

M-Plus'ta Latent Class Analizlerin Uygulanmasına İlişkin Bir Örnek Çalışma

Murat KAYRI¹

¹Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi, Bilgisayar Bölümü, Van

Özet : Popülasyona uygulanacak olan test istatistiğinin ve bu test istatistiği doğrultusunda kurulacak olan modele ait parametre tahminlemesinin doğru ve sağlam bir zeminde ele alınması dikkat edilmesi gereken bir unsurdur. Popülasyonu meydana getiren birey ya da değişkenlerin benzer özellikler taşımadığı durumlarda; tek gibi gözükən popülasyonun alt popülasyonlara bölünmesi istatistik süreç açısından daha sağlıklı sonuçlar üretebilmektedir. Zira, alt popülasyonlar için elde edilen tanımlayıcı istatistik sonuçları ve parametre tahminlemeleri daha sağlam (robust) olabilmektedir. Bu çalışmada, heterojen bir veri setinin, homojen olan alt-popülasyonlara (latent class) bölünmesini hedef alan karışımli modelin M-Plus yazılımında uygulanabilirliği ele alınmıştır. Ayrıca, M-Plus'ta karışımli modelin nasıl kodlandığı ve bu kodlamaların ne anlam içerdiği de tartışılmıştır. Araştırmacıları karışımli modelleri uygulamaya teşvik amacıyla, modelin literatür bildirişleri özenle seçilmiş; modelin matematiksel ve mantıksal boyutları da etraflıca irdelenmiştir. Özellikle heterojen olan popülasyonu, ideal homojen alt sınıflara (sub-population) ayırmada uyum kriterleri (Akaike Bilgi Kriteri, Bayesian Bilgi Kriteri ve Log-olabilirlik) örneklem üzerinde vurgulanmıştır. Örneklem, heterojen olan sürekli ve kesikli değişkenlerden oluşan 643 birey tarafından oluşturulmuştur. Bu çalışma sonucunda; analiz yapılacak olan veri setinin büyüklüğü durumunda; veri setinin heterojenliği riskini yok etmek üzere ve bu anlamda elde edilen homojen alt-sınıfların regresyon denklemlerinin ve parametre tahminlemelerinin Latent Class (mixture models) tekniği tarafından M-Plus ortamında sağlıklı bir şekilde yapılabileceği kanısına varılmıştır.

Anahtar sözcükler : Karışımli model, latent class, M-Plus, uyum kriteri

The Application of Latent Class By Using M-Plus Programme : A Case Study

Abstract : It is considered to be important to handle the test statistics that is going to be applied to population and the parameter estimation statistics on an accurate and robust manner. If individuals or variables do not carry some similar features then searcher must use some clustering method. By this way a heterogeneous population is divided into enough homogeneous sub-population. And the result of descriptive statistics and parameter estimation will be more robust and coherence. In these cases, using latent class statistics contribute to obtain robust parameter estimation. Also it provides the homogeneity assumption. Especially, the mixture model was studied on this article which belongs to the latent class. The purpose of mixture model is to divide heterogeneous population into homogeneous sub-populations. At the same time, in this article it was explained how to construct a mixture model by using M-Plus programme. To encourage the use of the mixture model in the educational sciences, the references of the mixture model were selected fastidiously and the mathematical, and logical dimensions of this model were explained comprehensively. In this study, a mixture model which means heterogeneous verdict is divided into homogeneous subpopulations (latent class) is used in an experimental study which is consisted of 643 individuals who are registered with an educational department. In this study, Akaike's Information criterion, Bayesian Information criterion and Log-likelihood terms were identified well. At the end of the study, the article concludes that the mixture model seems to be more efficient in analysing the similar data and verdict analysed in this study for experimental studies.

Key Words: Mixture model, latent class, M-Plus, information criterion

Giriş

Latent class analizler benzer özellikleri taşıyan bireylerin kategorize edilip, kategorize edilen bu bireylerin ayrı sınıflar şeklinde tanımlanması şeklinde düşünülebilir (Nylund ve ark. 2006; Clogg 1995; Lazarfelds ve Hanry 1968). Yine benzer şekilde Latent class modellerin (mixture models) genel teması, dağılımın ilk etapta sayısı bilinmeyen sınıflara ya da alt-sınıflara (latent class) bölünmesi şeklindedir (Everitt ve Hand 1981; Titterington, Smith ve Markov 1985). Bilindiği üzere istatistiksel algoritmalara ait gelişmeler daha çok model kurmaya yönelik olup, söz konusu bu modellerin istatistik paket programlarında analiz edilebilirliği yine gelişmekte olan bir sektör konumuna gelmiştir. Latent class analizlerin, örneğin karışımli model (mixture modeling) gibi analizlerin uygulanabildiği önemli yazılımlardan biri M-Plus paket programıdır. Genel itibarıyla M-Plus yazılımının tercih edilme sebebi; kullanımı kolay bir ara yüz desteği ile istatistikçinin matris işlemlerine (matrix algebra) ve istatistik formüllerine ihtiyaç duymaması ve araştırmacıya çok güçlü bir model yapısı sunabilmesidir (Muthen ve ark. 2006). Ayrıca M-Plus, farklı analiz tekniklerini de bünyesinde barındırabilmektedir: Yapısal eşitlik modelleri

(structural equation model), latent class (finite mixture; sonlu karışımli modeller), çok seviyeli modeller (multilevel modeling) ve yaşam analizleri (survival analysis) M-Plus'ın bünyesinde barındırdığı ileri düzey istatistik analizlerdir (Muthen ve ark. 2006). 1998 yılında Muthen tarafından geliştirilen M-Plus yazılımı her yıl yapısına yeni modüller eklemektedir. Esnek bir yapıya sahip olan M-Plus gelişen istatistik algoritmalarına karşı kendisini en hızlı şekliyle yenileyebilmektedir. Özellikle sürekli ve kategorik latent değişkenleri aynı anda modele dahil edebildiğinden araştırmacıya genel bir model yapısı sunabilmektedir (Muthen 2001; Muthen ve Muthen 1998; Muthen ve Muthen 2002).

İlk etapta genetik çalışmalar için geliştirilen latent class analizler, daha sonradan sosyal alanlar içinde kullanılmaya başlanmıştır. Özellikle heterojen olan büyük veri setleri için homojen olan alt popülasyonların elde edilmesinde latent class (mixture model) analizlerin etkili olduğu düşünülmektedir. Yine aynı şekilde elde edilen her alt sınıf için bir takım algoritmalar (Expectation Maximization gibi) vasıtası ile sağlam ve sapmasız

(unbiased) parametre tahminlemesi yapılabilmektedir (Duncan ve ark. 2002; Yeşilova 2003). Bu yönüyle latent class analizler; kitenin değişim sürecinin anlaşılması kadar, değişimdeki bireysel farklılıkların da tanımlanmasını içermektedir (Baltes ve Nesselroade 1979; Collins ve Horn 1991). Çünkü, popülasyon alt sınıflara bölünüp, uç bireyler de ayrı bir alt sınıfta değerlendirilmektedir. Geleneksel metotlarda bu durumu gözlemlemek mümkün olamamaktadır. Kategorik (kesikli) ve sürekli verileri bir arada modelde yürütebilen latent class modeller (karışımli model); aynı zamanda lojistik regresyon modelini de, indikatör (u_i) desteği ile yürütebilmektedir. Bağımlı değişkenin 1 ya da 0 gibi (binary trait) Bernoulli model içermesi durumunda da karışımli model U_i şeklinde bir indikatör tanımlamasıyla sağlam (robust) parametre tahminlemeleri yapabilmektedir (Zhang ve Merikangas 2000). Bilindiği üzere, lojistik regresyon modeli, "bir veya daha fazla kesikli veya sürekli bağımsız değişken ile kesikli bir bağımlı değişken arasındaki ilişkiyi incelemektedir" (Topuz ve Çakır 2002). Karışımli model ise, bu analizlerin özelliklerini kapsamaktadır ve "bu modelin üstünlüğü hususunda: şansa bağlı (random) fenomenler ışığında geniş varyasyonlu, Gaussian iterasyon tabanlı ve matematik merkezli yaklaşımı ile vazgeçilmez bir yöntemdir" yaklaşımı hâkimdir (Peel ve McLachan 2000). Lindsay (1995), özellikle biyoloji, psikoloji ve genetik çalışmalarda karışımli modelin yaygın olarak parametre tahminlemelerinde kullanıldığı vurgulanmaktadır.

Duncan ve arkadaşları (2002), genel itibarıyla latent class analizlerin aşamalarını; popülasyona ait alt sınıfların sayısının belirlenmesi (number of class), gözlenen değişkenlerin ilgili sınıflara düşme olasılıkları, gözlenen değişken ya da bireylerin hangi alt sınıflara yerleştirildiği ve son olarak model uyumunun yapılması (fit of model) şeklinde belirlemişlerdir.

Literatürde bilinen maksimum olabilirlik tahminleyicinin bünyesinde yürütmüş olduğu expectation maximization (EM) algoritması ile modelde düşünülen parametreler tahminlenebilmektedir. M-Plus karışımli modeller için kullanılan bu tür algoritmaları basit bir komut satırı ile yürütebilmektedir. Benzer şekilde, latent class analizlerinde kullanılan diğer algoritmalar da (parametre tahminlemesi için) M-Plus'ta işlenebilmektedir.

Bu çalışmada, latent class analizlerin (örneğin, karışımli modeller) M-Plus'ta uygulanabilirliğini göstermek üzere, genel itibarıyla M-Plus'ın kod yazılımı (syntax) ve bu kodların ne anlam ifade ettiği açıklanmaya çalışılmıştır. Ayrıca, popülasyona ait bilinmeyen alt sınıf sayısının tespiti için kullanılan bir takım kriterlerin (Akaike's Information Criterion ve Bayesian Information Criterion) hem matematiksel modeli, hem de analiz sonucu elde edilen bu değerlerin nasıl yorumlanabileceği ele alınmıştır. Bununla birlikte EM algoritması da matematiksel boyutuyla açıklanmaya çalışılmıştır.

Materyal ve Yöntem

Çeşitli değişkenlerden oluşan ve homojen olmayan bir veri seti (643 birey) materyal olarak kullanılmıştır. Bu veri setinin EM algoritması ile sonlu karışımli modelde nasıl homojen sınıflandığı da yöntem olarak belirlenmiştir. Deneme deseninde kullanılan değişkenler kesikli ve sürekli formda olup; bireylere ait olan Öss puanı bağımlı değişken, lise başarı puanı, anne ve baba mesleği,

yaşadıkları bölge, mezun oldukları lise kolu ise bağımsız değişken konumundadır. Ayrıca optimal sınıf sayısının belirlenmesinde kullanılan Akaike ve Bayesçi bilgi kriterleri de M-Plus yazılımı tarafından belirlenmiştir.

Karışımli modelin (mixture models) özellikleri : Latent class analiz çeşidi olan karışımli modellerde bireylerin alt sınıflarda olma olasılıklarını belirlemeye çalışılmaktadır. Veri setinin dağılışı formuna bakılmaksızın multinomial logit dağılışı kullanılarak, c. alt sınıf ile x arasındaki bağlantı kurulur (Wang ve ark. 1998; Chan ve Kou 2001; Yeşilova 2003). N (örnek büyüklüğü) verildiğinde multinomial logit dağılışında c_i 'nin olasılık dağılışı,

$$P\{C_{i1}=c_{i1}, C_{i2}=c_{i2}, \dots, C_{iK}=c_{iK}\} = \binom{N}{c_{i1}, \dots, c_{iK}} \pi_{i1}^{c_{i1}} \dots \pi_{iK}^{c_{iK}} \quad (1)$$

elde edilir (Duncan ve ark. 2002). Multinomial logit için model,

$$\pi_k = \frac{\exp\{\nabla_{ik}\}}{\sum_{k=1}^K \exp\{\nabla_{ik}\}} \quad (2)$$

yazılır (Duncan ve ark. 2002). Burada,

$$\nabla_{ik} = \log \frac{\pi_{ik}}{\pi_{iK}} \quad \text{olmaktadır (Duncan ve ark. 2002).}$$

Parametre tahminleri için kullanılan EM algoritmasında log-olabilirlik tabanlı maksimum olabilirlik kullanılmaktadır. Bütün veriler için log-olabilirlik fonksiyonu,

$$L(Y, X, \beta, \pi) = \sum_{i=1}^n \sum_{k=1}^K C_{ik} \log \pi_k + \sum_{i=1}^n \sum_{k=1}^K C_{ik} \log P(y_i / \lambda_i) \quad (3)$$

gibi ele alınır ve bu olabilirlik parametre tahmininde kullanılır (Yeşilova 2003). 3 numaralı eşitlikte, C_{ik} gözlenemeyen (unobserved) gözlemler olup, $C = \{c_{ik}, i = 1, 2, \dots, n; k = 1, 2, \dots, K\}$ olarak ele alınmalıdır.

Yine aynı şekilde c ile x arasındaki logit bağıntısı ise aşağıdaki gibi elde edilir:

$$P(c_{ik} = 1 | x_i) = \frac{\exp(a_{ck} + \gamma_{ck} x_i)}{\sum_{k=1}^K \exp(a_{ck} + \gamma_{ck} x_i)} \quad (4)$$

olarak elde edilir (Duncan ve ark. 2002). Bu denklemdeki C_{ik} aşağıdaki gibi düşünülebilir:

$$C_{ik} = 1 \quad C_{ik} \in K \\ C_{ik} = 0 \quad \text{diğer hallerde}$$

Karışımli modellerde parametreler tahmindikten sonra

$Y_{(ij)} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n$ gibi bir regresyon modeli tanımlanabilmektedir. M-Plus yazılımının tamamlanıp, programın çalıştırılmasından sonra model için ortalama değer (intercept; β_0) ve diğer değişkenler için parametre katsayıları ($\beta_1, \beta_2, \beta_3$ gibi) tespit edilmektedir. Sadece veri kümesi ortalaması esas alınıp alt popülasyonlar oluşturulabilirdiği gibi, hem ortalama hem de varyans birlikte esas alınıp alt popülasyonlar oluşturulabilir. Sadece ortalamaya göre oluşturulan

uygulama çok değişkenli istatistikte kullanılan kümeleme (cluster) veya profil analizinden çok farklı bir durum göstermemektedir. Bu nedenle verilerin hem ortalaması hem de varyansı dikkate alınarak multinomial logit prensipler uygulanıp, sınıflandırma (alt populasyonlara ayırma) yapmak daha karışık olmasına karşın homojen alt populasyonlar oluşturmada en uygun yol olmaktadır (Yeşilova 2003).

EM (Expectation Maximization): Modele ait log-olabilirlik fonksiyonunun maksimizasyonu EM algoritması ile tespit edilebilmektedir. Log-olabilirlik temeli doğrultusunda AIC ve BIC değerleri de tespit edilebilmektedir. Yine aynı şekilde modele ait parametre tahminleri de EM ile yapılmaktadır. EM algoritması E (Expectation) ve M (Maximization) aşamalarından oluşup; E aşaması heterojen olan veri setinin ideal olan homojen alt-sınıf sayısını belirlemede kullanılmaktadır. M aşaması ise bu tahminlemeyi maksimize etmektedir. Yani, gözlemlerin gözlenemeyen bireylerini eksik gözlem olarak kabul ederek ve model için tüm veri seti oluşturularak işlem yapılmaktadır (Dhanavanthan 2000).

EM algoritması E ve M aşamalarında oluşan bir iteratif işlemi kullanarak parametrelerin en çok olabilirlik tahminlerini elde etmektedir. E aşamasında, gözlenmiş veriler üzerinde koşullu beklenen değerleri ile eksik gözlemler türetilir. M aşamasında ise tüm veriler için beklenen log olabilirliği maksimize eden parametreler elde edilir (Dempster ve ark. 1977; Jansen 1993; Jamshidian ve Jennrich 1997; Yeşilova 2003).

Karışımli modellerde kullanılan EM algoritmasının aşamaları aşağıdaki gibi verilebilir.

Birinci aşama için (Expectation aşaması), $\beta^{(0)}$ ve $\pi^{(0)}$ başlangıç değerleri belirlenir. E aşamasında, $\beta^{(0)}$ ve $\pi^{(0)}$ başlangıç değerleri verildiğinde (X,Y) gözlenmiş verileri ve parametrelerin başlangıç değerleri üzerinden, C eksik gözlemleri elde edilir. $\hat{C}_{ijk}(\beta^{(0)}, \pi^{(0)})$ kullanılarak Ci'nin k'inci unsurunun koşullu olasılığı,

$$P(\text{kategori } k \mid y_i) = \frac{f(y_i | \text{kategori } k) P(\text{kategori } k)}{\sum_{k=1}^K f(y_i | \text{kategori } k) P(\text{kategori } k)} \quad (5)$$

$$\hat{C}_{i,k}(\beta^0, \pi^0) = \frac{\pi_k f_k(y_i | x_i, \beta_k^{(0)})}{\sum_{k=1}^K \pi_k f_k(y_i | x_i, \beta_k^{(0)})}, k=1,2,\dots,K \quad (6)$$

biçiminde verilebilir (Yeşilova 2003). Maksimizasyon (M) aşamasında ise,

$$\left\{ C_i(\beta^{(0)}, \pi^{(0)}) = (C_{i,1}, \dots, C_{i,K})'; i = 1, 2, \dots, n \right\}$$

koşullu olasılıklar verilmişken, parametre tahminleri log olabilirlik fonksiyonunun β ve π 'ya göre maksimize edilmesiyle,

$$Q(\beta, \pi | \beta^{(0)}, \pi^{(0)}) = E\{L(Y, C, \beta, \pi, X) | Y, X, \beta^{(0)}, \pi^{(0)}\} \quad (7)$$

$$Q = Q_1 + Q_2 \quad \text{ve} \quad \text{buradan,}$$

$$Q_1 = \sum_{i=1}^n \sum_{k=1}^K \hat{C}_{i,k}(\beta^{(0)}, \pi^{(0)}) \log(\pi_k) \quad (8)$$

$$Q_2 = \sum_{i=1}^n \sum_{k=1}^K \hat{C}_{i,k}(\beta^{(0)}, \pi^{(0)}) \log(y_i / \lambda_k) \quad (9)$$

biçiminde elde edilir (Yeşilova 2003). Aynı şekilde Yeşilova (2003), $\hat{\beta}$ ve $\hat{\pi}$ tahmin edicileri için denklemin aşağıdaki gibi sürdürülmesini vurgulamaktadır: Q_1 ve Q_2 eşitliklerinin π ve β 'ya göre türevlerinin alınması ile,

$$\frac{\partial Q_1}{\partial \pi_k} = \sum_{i=1}^n \left(\frac{\hat{C}_{i,k}}{\pi_k} - \frac{\hat{C}_{i,K}}{\pi_K} \right) = 0, k=1, \dots, K-1 \quad (10)$$

$$\frac{\partial Q_2}{\partial \beta} = \sum_{i=1}^n \sum_{k=1}^K \hat{C}_{i,k} \frac{\partial}{\partial c} P(y_i | \lambda_k) = 0 \quad (11)$$

şeklinde elde edilir. 6 numaralı eşitlik kullanılarak π_k ,

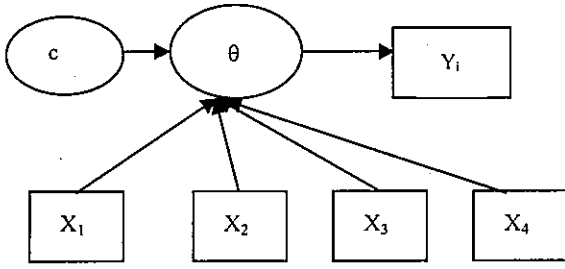
$$\pi_k = \frac{1}{n} \sum_{i=1}^n \hat{C}_{i,k}, k=1, \dots, K-1 \quad (12)$$

biçiminde elde edilmektedir (Wang ve Putterman 1998; Yeşilova 2003). EM algoritmalarının çözümü için yukarıdaki eşitliklerin dışında da bazı formüller geliştirilmiş olup, diğer formüller, yukarıdaki formun kapalılığını basitleştirmek üzere geliştirilmiştir. Diğer formüller de, yukarıda olduğu gibi E ve M aşamalarını ayrı ayrı ele almaktadır.

Karışımli modellerde kullanılan diğer algoritmalarından biri de Quasi olabilirlik (Quasi likelihood). Quasi olabilirlik de, EM'de olduğu gibi tahminleme yapan algoritmalarıdır. Karışımli modellerde, mevcut veri setinin kaç alt sınıfa ayrılacağı bilinmez; eğer K bilinmiyorsa, o zaman EM ya da Quasi olabilirlik (Quasi Likelihood) analizi yapılır. Model farklı alt populasyon sayıları için çalıştırılır ve bir takım uyum iyiliği ölçütlerine göre populasyonun, kaç alt-populasyona bölünebileceğine karar verilir. Bu uyum kriterlerine daha sonra değinilecektir.

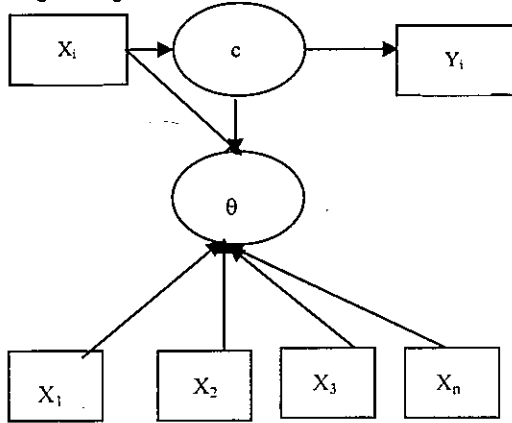
M-Plus'ta yazılımın (syntax) hazırlanması: M-Plus'ta yürütülecek olan veri seti programcı tarafından *.txt formatında hazırlanmalıdır. Bu yazılımda bağımlı, bağımsız değişkenler ve kovaryans ilişkilerinin (Xi) bağımlı değişken (dependent variable) üzerindeki etkiler M-Plus diliyle kodlanır. Karışımli analiz tekniği ile yürütülecek çalışma iki şekilde yürütülebilmektedir. Alt sınıf sayısını belirlemede her bağımsız değişkenin homojen alt sınıflardaki sadece ortalama (θ) üzerindeki etkileri (koşulsuz; unconditional); modeldeki bağımsız kovaryetlerin hem alt sınıf, hem de ortalama üzerindeki etkileri (koşullu; conditional) şeklinde karışımli modeller dizayn edilebilir. Bu iki farklı model türü (koşulsuz ve

koşullu) Nylund ve arkadaşlarının da (2006) benzer tarzda belirttiği gibi ele alınmaktadır (Şekil 1, şekil 2):



Şekil 1. Bağımsız değişkenlerin Karışımı Model üzerindeki etkisi (koşulsuz).

Şekil 1, karışım modellerde bağımsız değişkenlerin ortalama (intercept) üzerindeki etkisi gösterilmektedir. Bu model koşulsuz sonlu karışım olup; X_1, X_2, \dots, X_N bağımsız değişkenleri, θ alt sınıflara ait ortalama değerini, c değişkeni de alt sınıfları ve Y_i değişkeni de modele ait bağımlı değişkeni belirtmektedir. M-Plus'ta modeller hem koşulsuz, hem de koşullu olarak dizayn edilebilmektedir. Deneme desenine ait kurulacak olan modelin koşullu olma durumunda ise bağımsız kovaryetlerin hem sınıf, hem de ortalama üzerine olan etkileri yansıtılır. Bu model şekil 2'deki gibi de gösterilebilir:



Şekil 2. Kovaryet değişkeninin sınıf ve ortalama üzerindeki etkisi (koşullu).

Şekil 2, koşullu karışım modelleri göstermektedir. M-Plus programında yazılım hazırlanırken, modelin koşullu ya da koşulsuz olduğunu belirlemek gerekmektedir. Gözlenen bireylerin uzun bir zaman diliminde (longitudinal data) gözlenmesi söz konusu ise, bu durumda modele eğim parametresi (slope) de eklenebilmektedir. Bu tür modellere de "büyüyen karışım modeller (growth mixture model)" adı verilmektedir. Muthen (2001), büyüyen karışım modeli, yapısal eşitlik modelinin (SEM) ikinci bir üretimi olarak ele alıp, bu modeli SEM'e benzetmektedir. Zaman parametresinin (slope) modele dahil edilmesi durumunda, büyüyen karışım modelin kullanımı sağlıklı parametreler üretecektir (Duncan ve Duncan 1996).

M-Plus'ta karışım model için kullanılacak kodların işlevine değinmek üzere örnek bir veri seti oluşturulmuştur. Bu veri setinde kullanılan bir takım sosyo-demografik değişkenler (nitel) ve yine bu veri setine ait bir takım nicel değişkenler seçilmiştir. Modele dahil edilen

değişkenler hem sürekli hem de kesikli verileri içermektedir. Bu yönüyle latent class analiz hem sürekli hem de kesikli değişkenleri bir anda işleyebilmektedir. M-Plus'ta modelin koşulsuz ya da koşullu olma durumuna göre kodlama yapısı değişmektedir. Bu homojen olmayan veri setine ait alt-popülasyon oluşturma kodlamaları ele alınacak ve aynı zamanda M-Plus'ın çıktısından (output) parametre tahminleme yorumlarına değinilecektir.

M-Plus editöründe örneklem üzerine uygulanan modele ait kodlar (syntax) aşağıdaki gibidir:

```
Data: File is tum.txt; [Popülasyonun homojen olması durumu]
```

```
Variable: Names are birey kimlik OSS
          OOBP alan milli derece
          cins mekiksay tmek kan
          boy kilo babames anames bolge liscik;
Missing = All (-9);
Usevariables OSS OOBP cins babames anames
bolge liscik;
```

```
Classes = c(1);
Analysis: Type = mixture;
Estimator=MLR;
          miteration = 55000;
```

```
Model: %overall%
       ! [OSS];
       OSS;
```

```
OSS on OOBP cins babames anames bolge liscik;
c1# on OOBP cins babames anames bolge liscik;
```

```
%c#1%
[OSS*195];
! [pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge
liscik;
! savedata: file='tum-mur2.dat';
! save=cprob;
```

```
Output: tech1 tech7 ;
```

Yukarıdaki program bir sınıf (c1#) için oluşturulan kodlardan oluşmaktadır. Popülasyonun tek bir homojen alt popülasyondan geldiği ihtimali de söz konusu olduğu için, ilk etapta program bir sınıf için ele alınır. Daha sonra birkaç sınıf için ayrı yazılımlar hazırlanır ve bir takım uyum kriterlerine göre popülasyon ideal sınıf sayısına bölünür. Veri setinde gözlenemeyen değerler (missing values), -9 olarak kodlanmıştır. Yani veri setinde bireye ait bir takım gözlenemeyen değişkenler, -9 olarak kodlanmıştır. Bu herhangi bir karakter olabilirdi, hangi karakter olacağı programcıya bağlıdır.

```
Analysis: Type = mixture;
Estimator=MLR;
          miteration = 55000; komutlarıyla da M-Plus'ta yürütülecek analiz türünün "mixture model" olduğu, tahminleyicinin maksimum olabilirlik (MLR), iterasyon sayısının da (converge) 55000 olarak belirlendiği ifade edilmektedir. Muthen ve Muthen (1998), EM algoritmasının kullanımı için, Maksimum Olabilirlik (Maximum Likelihood) tahmin prensiplerinin uygunluğunu vurgulamaktadır.
```

```
c1# on OOBP cins babames anames bolge liscik;
komut satırı da kovaryet (covariate) dahilindeki (conditional) bir süreci anlatmaktadır. Yani bağımsız değişkenler birlikte değişimin etkisiyle, bağımlı değişkeni (Öss puanı) etkilemektedir. Kovaryetin söz konusu olmadığı bir durumda komut satırı, " ! c1# on OOBP cins
```

babames anames bolge liscik;" şeklinde değişecekti. Yani "!" işareti başa gelecekti.

Data: File is tum.txt; komut satırında da, tum.txt isimli veri setimiz çağrılmaktadır.

Yukarıdaki program çalıştırıldıktan sonra elde edilen çıktıdaki bir takım uyum kriterlerine bakılır. Karışımli modellemede kullanılan uyum kriterleri : a) Akaike'nin bilgi kriteri (AIC ; Akaike's Information Criterion), b) Bayesian bilgi kriteri (BIC ; Bayesian Information Criterion) ve c) Log-olabilirlik (Log-likelihood) değerleridir. Daha sonra iki ve daha fazla sınıf için program kodlanır ve eldeki tüm çıktılardaki bilgi kriterleri karşılaştırılır. Entropi'si en yüksek, AIC ya da BIC değeri en düşük programdaki sınıf sayısı, alt grup (latent class) sayısı olarak kabul edilir ki, kabul edilen sınıf sayısındaki parametre tahmin değerleri (ortalama, varyans gibi) ve etki unsurlarının kovaryans değerleri gibi istatistik değerler kabul edilip, yorumlanır. Yang'a (1998) göre, ideal sınıf sayısını belirleyen uygun ölçütün BIC olduğu şeklindedir. Buna karşın, AIC, BIC ölçütlerinin bir arada kullanılması da savunulmaktadır (Muthen ve Muthen 2002; Roeder ve ark. 1999). Bundan dolayı bu çalışmada, iki kriterin birlikte ele alınması uygun görülmüştür. Latent class modellerde iterasyon tamamlandıktan sonra elde edilen log-olabilirlik değerine bağlı olarak; $AIC = -\text{LogL} + 2p$, $BIC = -\text{LogL} + p \ln(n)$ eşitlikleri ile uyum kriterleri (AIC, BIC) elde edilir. AIC ve BIC eşitliklerindeki p, tahmin edilecek parametre sayısını göstermektedir. AIC ve BIC için belirtilen eşitlikte log-olabilirlik (LogL) değeri negatif (-LogL) ele alındığı için, optimal sınıf sayısı en küçük AIC ya da BIC değerine göre değerlendirilmektedir. Eşitlikte $-\text{LogL}$, pozitif ele alınmış olsaydı bu durumda en büyük AIC ya da BIC değeri optimal sınıf sayısını belirleyecekti. Ancak M-Plus yazılımının arka planında log-olabilirlik değeri negatif formülüne edildiği için en küçük AIC, BIC değeri dikkate alınmaktadır.

Bulgular

Çalışmada oluşturulan yazılım toplam 6 ayrı sınıf için değerlendirilmiştir. Diğer sınıflar için oluşturulan kodlamalar ise yukarıda belirtildiği gibi "ekler" kısmında gösterilmiştir. M-Plus'ta belli bir sınıf sayısından sonra AIC, BIC ve entropy değerleri bozulmaya doğru bir eğilim göstermektedir; bu durumda programlamaya son verilip, mevcut sınıf sayılarında ideal ölçüt değerlerine sahip programdaki sınıf sayısı, popülasyon için ideal alt sınıf sayısı olarak kabul edilmektedir. Çalışmada sınıf sayısının 2, 3, 4, 5, 6 olduğu durumlar için ayrı ayrı yazılım programları hazırlanmış ve bu yazılım programlarının uyum ölçütleri Çizelge 1'de verilmiştir.

Çizelge 1 incelendiğinde, popülasyon sayısı arttıkça BIC değerinin düşmekte ve buna karşın Entropy değerinin artmakta olduğu görülmektedir. Burada Entropi, alt sınıflara üye edilen bireylerin söz konusu alt sınıflarda olma olasılıkları şeklinde tanımlanabilir. Çizelge incelendiğinde, en ideal modelin, alt popülasyon sayısının beş olarak belirlendiği sınıf olduğu görülmektedir ($BIC = 5237.147$). İncelenmesi gereken parametreler ve kovaryans ilişkileri, bu programın elde ettiği çıktıları bakılarak yorumlanmalıdır. Alt popülasyon 6'daki uyum ölçüt kriterlerinin (AIC, Entropy) bozulmaya doğru bir eğilim gösterdiği gözlemlenmektedir. Bunun için alt popülasyon 5'le model sonlandırılmalıdır.

Çizelge 1. Farklı alt sınıflar için uyum ölçütleri

Alt Sınıflar	Uyum Ölçütleri		
	AIC (Akaike's Information Criterion)	BIC (Schwarz's Bayesian Criterion)	Entropy
1	5196.246	5552.876	0.542
2	5190.140	5287.200	0.615
3	5176.126	5334.951	0.558
4	5177.051	5397.642	0.560
5	5158.790	5237.147	0.700
6	5173.871	5517.992	0.682

Bu bilgilere göre popülasyonun kendi içinde 5 homojen popülasyon gibi düşünülüp, her alt popülasyon için farklı parametre tahmini yapılmalıdır.

Çizelge 2. Her alt popülasyona düşen birey sayısı

Alt Sınıflar	Birey sayısı
Alt Sınıf 1	139
Alt Sınıf 2	51
Alt Sınıf 3	174
Alt Sınıf 4	160
Alt Sınıf 5	85

Her sınıftaki birey sayısının diğer sınıflarda olma olasılığı ise program tarafından aşağıdaki gibi tespit edilmiştir:

	1	2	3	4	5
Sınıf 1	0.872	0.041	0.055	0.020	0.011
Sınıf 2	0.068	0.794	0.065	0.012	0.061
Sınıf 3	0.066	0.032	0.770	0.065	0.067
Sınıf 4	0.015	0.010	0.081	0.814	0.080
Sınıf 5	0.010	0.037	0.077	0.085	0.791

Yukarıdaki veriler incelendiğinde, birinci sınıfta olan bireylerin alt popülasyon 1'de olma olasılığı 0.872; birinci sınıf bireylerinin alt popülasyon 2'de olma olasılığı 0.068, alt popülasyon 3'de olma olasılığı 0.066, alt popülasyon 4'de olma olasılığı ise 0.015, alt popülasyon 5'de olma olasılığı ise 0.010 düzeyindedir. Benzer yorumlamalar diğer alt popülasyonlar için de yapılır.

M-Plus'ın ideal tespit ettiği alt-sınıflara ait parametre tahmin çıktıları ekler kısmında gösterilmiş olup, karışımli modelin çok kapsamlı olmasından dolayı çıktılarıdaki ayrıntılar ve parametre tahminlerinin yorumu ayrı bir çalışmada ele alınacaktır.

Sonuçlar ve Tartışma

Bu araştırmada latent class gibi ileri düzey istatistik tekniklerinin analiz edilebildiği M-Plus yazılımı tanıtılmaya çalışılmıştır. M-Plus'ı tanıtımı aşamasında latent class ailesinden karışımli model örneği uygulanmaya çalışılmıştır. 643 bireyden oluşan heterojen bir veri seti örneği üzerinde, M-Plus'ın karışımli model tekniği ile

homojen alt popülasyon oluştururken kullanmış olduğu kodların nasıl hazırlandığına da değinilmiştir. Optimal homojen alt sınıf sayısının tespitinde kullanılan uyum kriterlerinin (AIC, BIC) M-Plus çıktılarından nasıl yorumlanacağı ve aynı zamanda bu kriterlerin literatürdeki konumuna göz atılmıştır. Çok değişkenli analiz ailesinden kümeleme (clustering) analiz tekniğinin de homojen alt popülasyon oluşturabildiği bilinmekle birlikte (AIC ve BIC ile), alt popülasyonlara ait parametre tahminlemesini yeterli düzeyde yapamadığı bildirilmektedir (Otlu ve Alpar 2006). Bu anlamda Latent class analizlerin (mixture gibi) çok güçlü bir alternatif ve kullanmış olduğu güçlü algoritmalarla sağlam ve sapmasız parametre tahminleme yapabildiği kabul edilmektedir (Nylund ve ark. 2006; Lazarfelds ve Hanry 1968). Tüm bu güçlü yönleri ile Latent class analizlerin kolay bir ara yüz desteği ile M-Plus'ta modellenilebildiği bildirilmektedir (Muthen ve ark. 2006).

Bu çalışmada kısmen M-Plus tanıtılırken, aynı zamanda karışımli modellerde kullanılan bir takım algoritmaların (EM gibi) matematik alt yapısına da değinildi. Ayrıca uyum kriterlerinin log-olabilirlik tabanında nasıl bir matematik alt yapıya sahip olduğu tartışıldı.

Sonuç olarak, bu çalışmada gözlem sayısı çok olan heterojen veri setlerine ait modellemeler söz konusu olduğunda ve gözlenen değişkenlerin hem sürekli hem de kategorik (semi-parametric) olmaları durumunda M-Plus ortamında latent class analizlerin uygulama kolaylığı gösterilmiş olup, örnek bir veri seti M-Plus ortamında kodlanmıştır.

Kaynaklar

- Baltes, P. B., Nesselroade, J.R., 1979. History and Rationale of Longitudinal Research. In J.R. Nesselroade, & P.B. Baltes (Eds.), *Longitudinal Research in the Study of Behavior and Development*. Academic press. 40 p. New York, USA.
- Chan, Z., Kou, L., 2001. A note on the estimation of the multinomial logit model with random effects. *The American Statistician*, 55(2) : 89-95.
- Clogg, C.C., 1995. Latent Class Models. In G. Arminger, C.C. Clogg & M.E. Sobel (Eds.), *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. Plenum Pres. 311 p. New York, USA.
- Collins, L., Horn, J.L., 1991. *Best Methods for the Analyses of Change*. APA Press. 365 p. Washington, USA.
- Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of Royal Statistical Society*, 39 : 1-18.
- Dhanavanthan, P., 2000. Estimation of the parameters of compound intervened poisson distribution. *Biometrical Journal*, 42(3) : 315-320.
- Duncan, S. C., Duncan, T.E., 1996. A multivariate latent growth curve analysis of adolescent substance use. *Structural Equation Modeling*, 4: 323-347.
- Duncan, T. E., Susan, S. C., Strycker, L. A., Okut, H., 2002. Growth Mixture Modeling of Adolescent Alcohol Use Data: Chapter Addendum to an Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications. Oregon Research Institute, Oregon. 36.
- Everitt, B.S., Hand, D.J., 1981. *Finite Mixture Distributions*. Chapman and Hall. 159 p. London, England.
- Jansen, R. C., 1993. Maximum likelihood in an generalized linear finite mixture model by using the em algorithm. *Biometrics*, 49 (1) : 227-231.
- Jamshidian, M., Jennrich, R. I., 1997. Acceleration of the em algorithm by quasi-newton methods. *J.R. Statist. Soc. B.*, 59 (3) : 569- 587.
- Nylund, K.L., Asparouhov, T., Muthen, B., 2006. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 35 (3) : 439-460.
- Lazarfelds, P., Hanry, N., 1968. *Latent Structure Analysis*. Houghton Mifflin. 79 p. New York, USA.
- Lindsay, B. G., 1995. *Mixture Models : Theory , Geometry and Applications*. Institute of Mathematical Statistics. 365 p. Hayward.
- Muthen, B., Asparouhov, T., Rebollo, I., 2006. Advances in behavioral genetics modeling using mPlus: Applications of factor mixture modeling to twin data. *Twin Research and Human Genetics*, 9 (3): 313-324.
- Muthen, B. O., 2001. Second-generation structural equation modeling with combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In A. Collins L. & Sayer A. (Eds.), *New methods for the analyses for change* (pp. 291-322). New York, USA.
- Muthen, L. K., & Muthen, B., 1998. *Mplus: User's guide*. Muthen & Muthen CA. 389 p. Los Angeles, USA.
- Muthen, L. K., Muthen, B., 2002. *Mplus: User's guide*. Muthen & Muthen CA. 256 p. Los Angeles, USA..
- Otlu, H., Alpar, R., 2006. İki aşamalı küme analizi ve bir uygulama. 5-9 Eylül 2006 IX. Ulusal Biyoistatistik Kongresi, Zonguldak.
- Peel, D. ve McLachlan, G., 2000. *Finite Mixture Models*. Wiley – Interscience Publisher. 126 p. New York, USA.
- Roeder, K., Lynch, G. K., Nagin, S. D., 1999. Modeling uncertainty in latent class membership: A case study in criminology. *Journal of American Stat. Association*, 447: 766-776.
- Topuz, D. ve Çakır, M., 2002. Lojistik regresyon analiz tekniğinin eğitim bilimleri araştırmalarında uygulanabilirliği ile ilgili bir araştırma. *ALBÜ Eğitim Fakültesi Dergisi*, 3 (6): 56-81.
- Wang, P., Putterman, M.I., 1998. Mixed logistic regression models. *Journal of Agriculture, Biological and Environmental Statistics*, 3(2): 175-200.
- Wang, P., Cockburn, I.M., Putterman, M.L., 1998. Analyses of latent data-mixed poisson regression model approach. *Journal of Business and Economic Statistics*, 16(1) : 27-41.
- Titterington, D.M., Smith, A.F.M., Markov, U.E., 1985. *Statistical Analyses of Finite Mixture Distributions*. John Willey & Sons. 221 p. Chichester, U.K.
- Yang, C.C., 1998. *Finite mixture model selection with psychometric applications*. Unpublished doctoral dissertation. University of California, Los Angeles.
- Yeşilova, A., 2003. *Biyolojik Çalışmalardan Elde Edilen Kategorik Verilere Karışık Poisson Regresyon Analizinin Uygulanması*. (Doktora tezi). Yayınlanmamış Doktora Tezi, YYÜ. Fen Bilimleri Enstitüsü, Van.
- Zhang, H. P. ve Merikangas, K., 2000. A frailty model of segregation analysis: understanding the familial transmission of alcoholism. *Biometrics*, 56: 815-823.

Ekler

Ek 1: M-Plus'ın karışımli modele ait çıktısı.

THE MODEL ESTIMATION TERMINATED NORMALLY

TESTS OF MODEL FIT

Loglikelihood

H0 Value -2515.395

Information Criteria

Number of Free Parameters 64

Akaike (AIC) 5158.790

Bayesian (BIC) 5441.147

Sample-Size Adjusted BIC 5237.961

(n* = (n + 2) / 24)

Entropy 0.700

FINAL CLASS COUNTS AND PROPORTIONS OF TOTAL SAMPLE SIZE BASED ON ESTIMATED POSTERIOR PROBABILITIES

Class 1	139.38696	0.22888
Class 2	56.49506	0.09277
Class 3	164.31510	0.26981
Class 4	152.24504	0.24999
Class 5	96.55785	0.15855

CLASSIFICATION OF INDIVIDUALS BASED ON THEIR MOST LIKELY CLASS MEMBERSHIP

Class Counts and Proportions

Class 1	139	0.22824
Class 2	51	0.08374
Class 3	174	0.28571
Class 4	160	0.26273
Class 5	85	0.13957

Average Class Probabilities by Class

	1	2	3	4	5
Class 1	0.872	0.041	0.055	0.020	0.011
Class 2	0.068	0.794	0.065	0.012	0.061
Class 3	0.066	0.032	0.770	0.065	0.067
Class 4	0.015	0.010	0.081	0.814	0.080
Class 5	0.010	0.037	0.077	0.085	0.791

MODEL RESULTS

Estimates S.E. Est./S.E.

CLASS 1

OSS	ON			
OOBP		-0.094	0.246	-0.380
CINS		0.201	2.539	0.079
BABAMES		0.202	0.663	0.305
ANAMES		-1.273	0.866	-1.471
BOLGE		-0.153	1.311	-0.116
LISCIK		-0.560	1.816	-0.308

Intercepts				
OSS	204.753	24.711	8.286	

Residual Variances

OSS	54.685	6.499	8.415
-----	--------	-------	-------

CLASS 2

OSS	ON			
OOBP		0.281	0.438	0.640
CINS		7.269	3.373	2.155
BABAMES		-3.933	2.370	-1.660
ANAMES		-0.324	1.698	-0.191
BOLGE		-8.107	4.309	-1.881
LISCIK		5.873	4.334	1.355

Intercepts			
OSS	182.506	45.223	4.036

Residual Variances

OSS	54.685	6.499	8.415
-----	--------	-------	-------

CLASS 3

OSS	ON			
OOBP		-0.393	0.246	-1.600
CINS		-21.788	5.646	-3.859
BABAMES		-0.576	1.110	-0.519
ANAMES		12.276	3.096	3.965
BOLGE		0.593	1.732	0.342
LISCIK		-0.771	1.784	-0.432

Intercepts			
OSS	279.845	22.731	12.311

Residual Variances

OSS	54.685	6.499	8.415
-----	--------	-------	-------

CLASS 4

OSS	ON			
OOBP		-0.124	0.206	-0.599
CINS		15.970	4.277	3.734
BABAMES		1.422	0.554	2.569
ANAMES		-11.725	2.982	-3.931
BOLGE		2.803	1.244	2.253
LISCIK		0.130	1.499	0.087

Intercepts			
OSS	215.990	20.797	10.386

Residual Variances

OSS	54.685	6.499	8.415
-----	--------	-------	-------

CLASS 5

OSS	ON			
OOBP		1.871	0.318	5.880
CINS		-2.156	2.664	-0.809
BABAMES		-0.476	0.582	-0.817
ANAMES		3.309	0.823	4.020
BOLGE		-2.516	1.631	-1.543
LISCIK		6.536	1.269	5.150

Intercepts			
OSS	61.370	31.137	1.971

Residual Variances

OSS	54.685	6.499	8.415
-----	--------	-------	-------

LATENT CLASS REGRESSION MODEL PART

C#1	ON			
OOBP		-0.337	0.131	-2.576
CINS		-3.035	1.021	-2.971
BABAMES		-0.470	0.321	-1.465
ANAMES		-0.699	0.369	-1.897
BOLGE		0.524	0.636	0.825
LISCIK		1.434	0.521	2.755

C#2 ON

OOBP	0.051	0.080	0.641
CINS	1.421	0.893	1.591
BABAMES	-0.040	0.253	-0.158
ANAMES	0.044	0.218	0.203
BOLGE	-0.410	0.386	-1.062
LISCIK	-0.718	0.443	-1.620

C#3 ON

OOBP	-0.338	0.121	-2.800
CINS	-0.557	1.487	-0.374
BABAMES	-0.551	0.299	-1.842
ANAMES	-1.568	0.798	-1.964
BOLGE	0.712	0.529	1.347
LISCIK	1.281	0.606	2.115

C#4 ON

OOBP	-0.253	0.123	-2.059
CINS	-1.883	0.911	-2.066
BABAMES	-0.484	0.300	-1.616
ANAMES	-1.673	1.486	-1.125
BOLGE	0.463	0.528	0.877
LISCIK	1.163	0.571	2.036

Intercepts

C#1	31.442	11.760	2.674
C#2	-5.349	7.825	-0.684
C#3	28.572	10.998	2.598
C#4	24.927	10.260	2.430

Ek 2 : M-Plus editöründe hazırlanan yazılımlar (syntax)

Class 1 için hazırlanan yazılım kodları

```
Data: File is tum.txt;
Variable: Names are birey kimlik OSS
         OOBP alan milli derece
         cins mekiksay tmeK kan
         boy kilo babames anames bolge liscik;
Missing = All (-9);
Usevariables OSS OOBP cins babames anames bolge
liscik;
Classes = c(1);
Analysis: Type = mixture;
Estimator=MLR;
miteration = 55000;
Model: %overall%
! [OSS];
OSS;
OSS on OOBP cins babames anames bolge liscik;
c1# on OOBP cins babames anames bolge liscik;
%c#1%
[OSS*195];
! [pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
! savedata: file='tum-mur2.dat';
! save=cprob;
Output: tech1 tech7 ;
```

Class 2 için hazırlanan yazılım kodları:

```
Data: File is tum.txt;
Variable: Names are birey kimlik OSS
         OOBP alan milli derece
         cins mekiksay tmeK kan
         boy kilo babames anames bolge liscik;
Missing = All (-9);
Usevariables OSS OOBP cins babames anames bolge
liscik;
Classes = c(2);
Analysis: Type = mixture;
Estimator=MLR;
miteration = 55000;
Model: %overall%
! [OSS];
OSS;
OSS on OOBP cins babames anames bolge liscik;
c#1 on OOBP cins babames anames bolge liscik;
%c#1%
```

```
[OSS*195];
! [pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#2%
[OSS*212];
! [pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
savedata: file='tum-mur2.dat';
save=cprob;
Output: tech1 tech7 ;
```

Class 3 için hazırlanan yazılım kodları:

```
Data: File is tum.txt;
Variable: Names are birey kimlik OSS
         OOBP alan milli derece
         cins mekiksay tmeK kan
         boy kilo babames anames bolge liscik;
Missing = All (-9);
Usevariables OSS OOBP cins babames anames bolge
liscik;
Classes = c(3);
Analysis: Type = mixture;
Estimator=MLR;
miteration = 55000;
Model: %overall%
! [OSS];
OSS;
OSS on OOBP cins babames anames bolge liscik;
c#1 on OOBP cins babames anames bolge liscik;
c#2 on OOBP cins babames anames bolge liscik;
%c#1%
[OSS*175];
! [pint*4.231 pslp*-.789];
OSS on OOBP cins babames anames bolge liscik;
%c#2%
[OSS*222];
! [pint*6.231 pslp*-.1989];
OSS on OOBP cins babames anames bolge liscik;
%c#3%
[OSS*248];
! [pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
savedata: file='tum-mur3.dat';
save=cprob;
Output: tech1 tech7 ;
```

Class 4 için hazırlanan yazılım kodları:

```
Data: File is tum.txt;
Variable: Names are birey kimlik OSS
         OOBP alan milli derece
         cins mekiksay tmeK kan
         boy kilo babames anames bolge liscik;
Missing = All (-9);
Usevariables OSS OOBP cins babames anames bolge
liscik;
Classes = c(4);
Analysis: Type = mixture;
Estimator=MLR;
miteration = 55000;
Model: %overall%
! [OSS];
OSS;
OSS on OOBP cins babames anames bolge liscik;
c#1 on OOBP cins babames anames bolge liscik;
c#2 on OOBP cins babames anames bolge liscik;
c#3 on OOBP cins babames anames bolge liscik;
%c#1%
[OSS*195];
! [pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#2%
[OSS*212];
! [pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#3%
```



```
[OSS*218];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#4%
[OSS*222];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
savedata: file='tum-mur4.dat';
save=cprob;
Output: tech1 tech7 ;
```

Class 5 için hazırlanan yazılım kodları:

```
Data: File is tum.txt;
Variable: Names are birey kimlik OSS
         OOBP alan milli derece
         cins mekiksay tmeK kan
         boy kilo babames anames bolge liscik;
Missing = All (-9);
Usevariables OSS OOBP cins babames anames bolge
liscik;
Classes = c(5);
Analysis: Type = mixture;
         Estimator=MLR;
         miteration = 55000;
Model: %overall%
       ! [OSS];
       OSS;
       OSS on OOBP cins babames anames bolge liscik;
c#1 on OOBP cins babames anames bolge liscik;
c#2 on OOBP cins babames anames bolge liscik;
c#3 on OOBP cins babames anames bolge liscik;
c#4 on OOBP cins babames anames bolge liscik;
```

```
%c#1%
[OSS*195];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#2%
[OSS*212];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#3%
[OSS*218];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#4%
[OSS*222];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#5%
[OSS*222];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
```

```
savedata: file='tum-mur5.dat';
save=cprob;
Output: tech1 tech7 ;
```

Class 6 için hazırlanan yazılım kodları:

```
Data: File is tum.txt;
Variable: Names are birey kimlik OSS
         OOBP alan milli derece
         cins mekiksay tmeK kan
         boy kilo babames anames bolge liscik;
Missing = All (-9);
Usevariables OSS OOBP cins babames anames bolge
liscik;
Classes = c(6);
Analysis: Type = mixture;
         Estimator=MLR;
         miteration = 55000;
Model: %overall%
       ! [OSS];
       OSS;
       OSS on OOBP cins babames anames bolge liscik;
```

```
c#1 on OOBP cins babames anames bolge liscik;
c#2 on OOBP cins babames anames bolge liscik;
c#3 on OOBP cins babames anames bolge liscik;
c#4 on OOBP cins babames anames bolge liscik;
c#5 on OOBP cins babames anames bolge liscik;
%c#1%
[OSS*165];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#2%
[OSS*172];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#3%
[OSS*218];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#4%
[OSS*212];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#5%
[OSS*242];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
%c#6%
[OSS*262];
![[pint*7.231 pslp*-.989];
OSS on OOBP cins babames anames bolge liscik;
savedata: file='tum-mur6.dat';
save=cprob; Output: tech1 tech7 ;
```