Y. Y. Ü. Fen Bilimleri Enstitüsü Dergisi
Yıl : 2010, Cilt : 15, Sayı : 1, Sayfa :46-53

Çok Seviyeli Poisson Regresyonu için PQL, MQL ve MCMC Yöntemlerinin Karşılaştırılması

Suna AKKOL[1], Hayrettin OKUT[1]

[1] YYÜ Ziraat Fakültesi Biyometri ve Genetik ABD

Özet: Bu çalışmada, çok seviyeli genelleştirilmiş doğrusal modellerde (Multilevel Genaralized Linear Model: Çok seviyeli GLM) kullanılan Marjinal Quasi Olabilirlik (MQL), Penalized Quasi Olabilirlik (PQL) ve Markov Chain Monte Carlo (MCMC) yöntemlerinin karşılaştırmalı olarak çalışılması amaçlanmıştır. Bu amaçla, üç seviyeli yapıya sahip Poisson dağılışı gösteren veriler kullanılmıştır. Hiyerarşik yapı ve veriler, benzetim (simülasyon) yoluyla türetilmiştir. Bunun için, MLwiN istatistik paket programı kullanılmıştır. Veri setindeki hiyerarşi 3 seviyeli olup, birinci seviyede bireyler, ikinci seviyede aileler ve üçüncü seviyede semtler yer almıştır. Veri setinde, bağımlı değişken olarak aylık bira kullanım sıklığı ve aylık bira kullanım sıklığı üzerine etkili olabilecek değişkenler bulunmaktadır.
Sonuç olarak bu çalışmada, türetilen veri seti için en uygun tahminleme yönteminin MCMC yöntemlerinden adaptive hibrit Metropolis-Gibbs yönteminin olduğu sonucuna varılmıştır. Zira çok seviyeli GLM analizi yapılırken MQL, PQL yöntemleri olabilirlik fonksiyonunun yerine yaklaşık olabilirlik fonksiyonunu kullanmaktadır. Buna rağmen MCMC yöntemleri olabilirlik fonksiyonunun kendisini değerlendirmekte ve bunun için Bayes yorumlamadan faydalanmaktadır.
Anahtar kelimeler: Çok seviyeli genelleştirilmiş doğrusal modeller, Marjinal Quasi Olabilirlik, Penalized Quasi Olabilirlik, MCMC yöntemleri, Gibbs Örnekleme, Metropolis-Hastings.

The Comparison of Methods Such As PQL, MQL and MCMC for Multilevel Poisson Regression

Abstract: This study was aimed for comparison three estimation methods such as Marginal Quasi Likelihood (MQL), Penalized Quasi Likelihood (PQL) and Markov Chain Monte Carlo (MCMC) in a multilevel structure for Generalized Linear Models (GLM). Therefore a Poisson distributed data set with three levels structure were simulated for comparisons of methods Hence, MLwiN statistical software was used for data simulation and construction of hierarchical structure. In three-level hierarchy, children were considered to be first, family second and neighborhood third level. Response variable was monthly beer using among family members. Explanatory variables that assumed to have effects on beer use behavior included in data set as well.
In conclusion, the simulation results proved that the adaptive Metropolis-Gibbs method performed better in terms of the goodness of fir criteria and diagnosis MQL and PQL methods are used approximate likelihood function instead of full likelihood function. However MCMC methods evaluate different methods using both Gibbs sampling and Metropolis Hastings sampling.
Key Words: Multilevel generalized linear models, Marginal Quasi Likelihood, Penalized Quasi Likelihood, MCMC methods, Gibbs sampling, Metropolis-Hastings

Giriş

Hiyerarşik yapıya sahip bir populasyon ile ilgilenildiği durumda çok seviyeli bir problemin varlığından söz edilir. Birimlerin farklı seviyelerde gruplandırılması ile elde edilen veriler hiyerarşik verilerdir. Hiyerarşik bir gruplandırma veya sınıflandırmanın söz konusu olduğu verilere ait modeller çok seviyeli modeller olarak bilinir. Üzerinde çalışılan veri setinin normal dağılış göstermemesi durumunda çok seviyeli genelleştirilmiş doğrusal modellerden (Genaralized Linear Model: GLM) bahsedilebilir ve genelleştirilmiş karışık doğrusal modellerden (Genaralized Linear Mixed Model: GLMM) içinde yer alır. Zira çok seviyeli bir yapıya sahip modelin uyumunun yapılması için GLMM kullanılmaktadır (Breslow ve Clayton, 1993; Langford ve ark.,1998; Zhou ve ark., 1999; Agresti ve ark., 2000). İki seviyenin olduğu çok seviyeli bir modelleme probleminde, ilk olarak en yüksek seviyeden örnek birimler alınır. Daha sonra mevcut birimlerden alt birimler örneklenir. Söz konusu alt birimler birinci seviye alt birimleridir. Buna göre birinci seviyedeki birimler genellikle tamamen bağımsız olmazlar (Hox, 1998; Goldstein, 1995; Agresti ve ark., 2000). Aynı okuldaki öğrencilerin seviye açısından birbirlerine benzer olmaları, buna örnek olarak verilebilir. Bu örnekte okul ikinci ve öğrenci birinci seviye olmaktadır.

Standart istatistik testleri gözlemlerin bağımsızlık varsayımına dayanmaktadır. Bu varsayımların yerine gelmemesi durumunda hatalar şansa bağlılığa atfedilmeyecek kadar büyük olur ve tahminlemeler ve hipotez testlerinde yanıltıcı sonuçlar elde edilir (Hox; 1998)

Normallik varsayımının geçerli olduğu karışık modellerde En Çok Olabilirlik (Maximum Likelihood: ML) tahminlerinin elde edilebilmesi için beklenenlerin maksimizasyonu (Epectation-Maximization: EM) algoritması tercih edilmektedir (Searle ve ark., 1992). GLMM'de, EM algoritmasının E aşamasında çözülmesi gereken eşitliğin analitik olarak değerlendirilmesi oldukça güçtür (Lin ve Breslow, 1996). Bu amaçla yaklaşık ML yöntemi kullanılır. Çünkü yaklaşık ML yöntemi, olabilirlik fonksiyonunun kendisinin yerine olabilirlik fonksiyonunun analitik yaklaşımını maksimize etmektedir (Agresti ve ark., 2000).

Yaklaşık ML yönteminde doğrusal olmayan fonksiyondaki parametrelerin tahmini için ilk olarak Taylor serisi açılımı kullanılarak model doğrusallaştırılır (Goldstein, 1991; Breslow ve Clayton, 1993; Rodriguez ve Goldman, 1995; Goldstein ve Rasbash, 1996; Browne 1998). Daha sonra quasi-olabilirlik (McCullagh ve Nelder, 1983; Foulley ve Im, 1993) tahminlemesi kullanılarak

* AKKOL, Suna. 2004 (Doktora Tezi)

parametre tahminleri yapılır (Goldstein, 1991; Goldstein ve Rasbash, 1996). Çok seviyeli GLM'de iki yaklaşık ML veya quasi-olabilirlik yönteminden bahsedilebilir. Bunlardan ilki, Goldstein (1991) tarafından önerilen ve daha sonra Breslow ve Clayton (1993) tarafından isimlendirilen Marjinal Quasi Olabilirlik (MQL) yöntemidir. Diğeri, Browne (1998)'nin bildirdiğine göre; Laird (1978) tarafından önerilmiştir. Önerilen bu yöntem, yine Breslow ve Clayton (1993) tarafından Penalized veya Predictive Quasi Olabilirlik (PQL) olarak isimlendirilmiştir.

İki yöntem arasındaki fark şudur; Modeli doğrusallaştırmak için Taylor açılımı kullanıldığı zaman PQL'deki doğrusal olmayan fonksiyonun doğrusal unsurlarına daha yüksek seviye hataları eklenir. Ancak MQL'de bu durum söz konusu değildir (Breslow ve Clayton, 1993; Goldstein, 1995; Goldstein ve Rasbash, 1996; Browne, 1998; Zhou ve ark., 1999). Diğer bir ifadeyle Taylor açılımında kullanılan fonksiyonda (H_t), sadece sabit etkiler yer alıyorsa MQL yöntemine ait, hem sabit etkiler hem de şansa bağlı etkiler yer alıyorsa PQL yöntemine ait tahminler elde edilir. Dolayısıyla MQL ile herhangi bir kovaryetteki bir birim değişime göre cevap değişkenindeki değişim tahminlenir. Oysa PQL ile herhangi bir kovaryetteki bir birim değişim, verilen herhangi bir seviye 2 birimi için tahminlenir (Goldstein, 1995; Goldstein ve Rasbash, 1996).

Bilgisayar teknolojisinin ilerlemesi ile birlikte Bayes istatistik alanın önemi artmıştır. Dolayısıyla daha önce uygulaması imkansız olan teknikler etkin bir şekilde kullanılmaya başlanmıştır. Bu gelişmeler, çok seviyeli modellerin analizi için MCMC yöntemlerine dayanan tam (full) Bayes analizinin kullanımına olanak sağlamıştır. Çok seviyeli modellerde MCMC tahminlerinin kullanılmasına ilişkin ilk ciddi çalışma Browne (1998) tarafından yapılmıştır. Browne (1998) yaptığı çalışmada, normal dağılış ve binomiyal dağılış gösteren cevap değişkenleri için MCMC yöntemlerini kullanarak parametre tahminlerini yapmıştır.

Bayes modelleme için pratikte kullanılan MCMC yöntemlerini iki başlık altında incelemek mümkündür. Bunlar; 1) Gibbs Örnekleme ve 2) Metropolis-Hastings Örnekleme'dir (Browne, 1998; 2003; Hox, 1998; Goldstein ve ark., 2002b; Rasbash ve ark., 2000). Gibbs örnekleme genellikle parametrelerin şartlı posteriorları standart bir forma (Normal dağılış gibi) sahip olduğu durumda kullanılır. Eğer şartlı posteriorlar standart bir forma sahip değilse Metropolis Hastings (MH) örneklemeye dayanan MCMC yöntemi kullanılır (Browne, 1998; Browne ve Draper, 2000; Browne ve Draper, 2001). Bu çalışmada MCMC içinde yer alan adaptive Hibrit Metropolis-Gibbs yöntemi kullanılmıştır (Detaylar için bakınız Ek 1).

Bu çalışmanın amacı çok seviyli Poisson regresyonda parametre tahminlerini elde edilirken kullanılan tahminleme yöntemlerinin (MQL, PQl ve MCMC) teorik özellikleri verilerek birbirlerine olan üstünlüklerinin ortaya koyulmasıdır.

Materyal Ve Yöntem

Materyal: Bu çalışmada kullanılan veriler, Duncan ve ark.(2002), tarafından yapılan bir çalışmada kullanılan veriler esas alınarak, benzetim (simülasyon) tekniği ile türetilmiştir. Söz konusu çalışmada, aylık bira kullanım sıklığı üzerine etkili olabileceği düşünülen değişkenler yer almaktadır. Bu değişkenler, ailenin çocuklara rehberlik derecesi (AÇRD), ebeveynlerin içkiye karşı hoşgörü derecesi (EİKHD), yaş, cinsiyet ve bireyin risk alma durumu olmaktadır. Bu çalışmada üç seviyeli bir model tasarlanarak verilerin türetilişi yapılmıştır. Türetilen verilerin birinci seviyesinde bireyler, ikinci seviyesinde aileler ve üçüncü seviyesinde ise semtler olacak şekilde hiyerarşik yapı oluşturulmuştur. Diğer bir ifade ile, hiyerarşik yapı, semtler içinde aileler ve aileler içinde bireyler olacak şekilde düzenlenmiştir. Söz konusu hiyareşide bira kullanım sıklığının semt-aile-birey şeklindeki değişimi incelenmiştir.

Benzetim için SPSS ve MLwiN istatistik paket programları kullanılmıştır. Benzetim yapılırken tasarlanan üç seviyeli hiyerarşik modelde bira kullanımının semtten semte, aileden aileye ve bireyden bireye değiştiği garanti edilmiştir. İkinci seviyeye ($u \sim N(0,\sigma^2_u)$ ve üçüncü seviyeye ($v \sim N(0,\sigma^2_v)$ ilişkin şansa bağlı etkilerin tamamı benzetim ile çoğaltılmıştır. Bu amaçla ortalaması sıfır ve standart sapmaları için farklı değerler kullanılmıştır. Elde edilen katsayılar modelde yerine koyularak SPSS'de RV.POİSSON(mean) komutu ile şansa bağlı Poisson dağılımına sahip veri türetilmiştir. Elde edilen sonuçlar bireylere ait aylık bira kullanımını vermiştir. Cinsiyet, yaş ve bireyin risk alma durumlarının bireyden bireye değiştiği için, bu değişkenler sadece bireylerin bulunduğu birinci seviyeye dahil edilmiştir. Bu nedenle söz konusu olan bu değişkenler için basit regresyon katsayıları kullanılarak benzetim yapılmıştır.

Hiyerarşik yapı için benzetim, 25 semt ve her semt içinde 50 aile (50±1) ve her ailede ortalama 2 birey (2±1) olacak şekilde tasarlanarak yapılmıştır. Bu duruma göre toplamda, 25 semt, 1284 aile ve 2579 birey elde edilmiştir. Üzerinde çalışılan veri seti, 2579 bireye ait bira kullanım sıklığı olan cevap değişkenini ve bunun üzerine etkili olabileceği düşünülen değişkenleri (AÇRD, EİKHD, bireyin risk alma durumu, yaş ve cinsiyet) içermektedir.

Parametre tahminleri elde edilirken MLwiN istatistik paket programı kullanılmıştır.

Yöntem

Çok-seviyeli GLM: Doğrusal modeller için geçerli olan varsayımların yerine gelmemesi ve şans değişkeninin bir sınıflandırma göstermesi durumunda, çok seviyeli genelleştirilmiş doğrusal bir model kullanılmaktadır.

y_{ij}, Poisson dağılışından alınmış şans değişkeni olsun, $y_{ij} \sim Poisson(\lambda_{ij})$. Bu şans değişkeni için iki seviyeli model eşitliği, genel olarak,

$$\lambda_{ij} = \log^{-1}(X_{ij}\beta_j) = \exp(X_{ij}\beta_j) \quad (1)$$

olur (Goldstein, 1995; Barbosa ve Golstein, 2000; Golstein ve ark., 2002b). Burada y_{ij} gözlenen cevap değişkenleri için $(y_{ij}|\lambda_{ij})$, Poisson (λ_{ij}) ve $\log^{-1}$ ifadesi ters bağlantı fonksiyonu olmaktadır. Alt indislerden *i* ikinci seviye içindeki birinci seviyedeki gözlemlerin ve *j* ikinci sınıftaki gözlemlerin sayısı olmaktadır. Buna göre iki seviyeli bir model için

$$\lambda_{ij} = \exp(X_{ij}\beta_j + u_j) \quad (2)$$

eşitliği yazılır. Çok seviyeli genelleştirilmiş doğrusal bir model için daha genel bir ifade,

$$\log(\lambda_{ij}) = \eta = X\beta + Zu \quad (3)$$

olur (Goldstein, 1991; Breslow ve Clayton, 1993; Rodriguez ve Goldman, 1995; Browne ve Draper, 2001). Burada *X*, sabit (β), *Z* ise şansa bağlı (*u*) etkiler için desen matrisleri ve η şartlı doğrusal tahminleyicidir.

Marjinal quasi olabilirlik (MQL) ve penalized quasi olabilirlik (PQL) yöntemleri: Çok seviyeli genelleştirilmiş doğrusal modeller, bilinen şansa bağlı hata terimi haricinde seviyelerden kaynaklanan şansa bağlı terimlere de sahiptirler. Doğrusal olmayan iki seviyeli bir model için eşitlik;

$$\lambda_{ij} = \pi_{ij} = f\left(X_{ij}\beta + Z_{ij}^{(2)}u_j + Z_{ij}^{(1)}e_{ij}\right) \qquad (4)$$

verilir (Goldstein, 1995; Browne ve Draper; 2001). Burada $f(l)$ doğrusal olmayan bir özelliğe sahiptir. İlk olarak model doğrusallaştırılır ve daha sonra quasi-olabilirlik (McCullagh ve Nelder, 1983; Foulley ve Im, 1993) tahminlemesi uygulanır (Goldstein ve Rasbabh, 1995). Zira olabilirlik fonksiyonunun değerlendirilmesindeki zorluklar bu alanda çalışanların, yaklaşık quasi olabilirlik çözümlerini tercih etmesine neden olmuştur (Rodriguez ve Goldman, 1995).

Yukarıda da belirtildiği gibi modelin sabit kısmı için geçerli olan tahmin etrafında birinci sırada Taylor serisi açılımı ve modelin şansa bağlı kısmı için 0 (sıfır) etrafında ikinci sırada Taylor serisi açımı kullanılarak model doğrusallaştırılır. Daha sonra quasi olabilirlik yöntemleri uygulanarak tahminleme yapılır. *(t+1)*'inci iterasyonda elde edilen eşitlik için $f(.)$ aşağıdaki şekilde verilir (Goldstein, 1995; Goldman ve Rashbash, 1996; Browne ve Draper, 2001; Langford ve ark., 1998).

$$\begin{aligned} f(H_{t+1}) &= f(H_t) + X_{ij}(\beta_{t+1} - \beta_t)f'(H_t) \\ &+ \left(Z_{ij}^{(2)}u_j + Z_{ij}^{(1)}e_{ij}\right)f'(H_t) \\ &+ \frac{1}{2}\left(Z_{ij}^{(2)}u_j + Z_{ij}^{(1)}e_{ij}\right)f''(H_t) \end{aligned} \qquad (5)$$

Yukarıda verilen eşitlik modelin sabit ve şansa bağlı kısımlarını bir arada bulundurur. Modelin sabit kısmı bu eşitliğin ilk iki terimidir. Şansa bağlı kısmı üçüncü ve dördüncü terimlerdir. Bu eşitlikte, modelin şansa bağlı kısmı için hem birinci hem de ikinci sırada Taylor serisi açılımı kullanılır. Eşitlikteki üçüncü terim, Goldstein (1991) tarafından önerilen terimdir ve birinci sırada düzeltmeyi içerir. Dördüncü ve son terim ikinci sırada düzeltmeyi içerir ve bu terimin kullanılması ile daha sapmasız tahminlemeler elde edilir (Goldstein ve Rasbash, 1996).

Bu amaçla, 5 numaralı eşitlikteki H_t'nin belirlenmesi gerekir. Zira, şansa bağlı etkilere ilişkin tahminlemelerin yapılabilmesi için H_t'nin birinci ve ikinci türevlerine ihtiyaç duyulur. H_t'nin farklı iki seçimi vardır (Goldstein, 1995; Goldstein ve Rasbash, 1996). Bunlar,

a) $H_t = X_{ij}\beta_t$

b) $H_t = X_{ij}\beta_t + Z_{ij}^{(2)}\hat{u}_j + Z_{ij}^{(1)}\hat{e}_{ij}$

a'nın seçimi ile Taylor serisi açılımı için yalnızca sabit kısma ilişkin parametreler kullanılır. Dolayısıyla modeldeki sabit kısım için ilgilenilen tahminler etrafında Taylor serisi açılımı kullanılmaktadır. Bu yöntem, Goldstein (1991)'in teklif ettiği yöntemin kendisidir (Goldstein ve Rasbash, 1996). Bu yöntem, Breslow ve Clayton (1993) tarafından MQL olarak isimlendirilmiştir. Eğer b seçilmiş ise o zaman Taylor serisi açılımı, tahminlenmiş olan şansa bağlı etkiler için kullanılır ve bu yöntem PQL olarak isimlendirilir (Breslow ve Clayton, 1993).

MQL ve PQL algoritmasının sırası, doğrusallaştırmanın temelini oluşturan Taylor serisinde kaç tane açılımın kullanıldığını ifade eder. İkinci sıraya kadar yapılan Taylor serisi açılımı 5 numaralı eşitlikte verilmiştir. İkinci sırada Taylor serisi açılımının kullanılması ile H_t'nin seçimine göre MQL_2 ve PQL_2 tahminleri yapılır (Yang, 1997; Browne ve Draper, 2001).

Markov zinciri Monte Carlo (MCMC) yöntemleri: Genellikle çok seviyeli modellerde ilgilenilen parametre sayısı kadar bilinmeyen parametre vardır. Bayes yorumlamada parametrelerin ortak posterior dağılışı bulunması gerekir. Daha sonra bu dağılışın sahip olduğu uzaydan örnekleme yapılır. Ortak posterior dağılışın bulunabilmesi, çok boyutlu integrasyonların çözümünü gerektirir ve genellikle bu integrasyonun çözümü zor olmaktadır (Rasbash ve ark., 2000; Browne, 2005; Goldstein ve ark., 2002b). Her ne kadar ortak posterior dağılıştan doğrudan örnekleme yapmak mümkün olmasa bile bunun bir alternatifi vardır. Bu alternatif parametrelerin şartlı posterior dağılışının kullanılması ile örnek seti elde etmektir (Rasbash ve ark., 2000; Browne, 2005).

MCMC yöntemleri bu alternatif yaklaşımı kullanır (Browne ve Draper, 2000; Browne ve Draper, 2001; Browne, 2005). MCMC'de parametreler gruplara ayrılır ve sırasıyla her bir grup için şartlı posterior dağılıştan örnekler oluşturulur (Rasbash ve ark., 2000; Goldstein ve ark.,2002b; Browne, 2005). Her bir parametre için oluşturulan şartlı dağılışlardan sırasıyla yapılan örnekleme ortak posterior dağılıştan yapılan örneklemeye denk olmaktadır (Browne, 2001, 2003a; Rasbash ve ark., 2000; Goldstein ve ark., 2002b).

Bayes modelleme için pratikte kullanılan iki temel MCMC yöntemi vardır (Browne, 1998; Browne ve Draper, 2000; Goldstein ve ark., 2002b; Browne, 2005).

1) Gibbs Örnekleme
2) Metropolis -Hastings Örnekleme

Gibbs örnekleme genellikle, parametrelerin şartlı posteriorları standart bir forma (Normal dağılış gibi) sahip olduğu durumda kullanılır. Eğer şartlı posteriorler standart bir forma sahip değilse Metropolis Hastings (MH) yöntemi kullanılır.

Gibbs yöntemi kullanılarak kolaylıkla simüle edilemeyen şartlı dağılışlar söz konusu olduğunda MH kullanılmaktadır. Ancak Gibbs örneklemede sözü geçen 4 aşamadan (1- sabit etkiler, 2- seviye 2 hataları, 3- seviye 2 varyans matrisi ve 4-seviye 1 varyansı) sadece ilk ikisi, yani sabit etkiler (β'lar) ve hatalar (e hariç) kolay simüle edilememektedir. Bu nedenle kolaylıkla simüle edilen varyanslar için Gibbs samlping kullanılırken sabit etkiler ve hatalar için MH kullanılır (Browne, 1998; Browne ve Draper, 2000; Browne ve Draper, 2001). Bu, Hibrit Metropolis-Gibbs olarak bilinir. Farklı önerilen dağılışların kullanılmasına gore kendi içinde ikiye ayrılır. Bu çalışmada Hibrit Metropolis-Gibbs yöntemi I (Univariate Update) kullanılacaktır. Cevap değişkeninin Poisson dağılışına sahip olduğu bu çalışmada MCMC yöntemlerinden Hibrit Metropolis-Gibbs kullanılarak parametre tahminleri elde edildi.

Bulgular

Üzerinde çalışılan verideki hiyerarşinin üçüncü seviyesini semtler ikinci seviyesini aileler ve birinci seviyesini bireyler göstermektedir. Veri seti bu hiyerarşik yapı ile birlikte toplamda 2579 bireye ait aylık bira kullanım sıklığı ve bunun üzerine etkili olabilecek değişkenleri içermektedir. Bu değişkenlere ilişkin tanıtıcı istatistikler aşağıdaki tabloda özetlenmiştir.

Çizelge 1. Çalışmada kullanılan değişkenlere ilişkin tanıtıcı istatistikler

Değişkenler	N	$\bar{x} \pm S_{\bar{x}}$	Min.	Maks.	Dağılış formu
Bira kullanım sıklığı	2597	9.5111±4.0057	1.0000	27.0000	Poisson
AÇRD*	2597	4.2100±0.5765	2.2595	6.6099	Normal
EİKHD**	2597	10.149±2.2564	2.5151	19.5417	Normal
Risk alma durumu	2597	1.2868±0.3023	0.1804	2.3324	Normal
Yaş	2597	13.011±1.9913	7.0000	19.0000	Normal

*AÇRD : Ailenin çocuklara rehberlik derecesi
**EİKHD : Ebeveynlerin içkiye karşı hoşgörü derecesi

Çizelge 1'de yer almayan cinsiyet değişkeni binomiyal karakterli olup %48,7'si bayan ve geri kalan %51,3'ü erkeklerden oluşmaktadır. Bira kullanım sıklığı Poisson dağılışı gösterdiği için modellerde *log* bağlantı (link) fonksiyonu kullanılarak verilerin analizi yapılmıştır.

Bireyin aylık bira kullanım sıklığının Poisson dağılışına sahip olması nedeni ile parametre tahminleri *log* bağlantı fonksiyonu kullanılarak çok seviyeli GLM'e göre yapılmaktadır. Analizler yapılırken, üzerinde çalışılan veri setini açıklayabilecek en iyi model ve en iyi yöntem belirlenmeye çalışılmıştır. Çalışmaya başlarken tüm değişkenleri Bu nedenle tüm değişkenleri modelde bulunduğu eşitlik aşağıdaki gibi olur.

$$\log(\lambda_{ijk}) = \beta_{0jk} + \beta_{1jk}x_{ijk} + \beta_{2jk}x_{ijk} + \beta_{3jk}x_{ijk} + \beta_{4jk}x_{ijk} + \beta_{5jk}x_{ijk}$$

Burada

$$\beta_{0jk} = \beta_0 + u_{0jk} + v_{0k} \qquad (6)$$

$$\beta_{1jk} = \beta_1 + u_{1jk} + v_{1k}, \qquad (7)$$

$$\beta_{2jk} = \beta_2,\ \beta_{3jk} = \beta_3,\ \beta_{4jk} = \beta_4 \text{ ve } \beta_{5jk} = \beta_5$$

olmaktadır. Katsayılar sırasıyla β_0 (intercept), β_1: AÇRD, β_2: EİKHD, β_3: bireyin risk alma durumu, β_4: bireyin yaşı ve son olarak β_5: bireyin cinsiyetini ifade etmektedir. Ayrıca 6 numaralı eşitlikdeki u_{0jk}, β_0 etrafında ikinci seviyeye ait şansa bağlı değişim miktarını ve v_{0k}, β_0 etrafında üçüncü seviyeye ait şansa bağlı değişim miktarını göstermektedir. Bir sonraki 7 numaralı eşitliğin açılımındaki u_{1jk}, β_1 etrafındaki ikinci seviyeye ait şansa bağlı değişim miktarını ve v_{1k} ise β_1 etrafındaki üçüncü seviyeye ait şansa bağlı değişim miktarını göstermektedir. Kısacası alt indis *0* başlangıç (intercepti) ve alt indis *1* eğimi (slope) ifade etmektedir.

Üzerinde ilk çalışılan bu modelde EİKDH ve cinsiyet değişkeninin önemsiz olduğu görüldü. Dolayısıyla bu değişkenler modelden çıkarılarak analizler tekrar yapıldı. Buna göre elde edilen yapılan MQL_1 ve PQL_2 analiz sonuçları aşağıdaki çizelgede verilmiştir.

Çizeige 2. MQL_1 ve PQL_2 analiz sonuçları

Parametreler	MQL_1 Tahmin (Std. Hata)	PQL_2 Tahmin (Std. Hata)
Sabit kısım		
β_0	3.631 (0.075)	3.656 (0.075)
β_1	-0.275 (0.013)	-0.283 (0.013)
β_3	-0.401 (0.023)	-0.401 (0.023)
β_4	0.021 (0.003)	0.021 (0.003)
Şansa bağlı kısım		
$\sigma^2_{v_0}$	0.000 (0.000)	0.000 (0.000)
$\sigma_{v_{01}}$	0.000 (0.000)	0.000 (0.000)
$\sigma^2_{v_1}$	0.000 (0.000)	0.000 (0.000)
$\sigma^2_{u_0}$	0.319 (0.103)	0.351 (0.107)
$\sigma_{u_{01}}$	-0.083 (0.025)	-0.091 (0.026)
$\sigma^2_{u_1}$	0.022 (0.006)	0.024 (0.007)

Yukarıdaki çizelgede, dört varyans ($\sigma^2_{v_0}$, $\sigma^2_{v_1}$, $\sigma^2_{u_0}$, $\sigma^2_{u_1}$) ve iki kovaryans ($\sigma_{v_{01}}$, $\sigma_{u_{01}}$) terimi yer almaktadır. Çizelge 2 incelendiğinde kullanılan iki yöntemden elde edile parametre tahmin değerleri arasındaki değişimlerin önemli olmadığı anlaşılmaktadır. Modeldeki sabit etkiler gibi ikinci seviyeye ait varyanslar ($\sigma^2_{u_0}$ ve $\sigma^2_{u_1}$) ve kovaryans ($\sigma_{u_{01}}$) önemli bulunmuştur. MQL_1 ve PQL_2 yöntemleri ile üçüncü seviye varyans-kovaryans matrisine ait tahminleme yapılamamıştır. Zira Taylor serisinin birinci sırada açılımını kullanan MQL (MQL_1)'in sapmalı tahminler vermesi üzerine (Rodriguez ve Goldman, 1995; Yang, 1997), Goldstein ve Rasbash (1996) ikinci sırada MQL (MQL_2) kullanılmış ve tahminlerdeki sapmada önemli bir değişim olmadığını bildirmişlerdir. Aynı araştırmacılar, birinci ve ikinci sırada PQL (sırasıyla PQL_1 ve PQL_2) yöntemi ile elde edilen tahminlerde iyileşme olduğunu ancak yine de sapmanın var olduğunu ortaya koymuşlardır (Goldstein ve Rasbash, 1996; Yang, 1997; Browne ve Draper, 2001). Çok seviyeli GLM'de yapılan çalışmalar, uzun bir süre quasi-olabilirlik tabanlı yöntemlerden MQL_1

ve PQL_2'nin yoğun olarak kullanıldığını göstermektedir (Goldstein, 1991; Breslow ve Calyton, 1993; Yang, 1997; Langford ve ark., 1998; Langford ve ark., 1999; Leyland ve McLeod, 2000; Blatchford ve ark., 2002; Goldstein ve ark., 2002a; Naderi ve Mace, 2003).

Bu çalışmanın bir sonraki aşamasında MCMC yöntemleri kullanılarak analizler yapılmış ve adaptive yöntemi kullanan iki ayrı MCMC analizine ait sonuçlar Çizelge 3'de verilmiştir.

Çizelge 3. MCMC[1] ve MCMC[2] analiz sonuçları

Parametre Tahminleri	MCMC[2] Tahmin (Std. Hata)	MCMC[3] Tahmin (Std. Hata)
Sabit kısım		
β_0	3.568 (0.069)	3.688 (0.076)
β_1	-0.279 (0.015)	-0.295 (0.021)
β_3	-0.394 (0.020)	-0.400 (0.022)
β_4	0.023 (0.002)	0.020 (0.003)
Şansa bağlı kısım		
$\sigma^2_{v_0}$	0.038 (0.017)	0.038 (0.018)
$\sigma_{v_{01}}$	-0.007 (0.005)	-0.007 (0.006)
$\sigma^2_{v_1}$	0.010 (0.003)	0.010 (0.003)
$\sigma^2_{u_0}$	0.246 (0.041)	0.359 (0.072)
$\sigma_{u_{01}}$	-0.067 (0.011)	-0.094 (0.018)
$\sigma^2_{u_1}$	0.019 (0.003)	0.025 (0.005)
Devians	3099.238	3077.434

[2] Zincir uzunlukları default olarak adaptive yöntem kullanılmıştır.
[3] 5000 zincir burn-in için ve 50000 zincir monitorig zincir uzunluğu için adaptive yöntem kullanılmıştır.

Üzerinde çalışılan son model için MCMC[1] ve MCMC[2] analiz sonuçlarının verildiği çizelgede dikkat çeken en önemli nokta, üçüncü seviye varyansları ve kovaryansına ilişkin tahminlemelerin de yapılmış olmasıdır. Bu durum diğer modeller için de verildiği gibi MCMC yöntemlerinin yaklaşık olabilirlik tabanlı yaklaşımları kullanan yöntemlere (MQL ve PQL) üstünlüğünü göstermektedir. Buna ek olarak, Çizelge 3 ve 4'de verilen kovaryans terimlerinden $\sigma_{v_{01}}$, üçüncü seviyeye ait olup β_0 ve β_1 arasındaki ilişkiyi göstermektedir. Benzer şekilde $\sigma_{u_{01}}$, ikinci seviyeye ait olup söz konusu seviye için β_0 ve β_1 arasındaki ilişkiyi göstermektedir.

Yukarıdaki çizelgede verilen MCMC[1] ve MCMC[2] yöntemlerinin her ikisi de adaptive hibrit Metropolis-Gibbs algoritmasının kullanan MCMC yöntemidir. Bu iki yöntem arasındaki fark, kullanılan zincir uzunluklarının farklı olmasından kaynaklanmaktadır. MCMC[1]'de parametre tahminleri programın default olarak kullandığı zincir uzunluğundan elde edilmiştir. Default olarak kullanılan zincir uzunluğu adaptive kısım için 5000, bunu takip eden burn-in periyodu için 500 ve daha sonraki görüntülenen zincir uzunluğu için 5000 olmaktadır. Görüntülenen bu zincir uzunluğu, parametre tahminlerinin dağılışına ilişkin istatistiksel özetlerin hesaplanması için kullanılmaktadır. MCMC[1] ile elde edilen parametre tahminlerine ait zincir durumuna (trajectory) ilişkin bilgiler derlendirildiğinde, özellikle sabit etkiler ve şansa bağlı etkilerden ikinci seviye varyanları ve kovaryansı için daha fazla zincir uzunluklarına ihtiyaç duyulduğu belirlenmiştir. Bu nedenle aynı yöntem için zincir uzunlukları 5000 adaptive (bu MLwiN'de default olmaktadır), 5000 burn-in ve 50000 görüntülenen zincir uzunluğu için çalıştırılmış ve analiz sonuçlar MCMC[3] şeklinde kodlanmıştır.

Çizelge 3 devianslar bakımından değerlendirildiğinde, MCMC[2] ile elde edilen devians değerinin MCMC[1] ile elde edilen devians değerinden yaklaşık 20 birim daha küçük olduğu görülmektedir. Burada devians modelin açıklayamadığı kısmı ifade etmektedir. Zira devians, MCMC yöntemleri kullanılırken dikkate alınan ciddi bir ölçüttür. Söz konusu çizelge, Çizelge 2 ile karşılaştırıldığında üçüncü seviye varyansları ve kovaryansına ait tahminlemenin yapılmış olduğu dikkat çekmektedir. Bu durum MCMC yöntemlerinin yaklaşık olabilirlik tabanlı yaklaşımlara olan üstünlüğünü ortaya koymaktadır.

ML tabanlı tahminleme yöntemleri iteratif olarak çalışır ve cevaba yakınsama oluncaya kadar devam eder. Yakınsama için bir ölçüt kullanılır ve bu ölçüt sağlandığı zaman yakınsama gerçekleşmiş olur. Bu ölçüt elde edilen son tahmin ile bir önceki tahmin değeri arsındaki farktır. Bir Markov zincirinin yakınsaması, bu yöntemlerden farklıdır. Markov zinciri yöntemlerinde bir tahmine yakınsamak yerine bir dağılışa yakınsama söz konusudur. Bu dağılış ise ilgilenilen ortak posterior dağılıştır. Bir dağılışa yakınsama konusunda burn-in periyodu, zincir karışımının iyi olup olmaması gibi durumlar değerlendirilir ve bunun için birçok ölçüt vardır.

Tartışma Ve Sonuç

Hiyerarşik karakterli çok düzeyli modellerde her seviyedeki bireylerin cevap değişkeni bakımından birbirlerine benzer özellik göstermeleri beklenmektedir. Bu nedenle aynı semt içindeki bireylerin birbirlerine benzer özellikler göstermesi gibi, aynı aile içindeki bireylerde birbirlerine benzer özellikler göstermesi beklenmektedir. Dolayısıyla birinci seviyedeki bireylere ait cevap değişkenleri arasında bağımsızlık varsayımını yerine getirilmemektedir. Bu tip veriler söz konusu olduğunda, bilinen en küçük kareler (OLS: Ordinary Least Squares) yöntemi ile yapılan analizlere ait parametre tahminleri sapmalı olup, sonuçlara güvenilmemektedir. Zira veriler arasında bulunan bağımlılık, bu yöntemler tarafından göz ardı edilmektedir. Veri setinin sahip olduğu bu doğal hiyerarşiyi, çok seviyeli modeller dikkate almaktadır.

Çalışmada yapılan MQL_1 ve PQL_2 analizleri sapmalı tahminler üretmiştir. Bu yöntemlerin sapmalı tahminler üretmesi Breslow ve Clayton (1993), Goldstein ve Rasbash (1996), Yang (1997), Browne ve Draper'in (2001) çalışmaları ile paralellik göstermektedir. Zira, çok seviyeli GLM'de kullanılan MQL_1 ve PQL_2 yöntemleri, çoğu zaman sapmalı tahminler vermektedir (Breslow ve Clayton, 1993; Goldstein, 1995; Yang, 1997; Sutradhar ve Rao, 2000; Raudenbash ve ark., 2000; Browne ve Draper, 2001; McBroom, 2001; Rabe-Hesketh ve ark., 2002). Söz konusu yöntemler ile çalışmada değerlendirdiğimiz son modelde üçüncü seviyeye ait varyans-kovaryans için tahminleme yapılamamıştır. Bu sonuçlar normal dağılışlı olmayan cevap değişkenlerine sahip veri seti analizi için Browne ve Draper (2001)'in yaptığı çalışma ile benzerlik göstermektedir. Zira, adı geçen araştırmacıların çalışmasında da yaklaşık olabilirlik tabanlı yöntemler ile ikinci seviye varyansına ilişkin tahminleme yapılamamıştır. MQL_1 ve PQL_2 analizleri ile elde edilen tahminlemelerde ortaya çıkan bu sorunların nedeni, MQL_1 ve PQL_2 yöntemleri yaklaşık olabilirlik tabanlı yöntemler olmasıdır. Yaklaşık olabilirlik tabanlı yöntemler, olabilirlik fonksiyonu yerine, olabilirlik fonksiyonunun analitik yaklaşımını maksimize eder. Olabilirlik fonksiyonu, modelde çok miktarda şansa bağlı etki bulunması durumda yüksek boyutlarda integraller içermekte ve bu durum çözümü güçleştirmektedir. Ancak yaklaşık olabilirlik tabanlı yöntemler, hesaplanmasının kolay olması, tahminlerin kısa sürede elde edilebilmesi ve çok fazla bilgisayar hafızasına ihtiyaç duymaması nedeni ile tercih edilen yöntemlerdir. Bu çalışmada da MQL_1 ve PQL_2 yöntemleri ile ilgili parametre tahminleri çok kısa sürede elde edilmiştir. Gözlenen bu durum Breslow ve Clayton, 1993; Rodriguez ve Goldman (1995), Goldstein ve Rasbash (1996), Yang (1997) ve Browne (1998) tarafından yapılan çalışmalarda da elde edilmiştir.

Normal dağılış göstermeyen verilerin analizinde söz konusu sorunların aşılmasında MCMC yöntemleri tercih edilmektedir. Bu çalışmada da alternatif yöntem olarak MCMC kullanılmıştır. Yaklaşık olabilirlik tabanlı yöntemlerin yerine MCMC yöntemlerinin kullanılmasında iki önemli neden vardır. Bunlardan biri, bu yöntemlerin doğru (accurate) tahminler vermesi ve diğeri ise modellerin uyumunun Bayes çatısı altında yapılmasıdır. Üzerinde çalışılan son modelde kullanılan MCMC yöntemlerinin tamamı, sabit etkiler ve hatalar için MH örnekleyicisini ve varyanslar için Gibbs örnekleyicisini kullanmaktadır. Bu yöntem Hibrit yöntem I olarak bilinir. Zira çok seviyeli modellerde çalışan araştırmacılar (Browne, 1998; Hox, 1998; Browne, 2002; Rasbash ve ark., 2002; Goldstein ve ark., 2002b; Browne ve Rasbash, 2003) kesikli veriler için hibrit Metropolis-Gibbs yöntemini önermektedirler. Çünkü Gibbs örneklemenin kullanılabilmesi için bir grup parametreye ait şartlı posterior dağılışların Normal dağılış gibi standart forma sahip olması gerekir. Nitekim üzerinde çalışılan veri seti için Gibbs örnekleyicisi kullanılarak MCMC analizi yapılmaya çalışılmıştır. Elde edilen sonuçlar, bu öneride bulunan literatürü desteklemektedir.

Çalışmada kullanılan bir diğer MCMC yöntemi, adaptive hibrit Metropolis-Gibbs yöntemidir (MCMC[1]). Bu yöntem ile elde edilen analiz sonuçları değerlendirildiğinde, modellerin her birinde yer alan sabit ve şansa bağlı etkilere ilişkin tahminlerin elde edildiği görülmektedir. Alternatif bir yaklaşım olan adaptive yöntem, başlangıçta önerilen dağılışlara sahip olmayı ve daha sonra bu dağılışları adapte etmeyi gerektirir. Çok seviyeli modellemede bu yöntem, hibrit Metropolis-Gibbs yöntemi ile birlikte adaptive hibrit Metropolis-Gibbs yöntemi olarak bilinir. Bu algoritma Markov zincirinin karışımını iyileştirmek için kullanılır. Zira önerilen iyi bir dağılış ile işe başlamak Markov zincirinin karışımını iyileştirdiği için daha düşük otokorelasyona sahip bir zincir elde edilir. Bunun bir sonucu olarak, önerilen dağılışa yakınsama için daha az iterasyona ihtiyaç duyulur. Zira çalışmada kullanılan son model için prior varyans seçimi için skala faktör kullanıldığında elde edilen devians değeri yaklaşık 3219 birim iken adaptive yöntemin kullanılması ile bu değer yaklaşık olarak 3099 birim olmuştur. Dolayısıyla deviansta 119 birimlik bir azalma sağlanmıştır. Bu, prior varyans seçimi için adaptive yöntemin kullanılması gerekliliğini gösteren bir diğer ölçüt olmaktadır. Diğer bir ifade ile bu, adaptive yöntemi kullanan hibrit Metropolis-Gibbs yönteminin skala faktörü kullanan hibrit Metropolis-Gibbs yönteme tercih edilme nedeni olmuştur. Çünkü bu yöntem;

$$r^* \ge r, \sigma_p \to \sigma_p\left[2-\left(\frac{1-r^*}{1-r}\right)\right],$$ aksi durumda

$$\sigma_p \to \frac{\sigma_p}{\left(2-\frac{r^*}{r}\right)}$$

uygulayarak önerilen dağılışın standart sapması azaltılır veya artırılır. Burada r kabul oranını gösterir. Bir önceki 100 iterasyondaki kabul oranı r^*, ve verilen parametre için önerilen dağılışın standart sapması σ_p ile gösterilir Her bir parametreler için tolerans aralığı ölçütü sağlanınca, bu işlem biter. Tolerans aralığı (r-δ, r+ δ) olup (r, δ) değerleri için (0.50, 0.10) kullanılmıştır (bakınız, 3.2.5.4.3. başlığı). Bu konuda çalışan araştırmacılar Browne ve Draper (2000) ve Browne ve Draper (2001)'in kullandığı gibi tek değişkenli önerilenler için kabul oranı olarak %50 kullanılmıştır. Ancak bu oranın çok değişkenli önerilenler dağılışlar için %40 olması durumunda sonuçların daha doğru olacağı bildirilmiştir (Browne, 2003). Dolayısıyla daha sonra yapılan analizler için adaptive yöntemin kullanılmasına karar verilmiştir. Bu yöntem, benzer nedenlerden dolayı Browne (1998), Browne ve Draper (2000), Browne (2001) ve Browne (2003) tarafından da tercih edilmiştir.

MCMC yöntemlerini tamamımda iki problemden bahsedilebilir. Bunlardan ilki, burn-in periyodunda doğru posterior dağılışı elde edecek kadar yeterli iterasyona sahip olunduğundan emin olmaktır. Diğeri, iyi tahminler elde edecek bir posterior dağılıştan yeteri kadar birbirlerinden bağımsız örnek toplamaktır (Hox, 1998). MCMC yöntemlerinde parametre tahminleri için üretilen değerlerin bağımsız olması önemlidir. Zira, MH'de Gibbs örnekleyicisine göre söz konusu değerler arasında daha fazla bağımlılık söz konusudur. Benzer sonuçlar bu çalışmada kullanılan Hibrit yöntemlerde (MCMC[1] ve MCMC[2]) dikkat çekmektedir. Bu çalışmada kullanılan MCMC[1] yöntemi ile elde edilen parametre tahminlerine ait zincir durumunu ifade eden bilgiler, bağımsız örneklerin elde edilebilmesi için daha fazla iterasyona ihtiyaç olduğunu göstermiştir. Bunun üzerine 5000 iterasyonluk bur-in periyodu ve bunu 50000 iterasyonluk görüntülenen zincir uzunluğunu kullanan MCMC[3] ile analizler yapılmıştır. Bu analiz sonunda elde edilen devians değeri dikkate alındığında tercih edilen yöntem MCMC[3] olmuştur. Dolayısıyla bu yöntem ile elde edilen parametrelere ilişkin katsayılar ve standart hatalar dikkate alınmalıdır.

Bu çalışma, çok seviyeli GLM'de kullanılan tahminleme yöntemlerinin birbirlerine üstünlüklerini ortaya koymaktadır. Poisson dağılışı gibi kesikli dağılış gösteren verilerin analizinde kullanılan yaklaşık olabilirlik tabanlı yaklaşımlardan MQL_1 ve PQL_2 yöntemlerinin sapmalı tahminler verdiği ortaya koyulmuştur. Çalışmada kullanılan veri seti için bu sapmalar oldukça kötü olmuştur. Daha öncede belirtildiği gibi bu sapmalar, bazı veri setleri için oldukça kötü olabilir. Çalışmanın bir sonraki aşamasında olabilirlik fonksiyonunun kendisini maksimize eden MCMC yöntemlerine ilişkin sonuçlar elde edilmiştir. Elde edilen bilgiler, çok seviyeli bir yapıya sahip Poisson dağılışı gösteren verilerin Gibbs örnekleme kullanılarak analizinin yapılamayacağı göstermektedir. Ancak Hibrit Metropolis-Gibbs yöntemi normal dağılış göstermeyen verilerin analizi için uygun görülmüştür. Adaptive yöntem ile skala faktörü kullanan hibrit Metropolis-Gibbs yöntemleri karşılaştırılmış, adaptive hibrit Metropolis-Gibbs yöntemini üstünlükleri nedeniyle tercih edilen yöntem olmuştur. Bu yöntemin tercih edilme nedenleri yukarıda verilmiştir. Kullanılan MCMC yöntemlerinde herhangi iki parametre arasında bir ilişki olmadığı varsayılarak tek değişkenli önerilen dağılış kullanılmıştır. Zira söz konusu veri seti için tahminler çok değişkenli önerilen dağılışın kullanılması ile elde edilmiştir. Sonuç olarak, cevap değişkenlerinin Poisson dağılışına sahip olduğu çok seviyeli bir model için şunlar söylenebilir; GLM ve GLM'e ilişkin tahminleme yöntemleri kullanılmalıdır, MCMC yöntemleri (bu çalışmada adaptive

Metropolis-Gibbs yöntemi I), yaklaşık olabilirlik (MQL_1 PQL_2) yöntemlerine tercih edilebilir.

Kaynaklar

Agresti, A., Booth, J. G., Hobert, J. P. and Caffo, B., 2000. Random Effect Modelling of Categorical Response Data, Sociological Methodology. 30 (1) pp 27-80.

Blatchford, P., Goldstein, H., Martin, C., Browne, W., 2002. A Study of Class Size Effects in English Scholl Reception Year Classes. British Educational Research Journal. 28 (2):169-185.

Breslow, N. E. and Clayton, D. G., 1993. Approximate Inference Generalized Linear Mixed Models. JASA 88 (421) pp.9-25.

Browne, W.J., 1998. Applying MCMC Methods to Multilevel Models. PhD Dissertation.

Browne, W.J., 2001. Using the WinBUGS interface in MLwiN. http://seis.bris.ac.uk/~frwjb/materials/wbbugsps.pdf

Browne, W.J. 2005. MCMC Estimation in MLwiN Version (2.0)(Second edition). University of Bristol,UK. 291.

Browne, W.J., Draper, D., 2000. Implementation and Performance Issues in The Bayesian Fitting of Multilevel Models. Computational Statistics. 15: 391-420.

Browne, W.J., Draper, D., 2001. A Comparison of Bayesian and Likelihood-Based Methods For Fitting Multilevel Models. Computational Statistics. Institute of Education, University of London. London. UK.

Browne, W.J., Draper, D., Goldstein, H., Rasbash, J., 2002. Bayesian and Likelihood Methods for Fitting Multilevel Models with Complex Level-1 Variation. Computational Statistics&Data Analysis, 39:203-225.

Browne, W.J., Rasbash, J., 2003. Multilevel Modelling. http://www.cmm.bristol.ac.uk/team/mmsage.pdf

Duncan, T.E., Duncan, S.C., Okut, H., Lisa, A.,Strycker, M.A., Hollie Hix-Small, B.S., 2002. Multilevel Model of Neighborhood Collective Efficacy. Oregon Reserch Institute, 1715 Franklin Blvd., Eugene, OR 97403.

Foulley, J.L., Im, S., 1993. A Marginal Quasi-Likelihood Approach to The Analysis of Poisson Variables with Generalized Linear Mixed Models. Genet. Sel. Evol., 25:101-107.

Goldstein, H., 1991. Nonlinear Multilevel Models, with an Application to Discrete Response Data. Biometrika, 78(1): 45-51.

Goldstein, H., 1995. Multilevel Statistical Models, Second Edition. http://www.sosiologie.uni-halle.de/langer/multilevel/books/goldstein.pdf

Goldstein, H., Browne, W., Rasbash, J., 2002a. Partioning Variation in Multilevel Models. Understanding Statistics, 1(4):223-231.

Goldstein, H., Browne, W., Rasbash, J., 2002b. Multilevel Modelling of Medical Data. Statistics in Medicine, 21:3291–3315.

Goldstein, H., Rasbash, J., 1996. Improved Approximations For Multilevel Models With Binary Responses. Journal of The Royal Statistical Society. Series A. 159: 505-513.

Hox, J., 1995. Applied Multilevel Analysis. http://www.google.com.tr/search?hl=tr&q=Applied+Multilevel+Analysis+hox&btnG=Ara&aq=f&aqi=&aql=&oq=&gs_rfai=

Hox, J., 1998. Multilevel Modelling in Windows; A Reviev of MLwiN. Multilevel Modelling Newsletter, 10(2):2-5.

Langford, I.H., Bentham, G., Mcdonald, A.L., 1998. Mortality from Non-Hodgkin Lymphoma and UV Exposure in THe European Community. Health & Place, 4(4): 335-364.

Langford, I.H., Leyland, A.H., Rasbash, J., Goldstein, H., 1999. Multilevel Modelling of The Geographical Distributions of Diseases. Appl. Statist.,48(2): 253-268.

Leyland, A.,H., McLeod, A., 2000. An Introduction to Multilevel Modelling Using MLwiN. MRC & Public Healt Sciences Unit Occasional Paper. 2:1-35

Lin, X., Breslow, N.E., 1996. Bias Correction in Generalized Linear Mixed Models with Multiple Components of Dispersion. Journal of The American Statistical Association, 91: 1007-1016.

McCullagh, P., Nelder, J.A., 1983. Generalised Linear Models. London. Chapman and Hall.

Naderi, A., Mace, J., 2003. Education and Earnings: A Multilevel Analysis. A Case Study of The Manufacturing Sector in Iran. Economics of Education Review, 22: 143-156.

Rabe-Hesketh, S., Skrondal, A., Pickles, A., 2002. Reliable estimation of Generalized Linear Mixed Models Using Adaptive Qaudrature. Stata Journal, 2(1):1-21.

Rasbash, J., Browne, W.J., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I., Lewis, T., 2000. A User's Guide to MLwiN. Version 2.1. Institute of Education, University of London.

Raudenbush, S.W., Yang, M.L., Yosef, M., 2000. Maximum Likelihood for Generalized Linear Models with Nested Random Effects Via High-Order, Multivariate Laplace Approximation. Journal of Computational and Graphical Statistics, 9(1): 141-157.

Rodríguez, G., Goldman, N., 1995. An Assessment of Estimation Procedures For Multilevel Models With Binary Responses. Journal of The Royal Statistical Society. Series A, 158: 73-89.

Yang, M., 1997. Multilevel Models for Multiple Category Responses-A Simulation. Multilevel Modelling Newsletter, 9(1):1-17.

Zhou, X.H., Perkins, A., Hui, S.L., 1999. Comparisons of Software Packages for Generalized Linear Multilevel Models. The American Statistician, 53(3):282-290.

Ek 1;

Hibrit Metropolis-Gibbs yöntemi I (Univariate Update):

Bu, Hibrit Metropolis-Gibbs olarak bilinir. Hibrit Metropolis-Gibbs yöntemlerine ilişkin algoritmayı vermeden önce, 3 seviyeli Poisson regresyon modeline ait bir eşitlik yazalım.

$$y_{ijk} \sim Poisson(\lambda_{ijk})$$
$$\log(\lambda_{ijk}) = \beta_0 + \beta_1 X_{ijk} + u_{0ik} + v_{0k} + u_{1ik} + v_{1k} \quad 55)$$

Buna göre algoritmanın daha anlaşılır olabilmesi için modeli aşağıdaki gibi yazılır (Browne, 1998; Browne ve Draper; 2000).

$$\log(\lambda_{ijk}) = X_{1ijk}\beta_1 + X_{2ijk}\beta_{2jk} + X_{3ijk}\beta_{3k} \quad (56)$$

$\beta_{2jk} \sim MVN(0,\Omega_2)$ ve $\beta_{3k} \sim MVN(0,\Omega_3)$

Burada u_{ik} ve v_k sırasıyla β_{2jk} ve β_{3k} olmaktadır. Çok seviyeli Poisson regresyonunda seviye 1'e ait hata terimi yoktur. Çünkü hem ortalama hem de varyans yalnızca λ_{ijk} parametresinin fonksiyonudur $[E(y_{ijk}) = \lambda_{ijk}, \mathrm{var}(y_{ijk}) = \lambda_{ijk}]$. Bu nedenle ortalamanın tahminlenmesi yeterli olmaktadır (Browne, 1998). Dolayısıyla N seviyeli genel poisson regresyon modeli için MCMC'de algoritma yalnızca 3 aşamadan ibarettir. Yukarıdaki eşitliği N seviyeli bir modele genelleştirilmek için aşağıdaki tanımlamalar yapılsın.

M_T modeldeki tüm gözlemlerin seti olsun. M_{lj} seviye *l*'deki (Burada *l=1, 2, 3*) *j* kategorisinde olan gözlem seti olsun. X_{li} *l* seviyesindeki *i*'ninci gözlem için değişken vektörü olsun. Burada $l=1$ sabit etkileri göstersin. Seviye *l*'deki şansa bağlı parametreler β_{lj} ile gösterilsin (*l>1*) burada *l* yüksek seviye terimlerinin kombinasyonu olsun ve sabit etkiler β_1 olsun. Son olarak Ω_l seviye *l* varyans matrisi olsun. $(X\beta)_i$ kısıtlaması ile (Browne, 1998; Browne ve ark., 2001)

$$(X\beta)_i = X_{1ijk}\beta_1 + X_{2ijk}\beta_{2jk} + X_{1ijk}\beta_{3k} \quad (57)$$

kullanılırsa model daha kısa bir notasyonla aşağıdaki gibi ifade edilir.

$$y_i = Poisson(\lambda_i), \log(\lambda_i) = (X\beta_i),$$

$\beta_{li} \sim MVN(0,\Omega_l)$ olur. Çok seviyeli bir Poisson regresyonu için Hibrit Metropolis-Gibbs algoritması aşağıdaki gibi yazılır.

1. Aşama: Sabit etkiler (β_1) için;

$\beta_{1i}^{(t-1)}$ önerilen dağılıştan β_{1i}^* örneklenir. $\min(1, p(\beta_{1i}^*|y,...)/p(\beta_{1i}^{(t-1)}|y,...))$ olasılığı ile değerlendirilir. Eğer olasılık değeri birden küçük ise önerilen dağılış ve dolayısı ile önerilen değer kabul edilir. Aksi halde bir önce elde edilen değer ile algoritma devam eder. Bu durumda başka bir önerilen dağılış kullanılır. Buna göre algoritma aşağıdaki gibi özetlenir.

$i = 1,...,N_{sabit}$ için

$$\beta_{1i}^{(t)} = \beta_{1i}^* \ \min(1, p(\beta_{1i}^*|y,...)/p(\beta_{1i}^{(t-1)}|y,...))$$
$$= \beta_{1i}^{(t-1)} \text{ diğer durumda.}$$

Burada $\beta_{1i}^* = \beta_{1i}^{(t-1)} + \gamma_{1i}$, $\gamma_{1i} \sim N(0,\sigma_{1i}^2)$ olur ve sabit etkiler için şartlı dağılış;

$$p(\beta_{1i}|y,...) \propto p(\beta_1)\prod_{i\in MT} e^{-e^{(X\beta)_i}}\left(e^{(X\beta)_i}\right)^{y_i}$$

şeklinde verilir.

2. Aşama: *l* seviyesi hataları (β_l) için;

$\beta_{lji}^{(t-1)}$'den β_{lji}^* örneklenir. $\min(1, p(\beta_{lji}^*|y,...)/p(\beta_{lji}^{(t-1)}|y,...))$ olasılığı ile değerlendirilir. Eğer olasılık değeri birden küçük ise önerilen dağılış ve dolayısı ile önerilen değer kabul edilir. Aksi halde bir önce elde edilen değer ile algoritma devem eder. Bu durumda başka bir önerilen dağılış kullanılır. Buna göre algoritma aşağıdaki gibi özetlenir.

$l = 2,...,N, j = 1,...,n_l$ ve $i = 1,...,n_{rl}$

$$\beta_{lji}^{(t)} = \beta_{lji}^*, \ \min(1, p(\beta_{lji}^*|y,...)/p(\beta_{lji}^{(t-1)}|y,...))$$

olasılık ile

$$= \beta_{lji}^{(t-1)}, \text{ diğer durumda.}$$

Burada $\beta_{lji}^* = \beta_{lji}^{(t-1)} + \gamma_{lji}$, $\gamma_{lji} \sim N(0,\sigma_{lji}^2)$ ve hatalar için şartlı dağılış,

$$p(\beta_{lj}|y,...) \propto \prod_{i\in M_l} e^{-e^{(X\beta)_i}}\left(e^{(X\beta)_i}\right)^{y_i}|\Omega_l|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\beta_{lj}'\Omega_l^{-1}\beta_{lj}\right]$$

olur.

3. Aşama: Seviye *l* varyansı (Ω_l) için;

Varyanslar için önerilen dağılış ters (invers) Wishart dağılışıdır. Bu önerilen dağılışın her seferinde kabul edilmesi nedeni ile varyanslar için Gibbs örnekleme kullanılır. Çünkü Gibbs örnekleme her önerilen dağılışın kabul edildiği MH örneklemenin özel bir durumudur (Browne, 1998). Buna göre varyanslar için,

$$p(\Omega_l^{-1}|y,...) \propto p(\beta_l|\Omega_l)p(\Omega_l^{-1})$$
$$\Omega_l^{-1} \sim$$
$$Wishart_{n_{rl}}\left[S_{pos} = \left(\sum_{i=1}^{n_l}\beta_{li}\beta_{li}' + S_{pl}\right)^{-1}, v_{pos} = n_l + v_{pl}\right]$$

verilir. Burada *nl* seviye *l* birimlerinin sayısı ve n_{rl} seviye *l*'deki şansa bağlı değişkenlerin sayısıdır.