

Örtük Markov Analizlerinde Model Seçimi Üzerine Bir Monte Carlo Simülasyon Çalışması

Duygu Güngör
Dokuz Eylül Üniversitesi

Selva Ülbe
Dokuz Eylül Üniversitesi

Samet Baş
Dokuz Eylül Üniversitesi

Özet

Örtük Markov modelleri, niceliksel olarak ölçüm alınmanın mümkün olmadığı boylamsal psikoloji çalışmalarında, kategorik gözlenen ve örtük değişkenlerin zamana bağlı değişimlerini analiz etmek ve yorumlamak için iyi bir alternatif olarak karşımıza çıkmaktadır. Ancak, son yıllarda giderek artan kullanıma rağmen Örtük Markov modellerindeki model seçim süreci konusunda henüz bir görüş birliğine varılamamıştır. Bu bağlamda, çalışmanın ilk amacı, tek değişkenli bir görgül veri seti kullanarak bir uygulama örneği sunmaktır. Bir diğer amacı da Monte Carlo simülasyon yöntemi ile oluşturulmuş veri setinden yararlanarak, madde tepki olasılıklarının gücü, ölçüm alınan zaman sayısı ve örneklem büyüklüğüne göre bilgi kriterlerinin model seçimi ve parametre tahmin yanlışlıklarına etkisini incelemektir. Sonuçlara göre, madde tepki olasılıklarının hem güçlü hem de zayıf olduğu koşullarda, ölçüm alınan üç farklı zaman koşulunda ya da örneklem büyüklüğü 200'den 2000'e yükseldiğinde BIC ve CAIC bilgi kriterleri kullanılarak yapılan model seçimlerinde %100 doğru karar oranı gözlenmiştir. Bulgular, alanyazın ışığında tartışılacaktır.

Anahtar kelimeler: Örtük Markov modeli, boylamsal veri, model seçimi

Abstract

Latent Markov models emerge as a good alternative for longitudinal psychological studies where it is not possible to take quantitative measurements to analyze and interpret time-dependent changes of categorical observed and latent variable(s). However, despite its increasing use in recent years, a consensus on the model selection process in the Latent Markov models has not been reached yet. In this context, the first objective of this research was to provide an application example by using an empirical dataset with a single variable. Another aim was to examine the impacts of the strength of item response probabilities, the number of times the measurement being taken and sample size on model selection and parameter estimation bias based on using the dataset generated by Monte Carlo simulation method. As a result, using BIC and CAIC information criteria, 100% correct decision rate was observed regardless to item response probabilities (weak or strong), and number of measurement (2 or 3) when the sample size increased from 200 to 2000. The findings were discussed in the light of the related literature.

Keywords: Latent Markov model, longitudinal data, model selection

Teşekkür Notu: Bu çalışmadaki görgül örnek için bizimle verilerini paylaşan Günaydın Group şirketine ve desteklerini esirgemeyen Sayın Erman Culha'ya teşekkür ederiz.

Bu çalışmanın bir kısmında birinci yazar TÜBİTAK 2219 programı kapsamında desteklenmiştir.

Yazışma Adresi: Doç. Dr. Duygu Güngör, Dokuz Eylül Üniversitesi, Edebiyat Fakültesi, Psikoloji Bölümü, Tınaztepe Yerleşkesi Adatepe Mah. Doğuş Cad. No: 207/M 35390 Buca / İZMİR

E-posta: duygu.gungor@deu.edu.tr

Gönderim Tarihi: 19.12.2017

Kabul Tarihi: 21.06.2018

Psikoloji arařtırmalarında zamana baėlı deėişim pek çok arařtırmaya konu olmaktadır. Bu tip arařtırmalarda geleneksel olarak tekrarlı ölçümler için varyans analizi gibi istatistiksel yöntemler sıklıkla kullanılmaktadır. Son yıllarda ise yapısal eşitlik modelleri kapsamında yer alan örtük gelişme modelleriyle (*latent growth models*) örtük deėişkenin sürekli olduėu durumlarda zamana baėlı deėişim incelenmektedir. Bahsi geçen analiz yöntemlerinin tamamı zamana baėlı deėişimin niceliksel olduėu varsayımıyla, bireylerin arařtırmaya konu olan deėişkenlerden aldıėı puanların artış ya da azalış miktarına odaklanmaktadır. Bununla birlikte, psikoloji arařtırmalarında her zaman niceliksel ölçümler yapmak mümkün olmamaktadır. Örneėin, sürücü davranışlarının incelendiėi bir arařtırmada riskli davranışları puanlandırmak isteyen bir arařtırmacı için, kırmızı ışıkta geçmek ya da hız sınırını aşmak davranışlarından hangisinin daha riskli bir davranış olduėuna karar vermek mümkün olamayabilir. Bu tür niteliklerin ölçümünde başka bir deyişle gözlenen ve örtük deėişken ya da deėişkenlerin süreksiz yapıda olduėu durumlarda ise örtük geçiş modelleri (*latent transition models*) (Collins ve Lanza, 2010) olarak da adlandırılan örtük Markov modeli (*latent Markov model-ÖM model*) alternatif olarak karşımıza çıkmaktadır.

ÖM modelin kullanıldıėı uygulamalı alanlara madde kötüye kullanımı (Cosden, Larsen, Donahue ve Ny-lund-Gibson, 2015; Guo, Aveyard, Fielding ve Sutton, 2009; La Flair ve ark., 2013; Lanza ve Bray, 2010), yeme davranışları (Cain, Epler, Steinley ve Sher, 2010; Castellini ve ark., 2013), eğitim ve seyahat davranışları verilebilir. Son yıllarda bilgisayar yazılımlarındaki gelişmelere paralel olarak artan yaygın kullanıma rağmen ÖM modellerindeki model seçim süreci konusunda bir görüş birliėi bulunmamaktadır. Bu kapsamda, bu çalışmanın amacı Monte Carlo simülasyon yöntemini kullanarak örneklem büyüklüėü, madde koşullu olasılıklarının gücü ve ölçüm alınan zaman sayısına göre bilgi kriterlerinin örtük statü sayısını belirlemedeki gücünü ve parametre tahmin yanlılıklarını incelemektedir.

İzleyen bölümlerde öncelikle örtük Markov modeli tanımlanacaktır. Bu kapsamda tek deėişkenli bir görgül veri seti ile bir uygulama örneėi verilecektir. Sonrasında, analiz basamakları ve model seçim sürecine deėinilecek ve simülasyon çalıřması sunulacaktır.

Örtük Markov Modeli

ÖM modeli, ölçme modeli ve yapısal model olarak adlandırılan iki alt bölümden oluşmaktadır. Ölçme modeli gözlenen deėişkenlerin örtük deėişkenlerle ilişkisinin tespit edildiėi kısımdır. İkinci kısım Markov model ise zamana baėlı olarak örtük statüler arasındaki geçiş olasılıklarını vermektedir. Bu kapsamda üç temel para-

metre tahmin edilmektedir: koşullu olasılıklar (*conditional probabilities*), örtük statü olasılıkları (*latent status prevalences - initial probabilities*) ve geçiş olasılıkları (*transition probabilities*) (Collins ve Flaherty, 2002; Collins ve Lanza, 2010; Lanza ve Bray, 2010; Vermunt, Tran ve Magidson, 2008).

$i = 1, 2, 3, \dots, n$, $j = 1, 2, 3, \dots, p$ ve $t = 1, 2, 3, \dots, T$ olmak üzere y_{ijt} i. katılımcının j. deėişkene t. zamanda verdiėi tepki olsun. i. katılımcının t. zamandaki tepkileri y_{it} ve tüm zamanlardaki tepkileri y_i vektörüyle, örtük statülerin olası deėerleri $x_i = 1, 2, 3, \dots, S$ olmak üzere t zamanındaki örtük statüleri X_i ile gösterilsin. Model parametrelerinin vektörünü tanımlamak için kullanılan F sembolü ile Gudicha, Schmittman ve Vermunt'un (2015) gösterimiyle örtük Markov modeli eşitlik 1'deki gibidir:

$$P(y_i; \Phi) = \sum_{x_{t1}} \sum_{x_{t2}} \dots \sum_{x_{tT}} p(X_t = x_t) \prod_{t=1}^T p(X_t = x_t | X_{t-1} = x_{t-1}) \prod_{j=1}^p p(y_{ijt} | X_t = x_t) \quad (1)$$

Eşitlikteki üç temel parametre: örtük statü olasılıkları $p(X_t = x_t)$, koşullu tepki olasılıkları $p(y_{ijt} | X_t = x_t)$ ve geçiş olasılıkları $p(X_t = x_t | X_{t-1} = x_{t-1})$ şeklindedir. Bu parametrelerden ilk ikisi örtük sınıf analizindeki parametrelerle doğrudan ilişkilidir. Örtük sınıf analizinde örtük sınıf olasılıėı olarak adlandırılan parametre örtük Markov modellerinde sınıfın dinamik yapısını tanımlamak amacıyla genellikle örtük statü olasılıėı kavramı ile adlandırılmaktadır (Collins ve Lanza, 2010). Herhangi bir kısıt konulmamış modellerde örtük statü olasılıkları ölçüm alınan her zaman dilimi için tahmin edilmekte ve ait olduėu zaman dilimde gözlemlerin/katılımcıların örtük statüdeki yaygınlıėını ifade etmektedir. Sürücü davranışlarının incelendiėi hipotetik bir arařtırma üzerinden düşünülecek olursa, dikkatli sürüş ve riskli sürüş olarak tanımlanabilecek iki örtük statünün tespit edilmesi mümkün olabilir. Böylesi bir durumda I. zaman diliminde dikkatli örtük statüsünde olma olasılıėı örneėin .70 olarak tespit edildiyse, katılımcıların %70'inin ölçüm alınan I. zaman diliminde bu örtük statüde yer aldıklarını söylemek mümkün olacaktır. Her zaman dilimi için örtük statü olasılıklarının toplamı bire eşit olacaėından 2. örtük statüde olma olasılıėı da $1 - .70 = .30$ olarak hesaplanabilir. II. zaman diliminde ise bu örtük statü olasılık daėılımları örtük geçiş olasılıklarına baėlı olarak deėişecektir.

Koşullu olasılıklar ise faktör analizindeki faktör yüklerine benzemekte ve örtük sınıf analizindeki gibi yorumlanmaktadır. Yukarıdaki hipotetik örnek üzerinden düşünülecek olursak j gözlenen deėişkeni "Kırmızı ışık ihlali yapmam" olsun. t zamanında I. örtük statüde yer alan katılımcıların bu deėişkene ait "Evet" deme koşullu olasılıėı .80 olarak tahmin edilirse, I. örtük statüde yer alan bireyler riskli sürücü davranış sergileyenler olarak adlandırılabilir. Örtük statü olasılıkları ile benzer şekilde koşullu olasılıkların da toplamı bire eşit olaca-

ğından “Evet”, “Hayır” kategorili hipotetik örneğimizde t . zamanda riskli sürücü davranışı olarak adlandırdığımız 1. örtük statüde yer alan bir katılımcının kırmızı ışık ihlali yapmam demeye olasılığı $1 - .80 = .20$ olacaktır. ÖM modeli sunulan hipotetik örnekteki gibi özbidirime dayalı gözlenen değişkenlerle kullanılabilmesi gibi, ceza kayıtları, gözlem, görüşme gibi pek çok farklı yöntemle elde edilen kategorik gözlenen değişkenlerle de kullanılabilir.

Geçiş olasılıklarının hesaplandığı kısım modelin Markov kısmını oluşturmaktadır. $T-1$. zamanda 1. örtük statüde olan bir bireyin/gözlemin t . zamanda da 1. örtük statüde olma koşullu olasılığı, $\pi_{1|1} = p(X_t = 1 | X_{t-1} = 1)$ olacaktır. $\pi_{2|1} = p(X_t = 2 | X_{t-1} = 1)$ ise $t-1$. zamanda 1. örtük statüde yer alan bireyin t . zamanda 2. örtük statüye geçiş olasılığıdır. $T-1$ zamanında alınan ölçümlerde bir kişi ancak bir örtük statüde yer alabilir ve örtük statü olasılıkları toplamı her t zamanı için 1'e eşit olacaktır. Ancak t . zamanda kişiler A örtük statüsünden B örtük statüsüne geçiş yapabilir ve statülerin yaygınlığı t . ve $t-1$. zamanda farklılaşabilir. A ve B örtük statülerinin olduğu iki zamanlı bir araştırmada olası geçişler, A-A, A-B, B-A, B-B şeklindedir. ÖM modellerinde geçiş olasılıkları sıklıkla örtük geçiş matrisi ($A^{(t)}$) ile gösterilmektedir. Matrisin köşegen elemanları aynı örtük statüde kalma olasılıklarını gösterirken, diğer elemanlar farklı statülere geçme olasılıklarını göstermektedir. Aşağıda sunulan örnek matriste üç örtük statü ve iki zamanlı bir geçiş matrisi görülmektedir.

$$A^{(t)} = \begin{bmatrix} .90 & .10 & .00 \\ .20 & .80 & .00 \\ .50 & .45 & .05 \end{bmatrix}$$

Bu matris şu şekilde yorumlanabilir: Birinci örtük statüde yer alan bireylerin %90'ı I. zamandan II. zamana geçtiklerinde aynı örtük statüde yer alırken, %10'u 2. örtük statüye geçiş yapmıştır. İkinci örtük statüdeki katılımcıların ise %80'i aynı örtük statüde yer alırken, %20'si 1. örtük statüye geçiş yapmıştır. En yoğun geçiş ise 3. örtük statü üyelerinde gözlenmiştir. Bu örtük statüde yer alan bireylerin sadece %5'i aynı statüde kalırken, %45'i 2. örtük statüye, %50'si de 1. örtük statüye geçiş yapmıştır. Bu hipotetik örnek gelişimsel olarak 1. örtük statüdekilerin yetenek düzeyinin daha iyi olduğu ve 3. statüdekilerin gelişmesinin beklendiği durumlarda karşımıza çıkabilir.

Geçiş olasılıkları her zaman dilimi için eşdeğer zaman-homojen geçiş (*time-homogenous transition*) olabileceği gibi zamana bağlı olarak değişen biçimde de zaman-heterojen gözlenmesi mümkündür. İki örtük statü olan bir modelde zaman-homojen geçiş olasılıkları olduğunu ve ilk ölçüm alınan zamanda ait olunan örtük

statüde kalma olasılığının .80 olduğunu varsayalım. Bu durumda diğer örtük statüye geçiş olasılığı .20 olacaktır. Üç zamanda ölçüm alınan bir araştırma için ikinci zamandan üçüncü zamana geçişte de bu olasılıklar aynı olacaktır. Zaman-heterojen geçiş olasılıkları ise her zaman diliminde farklı geçiş olasılıklarının tahmin edilmesini gerektirmektedir. Bu modellerden hangisinin daha iyi olduğu da model seçim sürecinde karar verilmesi gereken durumlardan biridir.

Örtük Markov modelin iki temel sayıltısı bulunmaktadır: yerel bağımsızlık ve birinci düzey Markov varsayımı. Yerel bağımsızlık varsayımı t zamanında gözlenen değişkenlerin sadece geçerli zamanla ilişkili olduğu varsayımdır. Birinci düzey Markov varsayımı ise, t zamanındaki statülerin sadece $t-1$. zamandan etkilendiği yönündedir (Bartolucci, Farcomeni ve Pennoni, 2013; Vermunt, Langeheine ve Bockenholt, 1999).

Model seçimi. Örtük Markov analizinde ilk adım ölçüm alınan her zaman dilimi için sınıf sayısının belirlenmesini oluşturmaktadır. Örtük statü sayısının önsel olarak bilindiği durumlarda ise doğrulayıcı bir yaklaşım izlenerek bu adım atlanabilmektedir (Bartolucci ve ark., 2013). Model seçim sürecinde genel olarak analizlere ölçüm alınan her zaman dilimi için ayrı ayrı örtük statü sayılarının belirlenmesi ile başlanır. Daha sonra statü sayılarına kısıt konarak koşullu olasılıklar ve örtük geçiş olasılıkları tespit edilir. Bu modeller de incelenerek eğer geçiş olasılıklarına da kısıt konabilecekse; örneğin Zaman-1'de hiperaktif örtük statüsünde olan bir bireyin Zaman-2'de normal statüsünde olma olasılığına sıfır değeri atamak gibi, çeşitli kısıtlar koyarak iç içe geçmiş modeller test edilir. Benzer şekilde, koşullu olasılıkları da Zaman-1 ve Zaman-2 için eş değer tanımlamak mümkündür. İç içe geçmiş bu modellerden en iyi modeli seçmek için kullanılan yöntemlerden biri en çok olabirlik oranı-kikare (*likelihood ratio chi-squared- L²*) fark istatistiklerine göre karar vermektir. Karşılaştırılacak iki model için hesaplanan L^2 değerleri arasındaki fark hesaplanarak en çok olabirlik (LR) değeri elde edilmektedir. LR'nin istatistiksel olarak anlamlı olup olmadığı da yine söz konusu iki modelin serbestlik dereceleri arasındaki farkın kritik değeriyle karşılaştırılması ile elde edilir. Alternatif bir yol olarak Akaike Bilgi Kriteri (AIC) ve Bayesyen Bilgi Kriterinin (BIC) kullanılması genel olarak kabul görmüştür. “ df ” serbestlik derecesi ve N örneklem büyüklüğü olsun. Bu durumda AIC ve BIC aşağıdaki eşitliklerle tanımlanabilir:

$$AIC = L^2 - 2 \cdot df \quad (2)$$

$$BIC = L^2 - \ln(N) \cdot df \quad (3)$$

Eşitlik 3'den de izlenebileceği gibi, BIC genellikle örneklem büyüklüğü geniş olduğunda tercih edilen bir

Tablo 1. Yıllara Göre Hız İhlallerinin Frekans Tablosu

Ceza Dönemi	Ocak- Haziran 2013	Temmuz-Aralık 2013	Ocak-Haziran 2014	Temmuz-Aralık 2014	Ocak-Haziran 2015
Frekans	21	23	29	40	40
Yüzde	4.1	4.5	5.7	7.8	7.8

bilgi kriteridir. Bununla birlikte CAIC, AIC3, DBIC gibi türevleri de var olan kriterlerden hangisinin kullanılmasının uygun olduğuna dair bir görüş birliği bulunmamaktadır. Hangi bilgi kriterinin kullanıldığından bağımsız olarak bilgi kriterinin aldığı daha küçük değer daha iyi modele işaret etmektedir. Alan yazında model seçiminde örtük Markov modelleri benzer kriterlerin kullanıldığı örtük sınıf analizinin incelendiği pek çok yayın bulunmaktadır (Bauer ve Curran, 2003; Dias, 2006; Lin, 2006; 2012; Nylund, Asparouhov ve Muthen, 2007; Vrieze, 2012; Yang ve Yang, 2007). Örtük Markov modellerin psikoloji verilerinde kullanılmasına ilişkin yaptıkları simülasyon çalışmaları oldukça az sayıda olmakla birlikte Bartolucci ve arkadaşları (2013), Collins ve Wugalter (1992) ve Visser, Rajimakers ve Molenaar (2002) örnek olarak verilebilir. Bu araştırmalara ait bulgular bu çalışmanın tartışma bölümünde sunulacaktır.

Örtük Markov modelleri kullanılarak yapılan uygulamalı araştırmaların bazıları incelendiğinde ise model seçiminde kullanılan bilgi kriterlerinin şu şekilde olduğu görülmüştür. Catts ve arkadaşları (2012) 493 çocuk katılımcı ile yaptıkları araştırmada, anasınıfı, 2, 4 ve 8. sınıflardaki çocuklardan ölçümler almışlar ve okuma yeteneklerinin zamana bağlı değişimlerini incelemişler, model seçimi için ise L^2 istatistiklerini kullanmayı tercih etmişlerdir. Seaton ve arkadaşları (2012), 566 Afrika-Amerikalı ergen ile üç yıl boyunca yürüttükleri çalışmada ırk ayrımcılığı ve ırk sosyalizasyonu algılarının etnik kimlik gelişimleri üzerindeki etkisini inceledikleri çalışmada BIC bilgi kriterini kullanmışlardır. Marcoulides ve arkadaşları (2008) 130 katılımcının 9, 10, 13, 16, 17 yaşlarında aldıkları akademik motivasyonla ilgili ölçümlerin model seçimi için BIC kullanmışlardır. Schmittmann ve arkadaşları (2005) 101 çocuk ile yaptıkları deneysel araştırmada BIC ve AICc'yi kullanırken, Castellini ve arkadaşları (2013) 856 hastanın yeme bozuklukları ile ilgili yaptıkları çalışmada altı yılda üç ölçüm almışlar ve BIC, CAIC bilgi kriterlerini kullanmışlardır. Lanza ve Bray (2010) ise ölçme eşdeğerliğini de inceledikleri kadınların madde kullanımına ilişkin yaptıkları araştırmada BIC bilgi kriterinin teoriye uygun modeli seçmede zayıf kaldığını belirtmiş, AIC, CAIC ve ABIC bilgi kriterlerini de dikkate almışlardır.

Son yıllarda Latent Gold (Vermunt ve Magidson,

2013), Mplus (Muthen ve Muthen, 1998-2007) gibi ticari PROC LTA (Lanza ve Collins, 2008) ve R (Bartolucci ve Pandolfi, 2016) gibi özgür bilgisayar yazılımların gelişmesiyle uygulamalı alanda sıklıkla kullanılmaya başlanan örtük Markov modelleri için gerekli örneklem sayısı, gösterge değişkeni ve zaman sayısının parametreler üzerindeki etkisi ile ilgili yapılan çalışmalar ise oldukça sınırlıdır. Bu kapsamda bu çalışmanın ilk amacı bir örnek uygulama sunmak; ikinci amacı ise örneklem genişliği, madde sayısı, güçlü ve zayıf koşullu olasılıkların model seçimi ve parametre tahminine etkisini Monte Carlo simülasyon yöntemi ile incelemektir.

Yöntem

Örneklem

Bu örneğin örneklemini özel bir taşımacılık firmasında çalışan 511 profesyonel sürücü oluşturmaktadır. Sürücülerin tamamı erkek olup, günlük ortalama sekiz saat araç kullanmaktadır. Firmanın izni ile sürücülere ait 2013, 2014 ve 2015 yılının ilk yarısındaki hız ihlali kayıtlarına ulaşılmıştır. Yıllar iki yarı döneme bölünerek beş zamanlı bir veri seti elde edilmiştir. Her dönemdeki hız limiti ihlallerine ait frekans değerleri Tablo 1'de sunulmuştur. Yıllar içinde hız ihlali değerlerinin %4 ile %8 arasında olduğu görülmektedir. Böylesi bir frekans tablosu bize genel eğilim hakkında bilgi verirken, kişilere özgü eğilimleri belirlemek için daha detaylı analizlere ihtiyacımız vardır. Örtük Markov analizleri ile hangi katılımcının hangi örtük statüye geçtiği bilgisi elde edilebilir. Bununla birlikte bir statü üyeliğinden başka bir statüye geçiş olasılıkları hesaplandığı için geleceğe dair yordamalar yapmak mümkündür.

İşlem

Beş zaman diliminde alınan tek gözlenen değişkenli veri seti Latent GOLD 5.1 Syntax versiyonu ile analiz edilmiştir. İki tepki kategorisi varlığında tek değişken olduğu için iki örtük statüde ait parametreler tahmin edilmiştir. Model-1 iki örtük statü ve geçiş olasılıklarının zaman-homojen olduğu model olarak test edilmiştir. Model-2 ise yine iki örtük statünün olduğu ancak geçiş olasılıklarının zaman-heterojen olarak test edildiği modeldir. Modellere ait sentakslar ekte sunul-

Tablo 2. Hız İhlallerine ait Örtük Markov Model Bulguları

Model	BIC	AIC	AIC3	CAIC
Zaman-homojen model	-118.384	-8.238	-34.2384	-144.384
Zaman-heterojen model	-116.06	-14.39	-38.39	-140.06

Not. BIC = Bayesian Bilgi Kriteri, AIC = Akaike Bilgi Kriteri, AIC3 = Akaike Bilgi Kriteri 3, CAIC = Tutarlı Akaike Bilgi Kriteri.

muştur.

Bulgular

İki-örtük statülü zaman-homojen ve zaman-heterojen modellere ait bilgi kriteri değerleri yukarıda sunulmuştur:

Tablo 2’den de görüleceği üzere BIC ve CAIC bilgi kriterleri zaman-homojen modeline işaret ederken; AIC, AIC3 bilgi kriterleri zaman-heterojen modeline işaret etmişlerdir. Yıllara göre farklı geçiş olasılıkları da bu çalışmada merak konusu olabileceği için AIC3 bilgi kriterinin de örtük sınıf analizinde daha tutarlı sonuçlar verdiği (Güngör, Korkmaz ve Sazak, 2015) bilgisine dayanarak zaman-heterojen modele ait parametreler yorumlanmıştır.

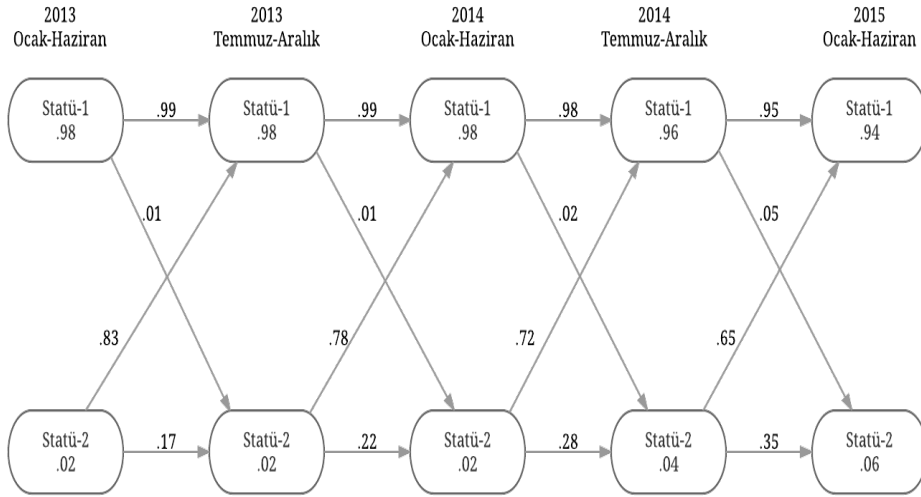
Madde-tepki olasılıkları incelendiğinde birinci örtük statüde olanların hız ihlali yapma olasılığı .01, hız ihlali yapmama olasılıkları .99; ikinci örtük statüde yer alanların ise hız ihlali yapma olasılıkları .66, hız ihlali yapmama olasılıkları .34 olarak bulunmuştur. Bu olasılıklara dayanarak birinci örtük statü “hız kuralına uyanlar”, ikinci örtük statü ise “hız ihlali yapanlar” olarak

adlandırılmıştır. Birinci ölçüm zamanındaki örtük statü olasılıkları incelendiğinde birinci örtük statüde olma olasılığının .98, ikinci örtük statüde olma olasılığının ise .02 olduğu bulunmuştur. Başka bir deyişle, 2013 yılının ilk yarısında sürücülerin %98’si hız ihlali yapmayanlar örtük statüsünde yer almıştır. 2014’ün ikinci yarısında ise bu olasılık .96’ya gerilemiş ve son olarak 2015 yılının ilk yarısında .93’e düşmüştür. Bu da 2015’in ilk yarısında hız ihlali yapan statüsünde yer alanların oranının %7’ye yükseldiği anlamına gelmektedir. Tahmin edilen örtük geçiş olasılıkları Şekil 1’de sunulmuştur.

Yöntem

Araştırma Deseni

Araştırma kapsamında öncelikle 18 farklı koşul için veri setleri Latent GOLD 5.0 (Vermunt ve Magidson, 2015) istatistik paket programının Monte Carlo simülasyon bölümü kullanılarak üretilmiştir. Araştırma deseninde üç farklı örneklem büyüklüğü (200, 600, 2000), iki farklı madde tepki olasılıklarının büyüklüğü (.1, .9 ; .3; .7), üç farklı ölçüm alınan zaman sayısı (2, 3,



Şekil 1. Tahmin Edilen Örtük Geçiş Olasılıkları

Tablo 3. Araştırma Deseni

		Örtük Statü Olasılıkları	
		Örtük Statü 1 .50	Örtük Statü 2 .50
Zaman I'den Zaman II'ye Örtük Geçiş Olasılıkları			
		Statü 1	Statü 2
Statü 1		.80	.20
Statü 2		.20	.80
Zaman II'den Zaman III'e Örtük Geçiş Olasılıkları			
		Statü 1	Statü 2
Statü 1		.80	.20
Statü 2		.20	.80
Madde Tepki Olasılıkları (Güçlü Madde Tepki Olasılığı Koşulu)			
Madde-1	0 1		Madde-5 0 1
Statü 1	.10 .90	...Madde-2...Madde-4...	Statü 1 .10 .90
Statü 2	.90 .10		Statü 2 .90 .10

4), model seçiminde kullanılan dört farklı bilgi kriterleri (BIC, CAIC, AIC3, AIC) olmak üzere dört bağımsız değişken yer almaktadır. Model seçiminde kullanılan bilgi kriterleri analiz aşamasında kullanılan bağımsız değişkendir. Bu kapsamda 18 (3x2x3) farklı koşul ve her koşul için 100 farklı tekrar olmak üzere 1800 veri seti üretilmiş ve analiz edilmiştir.

Üretilen tüm veri setlerinde; örtük statü sayısı iki, örtük statü olasılıkları .50, örtük geçiş olasılıkları .80 ve .20, madde sayısı beş ve madde kategori sayısı iki olarak alınmıştır. Örtük geçiş olasılıkları zamana bağlı olarak değişmeyen, başka bir deyişle homojen olarak değişimlenmiştir. Bu desene ait parametreler yukarıdaki tabloda sunulmuştur.

Bu tabloda sunulan parametreler matematik başarısının ölçüldüğü beş maddeli bir test üzerinden düşünülebilir. Bu durumda örneğin maddelere verilen doğru cevaplar "1", yanlış cevaplar "0" şeklinde kodlanmış olsun. Madde tepki olasılıkları incelendiğinde Statü-1 üyelerinin tüm maddeleri doğru cevaplama olasılıkları .90 olarak bulunmuştur. Başka bir deyişle, Statü-1'de yer alan bir bireyin soruları doğru cevaplama olasılığı %90 iken, Statü-2'de yer alan bir bireyin doğru cevap verme olasılığı sadece %10'dur. Bu olasılıklar göz önünde bulundurulduğunda Statü-1 *başarılı*, Statü-2 ise *gelişmeye açık* statüsü olarak adlandırılabilir.

Statülerin yaygınlığı hakkında bilgi veren örtük statü olasılıklarının .50 olduğu görülmektedir. Rastgele seçilen bir katılımcının 1. örtük statüde yer alma olasılığı .50'dir. Bu durumda örneğin 200 kişilik bir örneklemle çalışılıyorsa 100 kişinin *başarılı*, 100 kişinin *gelişmeye açık* statüsünde yer aldığı söylemek mümkündür.

Son olarak, örtük geçiş olasılıkları incelendiğinde I. zamanda *başarılı* örtük statüsünde yer alan katılımcıların II. zamanda da aynı örtük statüde yer alma olasılıkları .80 olarak bulunmuştur. Bu durumda *başarılı* örtük statüsündeki 100 kişiden 80'i aynı statüde yer alırken 20'si *gelişmeye açık* statüsüne gerilemiştir. *Gelişmeye açık* statüsünde yer alan 100 bireyden ise sadece 20'si II. zamanda *başarılı* statüsüne geçiş yapmıştır. Zaman-II ve Zaman-III arasındaki örtük geçiş olasılıkları da aynıdır. İyi bir eğitim ortamı ve yeterli düzeyde öğrenme olgunluğunun olduğu bir senaryoda ise örneğin Zaman-II'den Zaman-III'e *gelişmeye açık* statüsünden, *başarılı* statüsüne örtük geçiş olasılıklarının daha yüksek olması beklenebilirdi.

İşlem

Elde edilen veriler bir, iki, üç ve dört örtük statü olacak şekilde analiz edilmiştir. Analizler sonucunda her koşul için elde edilen L^2 değerleri kullanılarak bilgi kriterlerine ait değerler hesaplanmıştır. Bu bilgi kriterlerinin işaret ettiği modeller seçilerek raporlanmıştır. Örtük statü sayısı gerçekte iki olarak belirlendiğinden bu değerden sapma oranı yüzde olarak hesaplanmıştır. Ayrıca her koşula ait parametre değerlerinin ortalama ve standart sapma değerleri de hesaplanarak yorumlanmıştır.

Bulgular

Yakınsama Hatası Olan Model Sayısı

Yakınsama hatası (*convergence error*) en iyi model tahmini yapmaya programlanmış istatistik paket programlarında herhangi bir model tahmin edilememesi so-

Tablo 4. Örneklem Büyüklüğü, Madde Tepki Olasılıkları ve Ölçüm Alınan Zaman Sayısına Göre Model Seçim Değerleri

Madde Tepki Olasılıkları		Güçlü						Zayıf					
Ölçüm Alınan Zaman Sayısı		2		3		4		2		3		4	
Örneklem Büyüklüğü	Bilgi Kriterleri	Doğru Karar	Yanlış Pozitif	Doğru Karar	Yanlış Pozitif	Doğru Karar	Yanlış Pozitif	Doğru Karar	Yanlış Pozitif	Doğru Karar	Yanlış Pozitif	Doğru Karar	Yanlış Pozitif
200	BIC	100	0	100	0	100	0	100	0	100	0	100	0
	AIC	88	12	87	13	84	16	88	12	86	14	92	8
	AIC3	98	2	99	1	100	0	99	1	100	0	99	1
	CAIC	100	0	100	0	100	0	100	0	100	0	100	0
600	BIC	100	0	100	0	100	0	100	0	100	0	100	0
	AIC	87	13	93	7	87	13	89	11	85	15	89	11
	AIC3	100	0	100	0	99	1	100	0	99	1	100	0
	CAIC	100	0	100	0	100	0	100	0	100	0	100	0
2000	BIC	100	0	100	0	100	0	100	0	100	0	100	0
	AIC	90	10	81	19	74	26	84	16	94	6	78	22
	AIC3	100	0	99	1	98	2	100	0	99	1	99	1
	CAIC	100	0	100	0	100	0	100	0	100	0	100	0

nucu karşılaşılan hata kodu olarak tanımlanabilir. Yakınsama hatası gerçekte olduğundan fazla statü tahmin edilmeye çalışıldığında ya da iterasyona başlangıç değerleri uygun olmadığında karşılaşılan sorunlardan biridir. Bu araştırmada da öncelikle yakınsama hatası olan model sayısı incelenmiştir. Beklendiği üzere bir ve iki statülü modellerde herhangi bir hata ile karşılaşılması. Üç statülü modellerde ortalama %16 ($S = 17.12$) yakınsama hatası tespit edilirken bu oran dört statülü modeller tahmin edildiğinde %33.5'e ($S = 10.76$) yükselmiştir.

Madde Tepki Olasılıklarının Gücünün Model Seçimine Etkisi

Madde tepki olasılıklarının gücünün model seçimine etkisi incelendiğinde, Tablo 3'te görüldüğü üzere, olasılıkların hem güçlü hem de zayıf olduğu durumlarda örneklem büyüklüğünden bağımsız olarak BIC ve CAIC bilgi kriterleri yüzde 100 doğru karar oranı sergilemiştir. AIC bilgi kriteri kullanıldığında, madde tepki olasılığının güçlü olduğu durumlarda en yüksek yanlış pozitif karar oranı %26 iken, madde tepki olasılığının zayıf olduğu koşullarda bu oran %22'de kalmıştır. AIC3 bilgi kriteri kullanıldığında ise, madde tepki olasılığının güçlü ve zayıf olduğu bütün koşullarda doğru karar oranları %98 ve üzerindedir.

Ölçüm Alınan Zaman Sayısının Model Seçimine Etkisi

Ölçüm alınan zaman sayıları dikkate alındığında (bkz. Tablo 3), BIC ve CAIC kullanılarak yapılan model seçimlerinde, üç farklı zaman koşulunda da örneklem büyüklüğünden bağımsız olarak doğru karar oranları %100'dür. Ölçüm alınan zaman sayısı iki olduğunda, AIC kullanılarak yapılan model seçimlerinde en yüksek doğru karar oranı %90'dır. AIC3 bilgi kriteri kullanıldığında ise, en düşük doğru karar oranı .98'dir. Ölçüm alınan zaman sayısı üç olduğunda, AIC kullanılarak yapılan model seçimlerinde doğru karar oranı %81'e düşmüştür. Fakat, AIC3 kullanıldığında bu oran en düşük .99 olarak ölçülmüştür. Ölçüm alınan zaman sayısı dört olduğunda, AIC bilgi kriteri kullanıldığında yanlış pozitif karar oranı %74'e kadar yükselmektedir. AIC3 için en yüksek yanlış pozitif karar oranı %2'dir.

Örneklem Büyüklüğünün Model Seçimine Etkisi

Madde tepki olasılıklarının güçlü olduğu durumlarda örneklem büyüklüğü 200'den 2000'e çıktığında BIC ve CAIC bilgi kriterleri tüm koşullardan bağımsız olarak %100 doğru tahmin sayısı verirken, AIC bilgi kriterinin doğru karar oranı .78 ile .94 arasında kalmıştır. AIC3 bilgi kriteri için örneklem büyüklüğünde doğru karar oranının .98'e düşmüş olduğu görülmektedir. yine de %90'ın üzerinde olduğu görülmektedir.

Tablo 5. Parametre Tahmin Yanlılığı

Madde Tepki Olasılıkları	Örneklem Büyüklüğü	Ölçüm Alınan Zaman Sayısı	Örtük Statü Olasılıkları		Örtük Geçiş Olasılıkları		Madde Tepki Olasılıkları	
			Ort.	S	Ort.	S	Ort.	S
Güçlü	200	2,3,4	.52	3.56	.13	5.28	.21	4.37
Zayıf		2,3,4	.05	5.84	.79	15.95	.17	2.37
Güçlü	600	2,3,4	.35	2.09	.49	3.20	.20	2.38
Zayıf		2,3,4	.24	3.63	.97	1.73	.16	1.03
Güçlü	2000	2,3,4	.14	1.09	.27	1.56	.13	1.36
Zayıf		2,3,4	.16	1.95	.19	3.61	.17	.57

Parametre Tahmin Yanlılıkları

Çalışmada parametre tahmin yanlılığı aşağıdaki eşitlikten (Muthen ve Muthen, 2002) yararlanılarak hesaplanmıştır:

$$\text{Yanlılık} = (PT - PP) / PP * 100 \quad (2)$$

PT = Tekrarlamalar (replication) sonucu elde edilen parametre tahminlerinin ortalaması

PP = Önceden belirlenmiş popülasyon parametre değeri

Bu eşitliğe göre, 200, 600 ve 2000 olmak üzere üç farklı örneklem için madde tepki olasılıkları ve örneklem büyüklüğüne göre örtük statü, örtük geçiş ve madde tepki olasılıklarının parametre yanlılık değerleri hesaplanmıştır (bkz. Tablo 5).

Bulgulara göre, madde tepki olasılıkları güçlü olduğu durumda örneklem büyüklüğü 200'den 600'e çıktığında parametre yanlılık değerlerinin yükseldiği görülmektedir. Ancak örneklem büyüklüğü 600'den 2000'e çıktığında parametre yanlılık değerlerinin tekrar düştüğü gözlenmiştir. Madde tepki olasılıklarının zayıf olduğu koşullarda ise örneklem büyüklüğü 200'den 600'e yükseldiğinde örtük statü ve örtük geçiş olasılıkları için parametre yanlılık değerlerinin yükseldiği rapor edilirken, madde tepki olasılıkları için bu değer düşmüştür. Örneklem 600'den 2000'e yükseldiğinde tüm olasılıklarda parametre yanlılık değerlerinde düşüş görülmektedir.

Tartışma

Zamana bağlı değişim sosyal bilim çalışmalarında önemli bir yer tutmaktadır. Boylamsal çalışmalarda örtük değişken modelleri göz önünde bulundurulduğunda, normal dağılım varsayımı olmayan, kategorik örtük ve gözlenen değişkenlerle baş etme kapasitesi yüksek olan

örtük Markov modelleri iyi bir alternatif olarak karşımıza çıkmaktadır. Gelişen teknoloji ve güncellenen yazılımlarla popülaritesi artan örtük Markov modelleri ile yapılan çalışmalar ise henüz sınırlı sayıdadır. Özellikle model seçimi konusunda hangi kriterlerin kullanılması gerektiği hala cevaplanması gereken bir sorudur (Bartolucci ve ark., 2013).

Bu çalışmada öncelikle tek gözlenen değişken ve beş zaman dilimi olan bir görgül veri üzerinden araştırmacılara modelin tanıtılması hedeflenmiştir. Tek gözlenen değişken ile bir örtük değişken modeli test edilmesi ancak farklı zaman dilimlerinde elde edilmiş veri setleri ile mümkün olmaktadır. Trafikteki hız davranışları örneği kullanılarak yapılan analizlerde iki örtük statülü zaman-homojen ve zaman-heterojen olmak üzere iki model tahmin edilmiştir. BIC ve CAIC bilgi kriterleri zaman-homojen modeline işaret ederken, AIC ve AIC3 bilgi kriterleri zaman-heterojen modelini daha iyi model olarak göstermiştir. Araştırmacılar zaman-heterojen modelini yorumlamayı tercih etmiştir. Özellikle farklı bilgi kriterlerinin farklı modelleri işaret ettiği durumlarda her iki modele ait parametreleri incelemek faydalı olacaktır. Bu çalışmada da AIC3 genel olarak daha tutarlı sonuçlar veren bilgi kriteri olarak öne çıktığı için zaman-heterojen modeli tercih edilmiştir. Bu modelin seçilmesindeki bir diğer motivasyon ise zaman-heterojen modelin yıllar içindeki geçiş olasılıkları arasındaki farkı görmemize olanak sağlamasıdır. Geçiş olasılıkları incelendiğinde özellikle 2014'ün ikinci yarısı ile 2015'in ilk yarısında hız ihlali yapmayan gruptan, hız ihlali yapan gruba daha önceki yıllardan daha yüksek bir geçiş olduğu görülmektedir. Bu durum kontrollerin sıklaşması ile açıklanabilir. Genel olarak bakıldığında ise hız ihlali yapan grubun bu davranışı kronik olarak devam ettirdiği gözlenmektedir.

Bu bulguya paralel şekilde, trafik çalışmalarının bir kısmı hız davranışını planlı davranış kuramı ile açıklamakta ve kişilerin tutum düzeyinde hız ihlalinin yanlış olduğunu bilse bile örneğin sadece zevk aldığı için, bu davranışı devam ettirebileceği yönünde açıklanmaktadır (Pelsmaeker ve Janssens, 2007; Warner ve Aberg, 2008).

Çalışma kapsamında model seçimine ilişkin bir simülasyon çalışması sunulmuştur. Tüm diğer örtük değişken modellerinde olduğu gibi örtük Markov modellerde de en iyi modelin hangisi olduğuna karar verme yorumlama aşamasından önce aşılması gereken bir basamaktır. Collins ve Wugalter (1992) yaptıkları simülasyon çalışmasında bu çalışmada da kullanılan EM (*expectation maximization*) algoritması kullanıldığında örtük Markov modeline ait parametrelerin tahmin edilme gücünü araştırmışlardır. Bu kapsamda iki zaman dilimi ve üç örtük statü ile sınırlandırdıkları çalışmada geçiş olasılıklarını (doğrusal ve doğrusal olmayan), madde tepki güçlerini (zayıf ve güçlü), madde sayısını (3 madde ve 6 madde) ve örneklem büyüklüğünü (300 ve 1000) değiştirmiş ve farklı koşullardaki parametre tahmin yanlışlıklarını raporlamışlardır. Bu çalışma ile benzer şekilde, madde tepki güçleri güçlü olduğunda ve örneklem büyüdüğünde parametre kestirimleri daha doğru sonuçlar vermektedir. Model seçimine ilişkin simülasyon çalışmaları incelendiğinde ise örtük sınıf modellerinde model seçimi ile ilgili daha fazla çalışma bulunmasına rağmen örtük Markov modeli kullanıldığında model seçimine ilişkin az sayıda çalışmaya rastlanmıştır. Örtük Markov modellerin psikoloji verilerinde kullanılmasına ilişkin yaptıkları simülasyon çalışmasında ise Visser ve arkadaşları (2002) BIC, AIC, A-BIC ve A-AIC bilgi kriterlerinin model uyum performanslarını karşılaştırmışlar ve A-AIC ve AIC'nin doğru modelden daha fazla statü sayısına sahip modellere uyum gösterme eğilimi varken, BIC'nin de daha az statü sayısına sahip modellere uyum gösterme eğilimi olduğunu tespit etmişlerdir. A-BIC'nin ise doğru modele en yakın sonuçları verdiğini raporlamışlardır. Bartolucci ve arkadaşları (2013) da yakın zamanda yaptıkları simülasyon çalışmasında AIC ve BIC bilgi kriterlerinin sıklıkla kullanılmasına rağmen ÖM modellerindeki yerinin yeterince çalışılmadığını vurguladıkları kitapta bu iki bilgi kriterini karşılaştırmış ve bu çalışmada da BIC daha iyi sonuçlar vermiştir.

Bu çalışmanın bilgileri ışığında örtük Markov modeli kullanmayı planlayan araştırmacıların gösterge değişkenlerini seçerken titiz davranmaları ve madde tepki olasılığı .70 ve altına düştüğünde o maddeye ait yorumları dikkatli şekilde yapmaları önerilmektedir. Bulgular modelin farklı örneklem büyüklüklerinde benzer sonuçlar elde edildiğini göstermiştir. Optimum örneklem büyüklüğü güç analizi çalışmalarıyla elde edilebilecekken, bu çalışma 200 örneklem büyüklüğünde de madde tepki

olasılıklarının güçlü olması koşuluyla doğru sonuçlar elde edilebileceğini göstermektedir.

Boylamsal çalışmalarda ölçüm alınması gereken zaman sayısı bir diğer değişken olarak karşımıza çıkmaktadır. Bu çalışmada 2, 3 ve 4. zaman diliminde alınan ölçüm sonuçlarının tamamının uygun madde tepki olasılığı ve örneklem genişliğinde doğru sonuçlar verdiği bulgulanmıştır. Sonuç olarak, örtük Markov modeli bu çalışmada değişimlenen koşullarda güçlü bir analiz yöntemi olarak kendini göstermektedir.

Bu çalışmanın simülasyon aşaması ele aldığı değişkenlerle sınırlıdır. Örneğin, bu çalışmada örtük statü olasılıkları .50 olarak belirlenmiştir. Farklı örtük statü büyüklüklerinin model seçimi üzerine etkisini görmek mümkün olamamaktadır. Benzer şekilde, zaman-heterojen geçiş olasılıklarının simüle edildiği koşullarda ölçüm alınan zaman sayısı parametre tahminleri üzerinde bir takım etkiler görülmesine neden olabilir. Bununla birlikte, örneğin birden fazla grubun yer aldığı çoklu-grup örtük Markov modelleri, zaman-heterojen modellerin değişimlendiği simülasyon çalışmaları, boylamsal eş değerliğin test edildiği simülasyonlar hala üzerinde çalışılması gereken konulardan bazılarıdır. Araştırmacının bir diğer sınırlılığı ise tek bir paket program örneğine yer vermesidir. Bu kapsamda Latent GOLD programı kullanılarak analizler yapılmış ve ilgili sentaks ekte sunulmuştur. Ancak araştırmacılar özellikle son dönemde yaygınlaşan ve özgür yazılım olması nedeniyle oldukça yaygın kullanılmaya başlanan R programının örtük Markov eklentisi ile de aynı analizleri yapma olanağına sahiptir. Tek bir program üzerinden model tanıtılması bu çalışmanın sınırlılıklarından biridir.

Kaynaklar

- Bartolucci, F., Farcomeni, A. ve Pennoni, F. (2013). *Latent Markov models for longitudinal data*. Boca Raton, Florida: Chapman & Hall/CRC.
- Bartolucci, F. ve Pandolfi, S. (2016). *LMest package*. (01.06.2017 tarihinde tarihinde <https://cran.r-project.org/web/packages/LMest/index.html> adresinden alınmıştır.)
- Bauer, D. J. ve Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods*, 8, 338–363.
- Cain, A. S., Epler, A. J., Steinley, D. ve Sher, K. J. (2010). Stability and change in patterns of concerns related to eating, weight, and shape in young adult women: A latent transition analysis. *Journal of Abnormal Psychology*, 2, 255–267.
- Castellini, G., Fioravanti, G., Lo Sauro, C., Rotella, F., Lelli, L., Ventura, L., . . . Ricca, V. (2013). Latent profile and latent transition analyses of eating disorder phenotypes in a clinical sample: A 6-year follow-up study. *Psychiatry Research*, 92–99.
- Catts, H. W., Compton, D., Tomblin, J. B. ve Bridges, M. S. (2012). Prevalence and nature of late-emerging poor readers. *Journal of Educational Psychology*, 104(1), 166–181.
- Collins, L. M. ve Flaherty, B. P. (2002). Latent class models for longitudinal data. Jacques A. Hagenaars ve Allan L. McCutcheon (Ed.) *Applied Latent Class Analysis* içinde (287–303). U.K.: Cambridge University Press.
- Collins, L. M. ve Lanza, S. T. (2010). *Latent class and latent transition analysis with applications in the social, behavioral, and health sciences*. Hoboken, NJ: Wiley.
- Collins, L. M. ve Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research*, 27(1), 131–157.
- Cosden, M., Larsen, J., Donahue, M. T. ve Nylund-Gibson, K. (2015). Trauma symptoms for men and women in substance abuse treatment: A latent transition analysis. *Journal of Substance Abuse Treatment*, 50, 18–25.
- Dias, J. G. (2006). Latent class analysis and model selection. M. R. Kruse, C. Borgelt, A. Nürberger ve W. Gaul (Ed.). *From data and information analysis to knowledge engineering* içinde (95–102). Berlin: Springer-Verlag.
- Dural, S., Somer, O., Korkmaz, M., Can, S. ve Ogretmen, T. (2010). Zamana bağlı değişimin incelenmesi: Örtük gelişme modelleri. *Türk Psikoloji Dergisi*, 25(65), 1–13.
- Gudicha, D. W., Schmittman, V. D. ve Vermunt, J. K. (2015). Power computation for likelihood ratio tests for the transition parameters in latent Markov models. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–12.
- Guo, B., Aveyard, P., Fielding, A. ve Sutton, S. (2009). Using latent class and latent transition analysis to examine the transtheoretical model staging algorithm and sequential stage transition in adolescent smoking. *Substance Use & Misuse*, 2028–2042.
- Güngör, D., Korkmaz, M. ve Sazak, H. S. (2015). Örtük sınıf analizi ile yapılan ölçme eş değeri çalışmaları model seçimi. *Hacettepe Eğitim Fakültesi Dergisi*, 30(1), 90–105.
- La Flair, L. N., Reboussin, B. A., Storr, C. L., Letourneau, E., Green, K. M., Mojtabai, R., . . . Crum, R. M. (2013). Childhood abuse and neglect and transitions in stages of alcohol involvement among women: A latent transition analysis approach. *Drug and Alcohol Dependence*, 132, 491–498.
- Lanza, S. T. ve Bray, B. C. (2010). Transitions in drug use among high-risk women: An application of latent class and latent transition analysis. *Advances and Applications in Statistical Sciences*, 3(2), 203–235.
- Lanza, S. T. ve Collins, L. M. (2008). A new SAS procedure for latent transition analysis: Transition in dating and sexual behavior. *Developmental Psychology*, 44, 446–456.
- Lin, T. H. (2006). A comparison of model selection indices for nested latent class models. *Monte Carlo Methods and Applications*, 12(3-4), 239–259.
- Marcoulides, G. A., Gottfried, A. E., Gottfried, A. W. ve Oliver, P. H. (2008). A latent transition analysis of academic intrinsic motivation from childhood through adolescence. *Educational Research and Evaluation: An International Journal on Theory and Practice*, 14(5), 411–427.
- Muthen, B. ve Muthen, L. (1998-2007). M plus user's guide fifth edition. Los Angeles: Muthen ve Muthen.
- Muthen, L. K. ve Muthen, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9, 599–620.
- Nylund, K. L., Asparouhov, T. ve Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14(4), 535–569.
- Pelsmacker, P. D. ve Janssens, W. (2007). The effect of norms, attitudes and habits on speeding behavior: Scale development and model building and estimation. *Accident Analysis and Prevention*, (39) 6–15.

- Schmittman, V. D., Dolan, C. V., van der Maas, H. L. J. ve Neale, M. C. (2005). Discrete latent Markov models for normally distributed response data. *Multivariate Behavioral Research*, 40(4), 461–488.
- Seaton, E. K., Yip, T., Morgan-Lopez, A. ve Sellers, R. M. (2012). Racial discrimination and racial socialization as predictors of African American adolescents' racial identity development using latent transition analysis. *Developmental Psychology*, 48(2), 448–458.
- Vermunt, J. K., Langeheine, R. ve Bockenholt, U. (1999). Discrete-time discrete-state latent Markov models with time-constant and time-varying covariates. *Journal of Educational and Behavioral Statistics*, 24, 179–207.
- Vermunt, J. K. ve Magidson, J. (2013). LG-Syntax User's Guide: Manual for Latent GOLD 5.0 Syntax Module. Belmont, MA: Statistical Innovations Inc.
- Vermunt, J. K. ve Magidson, J. (2015). Technical Guide for Latent GOLD 5.0: Basic, Advanced, and Syntax. Belmont, MA: Statistical Innovations Inc.
- Vermunt, J. K. ve Magidson, J. (2016). Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax. Belmont, MA: Statistical Innovations Inc.
- Vermunt, J. K., Tran, B. ve Magidson, J. (2008). Latent class models in longitudinal research, *Handbook of longitudinal research: Design, measurement, and analysis*. Menard, S. (Ed.) içinde (s. 373-385). Burlington: Elsevier.
- Visser, I., Rajimakers, E. J. ve Molenaar, P. C. M. (2002). Fitting hidden Markov models to psychological data. *Scientific Programming*, 10, 185–199.
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion and the Bayesian information criterion. *Psychological Methods*, 17(2), 228–243.
- Warner, H. W. ve Aberg, L. (2008). Drivers' beliefs about exceeding the speed limits, *Transportation Research, Part F*, 376–389.
- Yang, C. ve Yang, C. (2007). Separating latent classes by information criteria. *Journal of Classification*, 24, 183–203.

Ek. Tek gözlenen değişkenli örtük Markov modele ait Latent GOLD sentaksı

```
.....
variables
caseidpersonelno;
dependent hız;
latent
Statedynamicnominal 2;
equations
State[=0] <- 1;
State<- (b1~tra) 1 | State[-1];
hız<- 1 + State;
```

Sentaksta “variables” komutunun altında araştırmanın değişkenleri tanımlanmaktadır.

“caseid” komutu katılımcıların etiketlendiği değişken sütunun tanımlandığı alt başlıktır.

“dependent” komutu gösterge değişkenlerin sıralandığı satırdır.

“latent” örtük değişkenlerin yazıldığı bölümdür ve State komutundan sonra örtük değişkenin ölçekleme seviyesi belirtilmektedir. Bu örnekte “nominal” yani sınıflayıcı olarak tanımlanmış ve iki kategori olduğu belirtilmiştir.

“Equations” komutunun altında ise tahminlenmesi istenen parametrelere ait eşitlikler yazılmaktadır. Bu örnekte,

Örtük statü olasılıkları için: “State[=0] <- 1”

Örtük geçiş olasılıkları için: “State<- (b1~tra) 1 | State[-1]”

Madde tepki olasılıkları için: “hız <- 1 + State” eşitlikleri kullanılmıştır.

State<- (b1~tra) 1 | State[-1] komutu, State<- (b1~tra) 1 | State[-1] + (b2~tra) time | State[-1];

şeklinde değiştirildiğinde zaman-heterojen model tahminlenmektedir.

Summary

A Monte Carlo Simulation Study on Model Selection in Latent Markov Models

Duygu Güngör

Dokuz Eylül University

Selva Ülbe

Dokuz Eylül University

Samet Baş

Dokuz Eylül University

Time-dependent change has always been the subject of many psychological researches. In recent years, latent growth models, which are part of structural equation models, have been used for investigating time-dependent change in the case of continuous latent variable(s). However, it is not always possible to make continuous measurements in psychological research. For instance, in the study of driver behaviors, it may not be possible for a researcher who wants to rate risky behaviors to decide which of behaviors of crossing the red light or exceeding the speed limit is riskier. As in this example when observed and latent variable(s) are discrete, the latent Markov models which are also known as latent transition models, are used as an alternative for the longitudinal psychology studies. These models were used in the applied research such as substance abuse (Couden, Larsen, Donahue, & Nylund-Gibson, 2015; Guo, Aveyard, Fielding, & Sutton, 2009; La Flair et al., 2013; Lanza & Bray, 2010), eating behaviors (Cain, Epler, Steinley, & Sher, 2010; Castellini et al., 2013) and so on. Although in recent years the latent Markov models have been commonly used in the applied field by means of the development of software such as Latent Gold (Vermunt & Magidson, 2013), Mplus (Muthen & Muthen, 2013), PROC LTA (Lanza & Collins, 2008) and R (Bartolucci & Pandolfi, 2016), which statistics should be used for model selection is still unclear. In this context, the first aim of the study is to present an example of model application by using an empirical dataset with a single variable. The second objective is to examine the effects of the strength of item response probabilities, the number of times the measurement being taken and sample size on model selection by means of the dataset generated by Monte Carlo simulation method.

Latent Markov models consist of two parts that are the measurement model and the structural model. In the measurement model, the relations of observed variables to the latent variables are determined. The struc-

tural model provides the transition probabilities among time-dependent latent states. In this regard, the basic parameters, which are item response probabilities, latent states probabilities, and transition probabilities, are estimated (Collins & Lanza, 2010; Collins & Flaherty, 2012; Lanza & Bray, 2010; Vermunt, Tran & Magidson, 2008). The parameter called latent class probability in the latent class analysis is usually known as latent states probability in the latent Markov models in order to emphasize the dynamic structure of class (Collins & Lanza, 2010). Considering a hypothetical study of driver behavior, it is possible to have two latent states which can be defined as careful driving and risky driving. In such a case, if the probability of being in a state of careful driving in the first time period is determined to be, such as .70, the probability of being in the second latent state can be calculated as $1.00 - .70 = .30$ since the sum of latent state probabilities for each time period is equal to 1. Item response probabilities are similar to factor loadings in factor analysis and can be interpreted as in the latent class analysis. The part where the transition probabilities are calculated constitutes the Markov part of model. In this regard, transition probabilities are the conditional probabilities that individuals or observations in the first latent state at time $t-1$ will move to the second latent state at time t .

The latent Markov model has two basic assumptions; local independence and first order Markov assumptions. The assumption of local independence is the assumption that the variables observed at time t are only related to the current time. The first order Markov assumption is that statuses at time t are influenced only by time $t-1$ (Bartolucci, Farcomeni, & Pennoni, 2013; Vermunt, Langeheine, & Bockenholt, 1999).

Model Selection

In the model selection process, firstly the number of latent state is determined separately for each time period. If the number of latent state is known in advance,

Table 1. Model Selection Values According to Sample Size, Item Response Probability and Number of Measurement Time

Item Response Probability		Strong						Weak					
Number of Measurement Time		2		3		4		2		3		4	
Sample Size	Information Criterion	True	False Positive	True	False Positive	True	False Positive	True	False Positive	True	False Positive	True	False Positive
200	BIC	100	0	100	0	100	0	100	0	100	0	100	0
	AIC	88	12	87	13	84	16	88	12	86	14	92	8
	AIC3	98	2	99	1	100	0	99	1	100	0	99	1
	CAIC	100	0	100	0	100	0	100	0	100	0	100	0
600	BIC	100	0	100	0	100	0	100	0	100	0	100	0
	AIC	87	13	93	7	87	13	89	11	85	15	89	11
	AIC3	100	0	100	0	99	1	100	0	99	1	100	0
	CAIC	100	0	100	0	100	0	100	0	100	0	100	0
2000	BIC	100	0	100	0	100	0	100	0	100	0	100	0
	AIC	90	10	81	19	74	26	84	16	94	6	78	22
	AIC3	100	0	99	1	98	2	100	0	99	1	99	1
	CAIC	100	0	100	0	100	0	100	0	100	0	100	0

this step can be skipped by adopting a confirmatory approach (Bartolucci, Farcomeni, & Pennoni, 2013). Then conditional probabilities and latent probabilities are determined by restricting the number of latent states. By examining these models, if the restrictions on the transition and conditional probabilities can be imposed, the nested models are tested by putting various restrictions. One of the methods that is used to select the best model among these nested models is to decide it based on L² statistics. Alternatively, the use of the AIC, BIC information criteria for model selection is generally accepted.

Empirical Example

The sample of this example is composed of 511 male drivers working in a private transportation company. With the permission of the company, the speeding records of the drivers in the first half of 2013, 2014, and 2015 were accessed. A five-time dataset was obtained by dividing the years into two semi-periods. When the frequency ratios of speed limit violations over years were examined, it was observed that they varied between 4% and 8%.

The dataset with a single variable taken in five-time periods was analyzed via the Latent GOLD 5.1 Syntax version. Parameters of model with two latent states were estimated. Model-1 was tested as a model in which there were two latent states and transition probabilities were

time-homogenous while Model-2 was analyzed as a model in which two latent states and but transition probabilities were time-heterogeneous. While the BIC and CAIC information criteria pointed to the time-homogeneous model, the AIC, AIC3 information criteria pointed to the time-heterogeneous model. The heterogeneous model parameters are interpreted based on the knowledge that the AIC3 information criterion gives more results that are consistent in latent class analysis (Güngör, Korkmaz, & Sazak, 2015). When item response probabilities are examined, the probability of those in the first latent state to violate speed limit was found to be .01 and the probability of those in the second latent status was found to be .66. Based on these probabilities, the first latent state was called as those complying with speed limit while the second latent state referred to those violating speed limit. In the first half of 2013, 98% of the drivers were in the latent state of those who comply with speed limit. In the second half of 2014, this probability declined to .96 and in the first half of 2015 dropped to .93. In other words, in the first half of 2015, the proportion of those in the latent state of violation of speed limit moved up to 7%.

Method

The datasets for 18 different conditions were generated by using Monte Carlo simulation section of La-

Table 2. Parameter Estimation Bias

Item Response Probability	Sample Size	Number of Measurement Time	Latent State Probability		Latent Transition Probability		Item Response Probability	
			Mean	<i>SD</i>	Mean	<i>SD</i>	Mean	<i>SD</i>
Strong	200	2,3,4	.52	3.56	.13	5.28	.21	4.37
Weak		2,3,4	.05	5.84	.79	15.95	.17	2.37
Strong	600	2,3,4	.35	2.09	.49	3.20	.20	2.38
Weak		2,3,4	.24	3.63	.97	1.73	.16	1.03
Strong	2000	2,3,4	.14	1.09	.27	1.56	.13	1.36
Weak		2,3,4	.16	1.95	.19	3.61	.17	.57

tent GOLD 5.0 statistical package program (Vermunt & Magidson, 2015). The study consisted of four independent variables which are the sample size of manipulated conditions (i.e., 200, 600, and 2000), the strength of item-response probabilities (i.e., .1, .9, .3, and .7), the number of measurement occasions (i.e., 2, 3, and 4), and the information criteria used in model selection (BIC, CAIC, AIC3, AIC). The information criterion used in the model selection was independent variable. In this context, 1800 datasets generated by 18 different conditions and 100 different replications for each condition were produced and analyzed. Furthermore, fixed conditions in the research were the number of the latent state (2), the latent state probabilities (i.e., .50), the latent transition probabilities (i.e., .80 and .20), the number of items (i.e., 5), the number of item category (i.e., 2). The latent transition probabilities were time-homogeneous.

As a result of the analyzes, the information criteria were calculated by using the values of L^2 obtained for each condition. Based on these information criteria, the models were selected and reported. Since the number of the latent states was fixed to 2, the proportion of deviation from this value was calculated as a percentage. In addition, the means and the standard deviations of the parameters of each condition were calculated and interpreted.

Results

Number of models with convergence error

Convergence errors were only detected for models with three and four states. In the case of models for three states, an average of 16% convergence error ($SD = 17.12$) was found, however for models with four status the rate was 33.5% ($SD = 10.76$).

The effect of strength of item response probabilities on model selection

As indicated in the Table 1, BIC and CAIC information criteria exhibited 100% true positive results, independent of the sample size, in the case of probabilities were both strong and weak. When the AIC3 information criterion was used, the correct decision ratios are 98% and above in all conditions where the item response probability was strong or weak.

The effect of number of measurement time on model selection

In model selection using BIC and CAIC, the true positive ratios were 100% regardless of the sample size in 3 different time conditions. When the number of measurement time was 2, in model selection by using AIC the highest true positive ratio was 90%. All other ratios for 3 and 4 time of measurement were indicated in Table 1.

The effect of sample size on model selection

In the case item response probabilities were strong and sample size increased from 200 to 2000, true positive ratios of the AIC information criterion were between .78 and .94 while the BIC and CAIC information criteria shows 100% accurate estimates, independent of all conditions.

Parameter estimation bias

The values of parameter estimation bias were estimated by the following equation (Muthén & Muthén, 2002):

$$\text{Bias} = (\text{PE} - \text{PP}) / \text{PP} * 100 \quad (1)$$

PE = Average value of parameter estimates as a result of replications

PP = Pre-defined population parameter value

According to this equation, parameter bias values of latent status, latent transition, and item response probabilities were calculated in reference to the item response probabilities and sample size for three different samples as 200, 600 and 2000 (see Table 2).

Discussion

In this study, which was designed as two parts, firstly, it was aimed to introduce the model through the empirical data with only one observed variable and five time occasions. In the first study, the models with two latent states which were time-homogeneous and time-heterogeneous were estimated by using the example of speed behavior in traffic. The BIC and CAIC information criteria pointed to the time-homogeneous model, while the AIC and AIC3 information criteria indicated the time-heterogeneous model as a better model. Since AIC3 is generally considered as an information criterion giving more consistent results (Güngör, Korkmaz, Sazak, 2015), in the study the time-heterogeneous model was preferred. In addition, time-heterogeneous model allows seeing the differences among the transition probabilities over the years. The transition from the group who did not violate the speed limit in the second half of 2014 and the first half of 2015 to the group who violated the speed limit had been found to have higher probabilities compared to previous years. Furthermore, it was observed that group who violated the speed limit continued to display the same behavior in the following years.

In the second study, a simulation study on model selection was presented. The findings of the simulation study indicated that when item response probabilities were strong and the sample size was large, the parameter estimates gave results that were more accurate. Therefore, we recommend the researchers planning to use the latent Markov model to be careful in choosing indicator variables and to make comments on that item when item response probabilities were .70 and less.

As for the limitations of the study, the simulation study was limited to the variables that addressed in this research. Furthermore, the simulations involving multi-group latent Markov models, simulations of time-heterogeneous models and longitudinal equivalence studies are still some of the issues to be investigated. Another limitation of the study was to introduce a model through an example of a single package program, Latent GOLD program. Researchers can conduct the same analysis with the latent Markov extension of R program, which has become an extensively used as a free software program.