M ile en iyi lineer yansız tahmin edici (BLUE) olan Aitken Genelleştirilmiş EKK (GEKK) tahmin edicisi elde edilir. Burada ise Mixed tahmin edici olarak adlandırılacaktır.

$$X'_M \Omega_M^{-1} X_M = \left[X'R'\right] \begin{bmatrix} \frac{1}{\sigma^2} I_T & 0 \\ 0 & \frac{1}{\sigma_v^2} \Omega^{-1} \end{bmatrix} \begin{bmatrix} X \\ R \end{bmatrix}$$

$$= \frac{1}{\sigma^2} X'X + \frac{1}{\sigma_v^2} R'\Omega^{-1} R$$

olmasından dolayı, $\lambda = \sigma^2 / \sigma_v^2$ şeklinde varyansların oranı olmak üzere b_M 'nin farklı bir formu ise,

$$b_M = \left[\frac{1}{\sigma^2} X'X + \frac{1}{\sigma_v^2} R'\Omega^{-1} R \right]^{-1} \left[\frac{X'y}{\sigma^2} + \frac{R'\Omega^{-1} r}{\sigma_v^2} \right]$$

$$= \left(X'X + \lambda R'\Omega^{-1} R \right)^{-1} \left(X'y + \lambda R'\Omega^{-1} r \right)$$

$$= \left(X'X + \lambda R'\Omega^{-1} R \right)^{-1} \left(X'Xb + \lambda R'\Omega^{-1} b_p \right) \qquad (1.8)$$

şeklindedir. Burada $b = \left(X'X\right)^{-1} X'y$, (1.1) modelindeki β'nın EKK tahmin edicisi ve $b_p = \left(R'\Omega^{-1}R\right)^{-1} R'\Omega^{-1} r$ ise (1.2) kısıtındaki β'nın önbilgi tahmin edicisidir. (1.8)' deki yapı b_p önbilgi tahmin edicisi ile b örneklem tahmin edicisinin ağırlıklandırılmış matris kombinasyonu şeklinde mixed tahmin ediciyi ifade etmektedir. Yani,

$$W_1 = \left(X'X + \lambda R'\Omega^{-1} R \right)^{-1} X'X$$

$$W_2 = \lambda \left(X'X + R'\Omega^{-1} R \right)^{-1} R'\Omega^{-1} R$$

olmak üzere

$$b_M = W_1 b + W_2 b_p \qquad (1.9)$$

yapısındadır.

2. Bayesyen Regresyon Modeli

Bayesyen regresyon analizinin karakteristiğini yansıtan özellik, analizde ön bilgiye yer verilmesidir. Herhangi bir örneklem bilgisi gözlemlemeden önce elde edilen bilgi varsa dağılıma ön (önsel) dağılım denir. Ön dağılım ve örneklem bilgisinden elde edilen dağılıma ise son (sonsal) dağılım denir. Bayesyen yaklaşımda, denemeler yapılmadan önce parametreye ilişkin sahip olunan ön bilgi ön olasılık yoğunluk fonksiyonu sayesinde analize dahil edilir. Ancak ön bilgi her aşamada mevcut olmayabilir veya farklı seviyelerde ön bilgi olabilir. Bundan dolayı ön dağılımın oluşturulması bilgi veren ve bilgi vermeyen ön olasılık yoğunluk fonksiyonu şeklinde iki başlık halinde incelenebilir. Bayesyen yaklaşımda $\beta_0, \beta_1, \ldots, \beta_p$ ve σ parametreleri birer rasgele değişkendir ve olasılık dağılımları vardır.

ϕ, β ve σ'yı içeren parametre vektörü, y, $f(y/\phi)$ ortak olasılık yoğunluk fonksiyonuna sahip örneklem gözlemlerinin bir vektörü olsun. $f(y/\phi)$, β ve σ için benzerlik (likelihood) fonksiyonuna benzer ve β ve σ hakkındaki tüm örneklem bilgisini içerir. ϕ bir rasgele vektör olarak alınırsa,

$$h(\phi, y) = f(y/\phi)g(\phi) = g(\phi/y)f(y) \tag{2.1}$$

eşitliği yazılabilir. Burada h, ϕ ve y'nin ortak olasılık yoğunluk fonksiyonu, g, ϕ'nın olasılık yoğunluk fonksiyonu, f, y'nin olasılık yoğunluk fonksiyonunu göstermektedir. (2.1) ile verilen eşitlik yeniden düzenlenirse,

$$g(\phi/y) = \frac{f(y/\phi)g(\phi)}{f(y)} \tag{2.2}$$

olarak yazılabilir. Bu ifade Bayes Teoremi olarak bilinmektedir. $g(\phi/y)$, y örneklemi elde edildikten sonra ϕ ile ilgili tüm bilgiyi özetlediğinden dolayı, ϕ için son olasılık yoğunluk fonksiyodur. $g(\phi)$, ϕ için ön olasılık yoğunluk fonksiyonudur ve ϕ ile ilgili örneklem olmayan bilgiyi özetler. $f(y)$, ϕ'ya göre sabit olarak ele alınabilir ve $f(y/\phi)$, $L(\phi/y)$ şeklinde benzerlik fonksiyonu şeklinde yazılabilirse, (2.2) eşitliği,

$$g(\phi/y) \propto L(\phi/y)g(\phi) \tag{2.3}$$

olarak ifade edilir. (1.3) ile verilen genel lineer regresyon modeline Bayesyen kuralı uygulandığı zaman analizin temelini oluşturan süreç,

Son oyf $\propto$ Ön oyf $\times$ Benzerlik Fonksiyonu

$$P(\beta,\sigma/y)\propto P(\beta,\sigma)L(\beta,\sigma/y) \tag{2.4}$$

şeklindedir. $y'=(y_1,y_2,\dots y_n)$, bilinmeyen μ ortalamalı, bilinen $\sigma^2=\sigma_0^2$ varyanslı normal bir kitleden çekilmiş n birimlik bağımsız gözlemler olsun. μ parametresinin son dağılımı (2.4)'den,

$$p(\mu|y,\sigma_0^2)\propto p(\mu)p(y|\mu,\sigma_0^2)$$

(2.5)

şeklinde elde edilir. Burada $p(\mu|y,\sigma_0^2)$, verilen bir y örneklem bilgisi ve bilindiği varsayılan σ_0^2 varyans ile μ parametresinin son olasılık yoğunluk fonksiyonu olup, $p(\mu)$ ise ön olasılık yoğunluk fonksiyonudur. $p(y|\mu,\sigma_0^2)$ ise bilinmeyen μ parametresinin $\prod_{i=1}^{n} p(y_i|\mu,\sigma_0^2)$ şeklinde benzerlik fonksiyonudur. Yani,

$$p(y|\mu,\sigma_0^2)=(2\pi\sigma_0^2)^{-n/2}\exp\left[-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(y_i-\mu)^2\right]$$

$$=(2\pi\sigma_0^2)^{-n/2}\exp\left[-\frac{1}{2\sigma_0^2}\left[vs^2+n(\mu-\hat{\mu})^2\right]\right] \tag{2.6}$$

olarak ifade edilebilir. Burada $v=n-1$, $\hat{\mu}=(1/n)\sum_{i=1}^{n} y_i$ örneklem ortalaması $s^2=(1/v)\sum_{i=1}^{n}(y_i-\hat{\mu})^2$ olan örneklem varyansıdır. μ için bir ön dağılım fonksiyonu göz önünde bulundurmak istenirse, bu parametre hakkındaki ön bilginin tek değişkenli normal dağılıma sahip ise,

$$p(\mu)=\frac{1}{\sqrt{2\pi}\sigma_a}\exp\left[-\frac{1}{2\sigma_a^2}(\mu-\mu_a)^2\right] \tag{2.7}$$

olarak ifade edilir. Burada μ_a ve σ_a^2 araştırmacı tarafından belirlenen ve onun elde ettiği başlangıç bilgilerine dayanan ön ortalama ve ön varyanstır.

Buradan Bayes teoremini (2.6)'daki benzerlik fonksiyonu ve (2.7)'deki ön oyf ile birleştirerek μ için son oyf,

$$p(\mu|y,\sigma_0^2) \propto p(\mu)p(y|\mu,\sigma_0^2)$$

$$\propto \exp\left\{-\frac{1}{2}\left[\frac{(\mu-\mu_a)^2}{\sigma_a^2}+\frac{n}{\sigma_0^2}(\mu-\hat{\mu})^2\right]\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left[\frac{\sigma_a^2+\sigma_0^2/n}{2\sigma_a^2\sigma_0^2/n}+\left(\mu-\frac{\hat{\mu}\sigma_a^2+\mu_a\sigma_0^2/n}{\sigma_a^2+\sigma_0^2/n}\right)^2\right]\right\} \quad (2.8)$$

olarak bulunur. Burada μ,

$$E(\mu)=\frac{\hat{\mu}\sigma_a^2+\mu_a\sigma_0^2/n}{\sigma_a^2+\sigma_0^2/n}=\frac{\hat{\mu}(\sigma_0^2/n)^{-1}+\mu_a(\sigma_a^2)^{-1}}{(\sigma_0^2/n)^{-1}+(\sigma_a^2)^{-1}} \quad 2.9)$$

ortalamalı ve

$$Var(\mu)=\frac{\sigma_a^2\sigma_0^2/n}{\sigma_a^2+\sigma_0^2/n}=\frac{1}{(\sigma_0^2/n)^{-1}+(\sigma_a^2)^{-1}} \quad (2.10)$$

varyanslı, normal dağılımlı bir son dağılımdır.

2.1. Bayesyen Regresyon Modelinde Kullanılan Bazı Ön Dağılım Çeşitleri

Ön dağılım, araştırmacının elindeki bilgileri analize yansıtma aracıdır. Ön dağılımın seçimi araştırmacının bilgi ve düşüncesine bağlıdır. Ön dağılımın seçiminde veya oluşturulmasında çok çeşitli sınıflandırmalar yapılabilir. Örneğin, Modern Parametrik Bayesciler, tasarlanmış özelliklere sahip eşlenik ön dağılım belirlemeyi uygun görürken, Subjektif Bayesciler çoğunlukla uzman görüşünden elde edilen bilgiye göre ortaya çıkarılan ön dağılımı tercih ederler (Ekici, 2005). Burada genel lineer modelde eşlenik ön dağılım ve belirsiz ön dağılım çeşitleri ele alınacaktır. Eşlenik ön dağılım, ön dağılım ile benzerlik fonksiyonu birleştirildiğinde elde edilen son dağılımının da ön dağılım ile aynı fonksiyon formunda olduğu bir dağılımdır. Eğer parametreler hakkında hiçbir bilgi yoksa bilgi verici olmayan ya da belirsiz ön dağılım şeklinde adlandırılan ön dağılım kullanılmaktadır.

2.1.1. Eşlenik Ön Dağılım

$u \sim N\left(0, \sigma^2 I\right)$ olmak üzere y, $X\beta$ ortalama vektörlü, $\sigma^2 I$ varyans kovaryans matrisli çok değişkenli normal dağılıma sahip olsun. Yani $y \sim N\left(X\beta, \sigma^2 I\right)$. y'nin yoğunluk fonksiyonu (verinin olasılığı),

$$f\left(y/\beta,\sigma,X\right)=\frac{1}{(2\pi)^{T/2}\sigma^T}\exp\left[-\frac{1}{\sigma^2}\left(y-X\beta\right)'\left(y-X\beta\right)\right] \quad (2.11)$$

şeklinde yazılabilir. Bu da, y ve X üzerine verilmiş bir veri kümesi için β ve σ'ya bağlı benzerlik fonksiyonu olarak ifade edilebilir. Yani,

$$L(\beta,\sigma)=L(\beta,\sigma/y)=f\left(y/\beta,\sigma,X\right) \quad (2.12)$$

olur. Benzerlik fonksiyonundaki karesel formu,

$$\begin{aligned}(y-X\beta)'(y-X\beta)&=(y-Xb)'(y-Xb)+(\beta-b)'X'X(\beta-b)\\ &=SSE(b)+(\beta-b)'X'X(\beta-b) \quad (2.13)\end{aligned}$$

şeklinde yazabiliriz. Böylece benzerlik fonksiyonu,

$$L(\beta,\sigma)=\frac{1}{(2\pi)^{T/2}\sigma^T}\exp\left[\frac{1}{2\sigma^2}\left\{SSE(b)+(\beta-b)'X'X(\beta-b)\right\}\right]$$

$$\propto$$

$$\sigma^{-T}\exp\left[\frac{1}{2\sigma^2}\left\{SSE(b)+(\beta-b)'X'X(\beta-b)\right\}\right] \quad (2.14)$$

Burada $\propto$ orantılılığı ve sabitlerin atılarak ifadenin basitleşebileceğini göstermektedir. İki tür durumla karşılaşılabilir.

Durum 1. (Bilinen σ ile) Bu durumda (2.14) eşitliği,

$$L(\beta)\propto \exp\left[-\frac{1}{2\sigma^2}(\beta-b)'X'X(\beta-b)\right]$$

olarak yazılabilir. Çünkü $SSE(b)/2\sigma^2$ ifadesi β'yı içermemektedir. Benzerlik fonksiyonunun bu yapısından dolayı β'nın ön dağılımı için normal dağılım kullanılabileceği önerilmiştir. β_0 önsel ortalama vektörü ve $\sigma^2\Omega$ önsel varyans kovaryans matrisli çok değişkenli normal dağılım,

$$p(\beta)=\frac{1}{\left(2\pi^{p/2}\left|\sigma^2\Omega\right|^{1/2}\right)}\exp\left[\frac{1}{2\sigma^2}(\beta-\beta_0)'\Omega^{-1}(\beta-\beta_0)\right]$$

$$\propto \exp\left[-\frac{1}{2\sigma^2}(\beta-\beta_0)'\Omega^{-1}(\beta-\beta_0)\right] \qquad (2.15)$$

şeklindedir. Daha sonra Bayes teoremi kullanılarak sonsal yoğunluğu oransal ön dağılım ile benzerlik fonksiyonun çarpımı şeklindedir. Yani,

$$p(\beta/y) \propto p(\beta)L(\beta)$$

$$\propto$$

$$\exp\left[-\frac{1}{2\sigma^2}\left\{(\beta-b)'X'X(\beta-b)+(\beta-\beta_0)'\Omega^{-1}(\beta-\beta_0)\right\}\right] \qquad (2.16)$$

olarak ifade edilir. $(\beta-b)'X'X(\beta-b)+(\beta-\beta_0)'\Omega^{-1}(\beta-\beta_0)$ ifadesi,

$$=\beta'\left(X'X+\Omega^{-1}\right)\beta-2\beta'\left(X'Xb+\Omega^{-1}\beta_0\right)+c$$

$$=\left(\beta-b^*\right)'\Omega^{*-1}\left(\beta-b^*\right)+c \qquad (2.17)$$

olarak da yazılabilir. Burada $c=b'X'Xb+\beta_0'\Omega^{-1}\beta_0$ olarak kısaltılmıştır ve β'yı içermediği görülmektedir. Ayrıca burada,

$$b^*=\left(X'X+\Omega^{-1}\right)^{-1}\left(X'Xb+\Omega^{-1}\beta_0\right)$$

$$\Omega^*=\left(X'X+\Omega^{-1}\right)^{-1} \qquad c^*=c-b'^*\Omega^{*-1}b^* \qquad (2.18)$$

şeklindedir ve bu ifadeler de β'yı içermemektedir. (2.18) kullanılarak sonsal yoğunluğu,

$$p(\beta/y)\propto\exp\left[-\frac{1}{2\sigma^2}\left(\beta-b^*\right)'\Omega^{*-1}\left(\beta-b^*\right)\right] \qquad (2.19)$$

şeklinde yazılabilir. Burada $(2\pi)^{-T/2}\left|\sigma^2\Omega^*\right|^{-1/2}$ normallik sabitidir. Benzer şekilde yoğunluk b^* ortalama vektörlü, Ω^* varyans kovaryans matrisli çok değişkenli normal dağılımlıdır. Bu sonsal yoğunlukta genellikle $\beta\sim N\left(b^*,\sigma^2\Omega^*\right)$ şeklinde ifade edilir. Önsel ve sonsalın normal olmasından dolayı, yukarıda β için verinin ön dağılımı eşlenik ön dağılım olarak adlandırılır.

Genel olarak ise,

$$b_1^* = \left(\frac{X'X}{\sigma^2} + \frac{\Omega^{-1}}{\sigma_\beta^2} \right)^{-1} \left(\frac{X'Xb}{\sigma^2} + \frac{\Omega^{-1}\beta_0}{\sigma_\beta^2} \right)$$

$$\Omega_1^* = \left(\frac{X'X}{\sigma^2} + \frac{\Omega^{-1}}{\sigma_\beta^2} \right)^{-1} \qquad (2.20)$$

olmak üzere ön dağılım $\beta \sim N\left(\beta_0, \sigma_\beta^2 \Omega\right)$ ise son dağılım $\beta \sim N\left(b^*, \sigma^2 \Omega_1^*\right)$ 'dır. Ayrıca $\sigma_\beta^2 = \sigma^2$ için, $b_1^* = b^*$ ve $\Omega_1^* = \Omega^*$ olur.

Durum 2. (Bilinmeyen σ ile) Bu durumda (2.14)'deki benzerlik fonksiyonuna yakın bir ön dağılım söz konusudur. $a > 0,\ l > 0$ olmak üzere eşlenik ön dağılım,

$$p(\beta,\sigma) = p(\beta/\sigma)p(\sigma) \propto$$

$$\sigma^{-(l+p)} |\Omega|^{-1/2} \exp\left[-\frac{1}{2\sigma^2} \left[(l-1)a^2 + (\beta - \beta_0)' \Omega^{-1} (\beta - \beta_0) \right] \right] \qquad (2.21)$$

şeklinde bir ortak dağılımdır. $l > 1$ ve $a > 0$ olmak üzere,

$$p(\sigma) \ \propto \ \sigma^{-1} \exp\left[\frac{(l-1)a^2}{2\sigma^2} \right]$$

ters gama fonksiyonu ile,

$$p(\beta/\sigma) \propto |\Omega|^{-1/2} \sigma^{-p} \exp\left[\frac{1}{2\sigma^2} (\beta - \beta_0)' \Omega^{-1} (\beta - \beta_0) \right]$$

koşullu dağılımı ise verilen bir σ ile β'nın çok değişkenli dağılımıdır. $p(\beta,\sigma) = p(\beta/\sigma)p(\sigma)$ eşitliği ortak olasılık için ailesel ilişkinin benzer olduğu doğrulanabilir. $p(\beta,\sigma)$'nın $\sigma > 0$ için integrali alındığında β'nın marjinal dağılımı çok değişkenli student t olacaktır.

(2.14)'deki benzerlik fonksiyonu (2.20) ve (2.21)' deki ön dağılımlar ile çarpıldığında β ve σ'nın son dağılımı,

$$p(\beta,\sigma/y) \propto \sigma^{-n} \exp\left[\frac{1}{2\sigma^2}\left\{c_0 + \left(\beta - b^*\right)'\Omega^{*-1}\left(\beta - b^*\right)\right\}\right] \tag{2.22}$$

elde edilir. Burada $n = (T + l + p)$ ve $c_0 = SSE(b) + (l-1)a^2 + c^*$ 'dır. (2.22)'nin σ'ya göre integrali alındığında β'nın marjinal son dağılımı,

$$p(\beta/y) \propto \left[c_0 + \left(\beta - b^*\right)'\Omega^{*-1}\left(\beta - b^*\right)\right]^{-(n-1)/2} \tag{2.23}$$

olur. Bu da b^* ortalama vektörlü, $\dfrac{c_0}{n-p-1}\Omega^*$ varyans kovaryans matrisli çok değişkenli t dağılımının yoğunluk fonksiyodur.

Böylece, β'nın son yoğunluk fonksiyonu $\sim$ çok değişkenli $t\left(b^*, \dfrac{c_0}{n-p-1}\Omega^*\right)$ olur.

2.1.2. Belirsiz Ön Dağılım

β ve σ parametreleri hakkında bilgimiz olmadığı durumda β ve σ'nın ortak ön dağılımı,

$$\begin{aligned} p(\beta,\sigma) &= p(\beta/\sigma)p(\sigma) \\ &\propto \frac{1}{\sigma} \end{aligned} \tag{2.24}$$

şeklindedir. Burada $p(\beta/\sigma)$ bir sabit ve $p(\sigma)$, $1/\sigma$ ile orantılıdır. Bu da (2.21)'deki β ve σ'nın ön dağılımlarının varyanslarının $\Omega^{-1} \to 0$ ve $a \to 0$ olmasıyla elde edilebilir. (2.24)'deki ön dağılım, "belirsiz" ön dağılım ya da "uygunsuz" ön dağılım olarak bilinir. Uygunsuz denmesinin sebebi olasılık yoğunluk fonksiyonunun integrali alındığında sonsuz çıkar ve olasılıklar toplamının 1'e eşit olma aksiyomunu bozar. Ayrıca bu Jeffreys' in ön dağılımı olarak da bilinmektedir (Jeffreys, 1957). Belirsiz ön dağılım kullanarak son dağılım,

$$p(\beta,\sigma/y) \propto \sigma^{-(T+1)} \exp\left[-\frac{1}{2\sigma^2}\left\{SSE(b) + (\beta - b)'X'X(\beta - b)\right\}\right] \quad 2.25)$$

şeklinde elde edilir. Böylece verilen bir σ ile β'nın son dağılımı $\beta \sim N\left(b, \sigma^2 (X'X)^{-1}\right)$ olur. Ancak, σ'ya göre integral alındığında β'nın marjinal son dağılımı,

$$p(\beta/y) \propto \left[SSE(b) + (\beta - b)' X' X(\beta - b)\right]^{-T/2} \tag{2.26}$$

olur. Bu da b ortalama vektörlü, $\dfrac{SSE(b)}{T-p}(X'X)^{-1}$ varyans kovaryans matrisli çok değişkenli t dağılımıdır. σ hakkında yorum yapabilmek için ortak son dağılımın β'ya göre integrali alınmalıdır.

3. Bayesyen Tahmin Edici ve Mixed Tahmin Edici Arasındaki İlişki

β ve σ üzerine eşlenik ya da bilgi verici olmayan ön dağılımlar kullanıldığında β'nın marjinal son dağılımı çok değişkenli t dağılımı elde edilmiştir. Böylece β parametresi çok değişkenli t dağılımının olasılık yoğunluk fonksiyonunun bilinen özellikleri kullanılarak analiz edilebilir. Bu simetrik son yoğunluk fonksiyonunun ortalaması moduna eşittir ve parametrenin Bayes tahmin edicisini sağlamaktadır. Bayes tahmin edici karesel kayıp fonksiyonuna göre beklenen kaybın minimum olduğu bir tahmin edicidir (MELO). (2.26)' dan β ve σ hakkında belirsiz ön bilgi olduğu varsayılırsa β'nın Bayes tahmin edicisi, EKK ya da Maksimum Likelihood tahmin edicisi (MLE) ile aynıdır, yani $b = (X'X)^{-1} X'y$ olur. β ve σ üzerine eşlenik ön dağılım kullanılması durumunda ise β'nın Bayes tahmin edicisi (2.18)'den,

$$\begin{aligned} b^* &= \left(X'X + \Omega^{-1}\right)^{-1}\left(X'Xb + \Omega^{-1}\beta_0\right) \\ &= \left(X'X + \Omega^{-1}\right)^{-1} X'Xb + \left(X'X + \Omega^{-1}\right)^{-1}\Omega^{-1}\beta_0 \end{aligned} \tag{3.1}$$

şeklindedir. b^* tahmin edicisi, veriye ve β_0 önsel ortalamaya dayalı b'nin ağırlıklandırılmış matris kombinasyonudur. Ayrıca b ve β_0'ın ağırlık matrisleri sırasıyla b ve β'nın varyans ve kovaryans matrislerinin normalleştirilmiş tersleridir. $V(b) = \sigma^2 (X'X)^{-1}$ ve $V(\beta) = \sigma^2\Omega$ olduğundan $V(b)^{-1} + V(\beta)^{-1} = \dfrac{1}{\sigma^2}\left(X'X + \Omega^{-1}\right)$ olur.

Ayrıca (3.1)'deki Bayesyen tahmin edici b^*'daki ikinci parantezin içine $X'X\beta_0$ ekleyip çıkartma işlemi yapılarak,

$$b^* = \left(X'X + \Omega^{-1}\right)^{-1}\left[X'X(b - \beta_0) + \left(X'X + \Omega^{-1}\right)\beta_0\right] \tag{3.2}$$

$$= \beta_0 + \left(X'X + \Omega^{-1}\right)^{-1} X'X(b - \beta_0) \tag{3.3}$$

şeklinde yazılabilir. Ya da $X'X = X'X + \Omega^{-1} - \Omega^{-1}$ eşitliği kullanılarak,

$$b^* = \beta_0 + \left(I - \left(X'X + \Omega^{-1}\right)^{-1}\Omega^{-1}\right)\left(b - \beta_0\right) \tag{3.4}$$

elde edilebilir. (3.1) ve (3.4)'daki Bayesyen tahmin edici b^*'ın her iki formu da $R = I$, $\lambda = 1$ ve $r = \beta_0$ için mixed tahmin edici $\left(b_M\right)$ olarak bilinmektedir. Bunun anlamı ise $p\left(\beta/\sigma\right)$'nın β_0 ortalamalı, $\sigma^2\Omega$ varyans kovaryans matrisli çok değişkenli normal dağılımlı olduğudur. Böylece $E\left(v\right) = 0$ ve $E\left(vv'\right) = \sigma^2\Omega$ olmak üzere $\beta = \beta_0 + v$ olarak yazılabilir. Bu da $y = X\beta + u$ modelini $r = \beta_0$ ve $R = I$ olmak üzere $r = R\beta + v$ stokastik kısıtlar ile çözebileceğimizi göstermektedir. Ancak Bayesyen sonuçlar ile mixed regresyon sonuçları benzer olsalar bile yorumları farklıdır. Bayesyen modelinde β bir rassal değişken iken mixed regresyon modelinde değildir.

KAYNAKLAR

[1] Ekici, O., (2005). Bayesyen Regresyon ve Win-Bugs İle Bir Uygulama. İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi, İstanbul.

[2] Graybill, F. A., (1961). An Introduction to Linear Statistical Models. Mc Graw Hill, New York, 463p.

[3] Jeffreys, H., (1957). Scientific Inference. Cambridge University Press, 2th Ed. Cambridge, USA, 268p.

[4] Lindley, D. V., (1971). Bayesian Statistics. A Review (Philadelphia, Pa.: Society For Industrial and Applied Mathematics.

[5] Lindley, D. V., Smıth, A. F. M., (1972). Bayes Estimates for the Linear Model. Journal of the Royal Statistical Society, B Series, 1-41.

[6] Myers, R. H., (1990). Classical and Modern Regression with Application. Duxbury Press, California, 488p.

[7] Montgomery, D. C., Peck, E. A., (1992). Introduction to Linear Regression Analysis. John Wiley and Sons, New York, 526p.

[8] Theil, H., Goldberger, A.S., (1961). On Pure and Mixed Statistical Estimation in Economics, International Economic Review, 2:65-78.

[9] Vinod H. D., Ullah A., (1981). Recent Advances in Regression Methods. Marcel Dekker, Inc, 361p.