T.C.
MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI
İSTATİSTİK YÜKSEK LİSANS TEZİ

DİKEY KESİT VERİLERİNİN İNCELENMESİNDE ÇOK DÜZEYLİ ANALİZ YÖNTEMLERİ VE BİR UYGULAMA

Özlem DENİZ

DANIŞMAN : Prof.Dr. Gülay KIROĞLU

İSTANBUL – HAZİRAN 2005

İÇİNDEKİLER

ÖZET

Dikey kesit verilerin analizi, incelenen verilerde değişimin ölçülmesine ve buna bağlı olarak daha anlamlı yorumlamalara olanak sağlaması nedeniyle günümüzde sıkça kullanılan bir yöntem olmuştur. Bu analizlerde amaç, aynı bireylere farklı zaman dilimlerinde, aynı ölçümlerin uygulanması sonucunda elde edilen bilgilerin değerlendirilmesidir.

Analizlerde ilgilenilen sonuca ulaşmak için incelenen değişkenlerin, farklı başka değişkenlerden de etkilenip etkilenmediğinin belirlenmesi için verileri inceleme yöntemlerinden bir tanesi de çok düzeyli analiz olup bu yöntemlerle alt düzeylerdeki daha fazla bilgi kullanılarak sonuçlara ulaşılabilmektedir.

Bu amaca yönelik olarak hazırlanan çalışmada dikey kesit verilerin tanımlanmış, dikey kesit verilere çok düzeyli analizlerin uygulanması ve elde edilecek modellerin birbirleriyle karşılaştırılmaları için kullanılabilecek kriterler incelenmiştir.

Uygulama aşamasında Mature Eğitim ve Aile Danışmanlığı Merkezi'nden elde edilen verilere HLM paket programı kullanılarak çok düzeyli analiz uygulanmış, elde edilen modeller yorumlanarak tartışmaları verilmiştir.

Anahtar Kelimeler: Dikey kesit veriler, çok düzeyli analizler, sapma istatistiği, AIC kriteri, BIC kriteri

SUMMARY

Longitudinal data analysis is a frequently used method for calculating the variation in the researched data and providing more significant explanations. The main objective of this analysis is to evaluate the information calculated from the application of the same measurements to the same people at different time periods.

Multilevel analysis is one of the methods used to measure whether the researched variables effected from other variables or not. In this method more information in sublevels is used to calculate the results.

In this study, longitudinal data is defined and criteria for the application of multilevel analysis to the longitudinal data and comparing the obtained models are studied.

In the application phase, HLM software is used to apply longitudinal analysis to the data taken from “Mature Eğitim ve Aile Danışmanlığı Merkezi” and the obtained models are explained and discussed.

Key words: Longitudinal data, multilevel analysis, deviance statistics, AIC criteria, BIC criteria

GİRİŞ

Dikey kesit verileri, değişkenleri oluşturan gözlemlerin zamana bağlı olarak çok kez incelenmesi ile elde edilmiş veriler olarak tanımlanabilir. Dikey kesit verilerin analizleri yardımıyla bireyler hem kendi içlerinde, hem de birbirleri ile karşılaştırılarak zaman içindeki gelişimleri gözlenebilmektedir. Analizlerde ilgilenilen sonuca ulaşmak için incelenen değişkenlerin, farklı başka değişkenlerden de etkilenip etkilenmediğinin belirlenmesine yönelik olarak veriler, çok düzeyli analiz yöntemleriyle de incelenebilir.

Bu amaca yönelik olarak çalışmanın birinci bölümde dikey kesit veriler tanımlanmış, yatay kesit ve dikey kesit veriler arasındaki farklılıklar ve yaş, dönem, kuşak etkilerinden bahsedilmiştir. Bunların yanı sıra dikey kesit verilerin analizlerinin uygulanabilmesi için verilerin düzenlenmesinde kullanılabilecek bazı düzenler ve sonrasında uygulanabilecek bazı analiz teknikleri ile bu tekniklerin özellikleri gösterilmiştir.

İkinci bölümde ise dikey kesit verilerin incelenmesinden, bu verilere ait açıklayıcı istatistiklerin elde edilmesinden, analizlere başlamadan önce çalışılacak örneklem büyüklüğünün belirlenmesi ve kullanılacak testlerin gücünün hesaplanabilmesi için gerekli açıklamalar yapılmıştır. Bunun yanı sıra çalışmalarda kayıp gözlemlerle karşılaşılması durumunda izlenecek yol bu bölümde açıklanmıştır.

Üçüncü bölümde çok düzeyli veriler ile çok düzeyli analiz yönteminin tanımı, iki, üç ve daha çok düzeyli modellerin ve parametre kestirimlerinin elde edilmesi, dikey kesit verilerin söz konusu olduğu durumlarda çok düzeyli analiz yöntemlerinin uygulanabilmesi hakkında bilgi verilmiştir.

Dördüncü ve son bölümde ise, Mature Eğitim ve Aile Danışmanlığı Merkezi'nden hizmet alan 40 bebeğe ait 12, 24 ve 36 aylık dönemlerdeki dil gelişimlerini etkileyen değişkenler incelenmiştir. Bunlar dışında gelişimi etkileyen farklı değişkenlerin de modele etkilerinin görülmesi için çok düzeyli analiz uygulanmıştır. Daha sonra modelde anlamsız oldukları belirlenen değişkenler modelden çıkartılarak iç içe geçmiş modeller oluşturulmuştur. Uygulamanın son aşamasında ise sapma istatistikleri ve bilgi kriterleri kullanılarak en uygun model seçilmiş, yorumlanarak tartışmaları verilmiştir.

BİRİNCİ BÖLÜM
DİKEY KESİT VERİ ANALİZİ

Dikey kesit analizi, toplanan verilerde değişimin ölçülmesine ve buna bağlı olarak daha anlamlı yorumlamalara olanak sağlaması nedeniyle günümüzde yaygın biçimde kullanılan bir yöntem olmuştur.

Dikey kesit veriler ilk olarak, 1665 yılında Yeni Fransa'da (Kanada) başlayan, Quebec'de sürdürülen ve 1754 yılına kadar süren periyodik nüfus sayımlarında kullanılmaya başlanmıştır. Bunlar yapılmış olan ilk nüfus sayımları olmamakla birlikte bu sayımlarla nüfus verileri ilk defa belli ve sabit dönem aralıklarında, periyodik olarak toplanmaya başlamıştır.

Bireysel çalışmalar incelendiğinde, bu alandaki ilk çalışmaların Baltes ile Nesselroade (1979) ve Wall ile Williams (1970) tarafından gerçekleştirildiği görülebilir.

1.1. Dikey Kesit Verilerinin Tanımı

Dikey kesit verileri, değişkenleri oluşturan gözlemlerin en az iki kez incelenmesi ile elde edilmiş veriler olarak tanımlanabilir. Yatay kesit ve dikey kesit verilerin kullanıldığı analizler karşılaştırılırsa ve T ölçüm sayısı olarak tanımlanırsa, yatay kesit analizlerde T=1 ve dikey kesit analizlerde T>1 olacaktır.[1]

Dikey kesit araştırmalarda;

(a) Her bir birey veya değişken için iki veya daha fazla farklı dönemlerde veri toplanır.

---

[1] C. J. Catrien, L J. Van Der Kamp, "Longitudinal Data Analysis, Designs, Models and Methods", Sage Publications, California, 1998, p:1.

(b) Analiz edilen gözlemler aynıdır ve en azından herhangi bir dönemdeki durumuyla karşılaştırılabilir niteliktedir.
(c) Analiz, verinin dönemler içinde veya arasında karşılaştırılmasını içermelidir.

Dikey kesit düzenlemelerle bir değişkende dönemler arasındaki farkların veya değişimlerin ölçümü yapılır.[2]

Dikey kesit verilerin çözümünde özel istatistiksel yöntemler kullanılır. Çünkü farklı dönemler için aynı gözlemler kullanıldığından analizlerde istenmeyen bir durum olan gözlemler arası korelasyon ortaya çıkmaktadır.[3]

1.2. Yatay Kesit Veri Analizi ile Dikey Kesit Veri Analizinin Farkları

Dikey kesit veri analizi ile yatay kesit veri analizi arasındaki farklar ve benzerlikler aşağıdaki gibi özetlenebilir.[4]

(1) Dikey kesit verilerle çalışılırken veri toplamak yatay kesit analizler için veri toplamaktan çok daha zor ve maliyetlidir. Yatay kesit veri analizinde, N veri toplanacaksa, dikey kesit veri analizinde en az 2N veri toplanmalıdır. Bu durum maliyeti arttıracaktır.[5]
(2) Dikey kesit çalışmalarda verinin niteliği ve örneklemin yeterliliğinde karşılaşılabilecek sorunlar yatay kesit çalışmalarda karşılaşılan sorunlardan farklı değildir.
(3) Yatay kesit çalışmalar gelişimsel (yaş) trendi, tarihsel (dönem) trendi ve kuşak etkisini birbirinden ayıramamaktadır. Bu üç etkinin de varolduğu çalışmalarda dikey kesit veri analizinin kullanılması uygun olmaktadır.

---

[2] S. Menard, "Longitudinal Research", Sage Publications, Series: Quantitative Applications in the Social Sciences, California, 1991, p:4.
[3] P. J.Diggle, P. Heagerty, "Analysis of Longitudinal Research", Oxford University Press, Oxford, Second Edition, 2003, p:2.
[4] A. g. e., Menard, p:66.
[5] A. g. e., Catrien, p:2.

(4) Tarihsel değişimin tanımlanması ve analizi, dikey kesit verilerin kullanılmasını gerektirmektedir.

(5) Gelişimsel trendlerin tanımlanması ve analizi zamana bağlı yatay kesit çalışmalarda incelenebilir ancak elde edilecek sonuçlar dikey kesit çalışmalar yardımıyla elde edilecek sonuçlardan daha kesin ve tutarlı olmayacaktır.

(6) Dikey kesit veriler sosyal bilimlerde herhangi bir dinamik sürecin parametrelerinin kestirimlerinin etkin ve önyargısız olması için gereklidir.[6]

Dikey kesit veri analizlerinde maliyet gibi ciddi bir olumsuzluk söz konusu iken, yatay kesit veri analizi yerine tercih edilmesinin iki esas nedeni olabilir:

- Aynı gözlem için birbiri ardı sıra veri toplayarak incelenen gözlemdeki zamana bağlı değişimler daha rahat görülebilir.
- Olayların zamana göre elde edilmesi ve sıralanması “nedensellik”e (causality) ulaşmada kolaylık sağlayacaktır.[7]

Yatay kesit ve dikey kesit çalışmalar arasındaki fark basit doğrusal regresyon modelinde daha açık görülmektedir. Dikey kesit verilerin analizinde elde edilen regresyon modelinde, yatay kesit veri analizinden farklı olarak sabit katsayı bulunmamaktadır. i indisi birimleri (i=1, ..., m), j indisi gruplardaki birimleri (j=1, ..., n_i) ve p’de küme sayısını (p=1, ..., n_k) göstermek üzere, dikey kesit veriler için doğrusal regresyon modeli

$$\begin{aligned} Y_{ij} &= \beta_1 x_{ij1} + \beta_2 x_{ij2} + \ldots + \beta_p x_{ijp} + \varepsilon_{ij} \\ &= \mathbf{x}_{ij}' \boldsymbol{\beta} + \varepsilon_{ij} \end{aligned} \qquad (1.1)$$

şeklinde ifade edilebilir. Burada $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_p)$ bilinmeyen regresyon katsayılarını göstermektedir.[8]

---

[6] A. g. e., Menard, p:66.
[7] A. g. e., Catrien, p:2.

Yatay kesit veri analizinin aksine, dikey kesit veri analizi ne kadar iyi yönetilirse yönetilsin, bazı sorunlarla karşılaşılabilir. Bu sorunlar, yöntemsel nedenlerden ve/veya aynı değişkenin tekrarlı gözlemlerden oluşmasından kaynaklanabilir. Böylece istatistikteki en önemli varsayımlardan biri olan, gözlemlerin birbirinden bağımsız olması varsayımı sağlanamayabilir. Bu yüzden, gözlemlere ait ölçümlerin seri olarak bağımlı (serially dependent) -(t + 1) zamanındaki gözlemlerin, (t) zamanındaki gözlem değerlerinden kestirilebilmesi- olduğu söylenir. Seri bağımlılık durumu birçok test istatistiğinin uygulanmasını geçersiz kılmaktadır.[9]

1.3. Yaş - Dönem – Kuşak Etkileri

Dikey kesit çalışmalar bireylerin temel alınan düzeyden başlayarak (kuşak etkisi), zaman içindeki değişimlerini (yaş etkisi) incelemektedir.[10] Değişimin ölçüldüğü çalışmalarda "zaman" ve "yaş" değerleri arasındaki ayrımın kavramsal olarak açıklanmaması, sonuçların yorumlanması aşamasında ciddi problemlerin ortaya çıkmasına neden olur. Bu problemleri anlayabilmek, tarihsel ve gelişimsel değişimlerinin ölçülmesindeki tartışmaları sabit bir şekilde açıklayabilmek için analizin bir birimi olan yaş, dönem ve kuşak değişkenleri arasındaki farklılıkların açıklanması gerekir.

Yaş, dönem ve kuşak kelimeleri sırasıyla, "kaç yaşındasın?", "hangi yıldayız?" ve "kaç yılında doğdun?" sorularını cevaplamaktadır. "Kaç yaşındasın?" sorusu davranışın gelişimsel açıklamasının yapılmasını sağlamaktadır. Örneğin emeklilik daha çok 65 yaşından sonra görülmektedir. "Hangi yıldayız?" sorusu davranışın tarihsel yapısını ve tarihsel açıklamalarını göstermektedir. "Kaç yılında doğdun?" sorusu ise kişinin

[8] A. g. e., Diggle, p:16.
[9] A. g. e., Catrien, p:4.
[10] A. g. e., Diggle, p:1.

belirlenen yıllarda veya dönemlerde kaç yaşında olduğunun kestirilmesini sağlar ve kişinin gelişim açısından incelenmesini kolaylaştırır.

Gelişimsel, tarihi ve kuşak birliklerinin etkilerinin dikey kesit veriler olmadan birbirlerinden kolayca ayrılamayacağı bilinmelidir.

Yaş, dönem ve kuşak etkilerinin doğrusal bir ilişki içerisinde olduğu varsayılırsa, herhangi iki etkinin üçüncü etkiyi de kontrol ettiği söylenebilir. Bu açıklamaya bağlı olarak da bu etkileri ifade eden doğrusal fonksiyon aşağıdaki gibi yazılır.[11]

Kuşak (doğum yılı) = Dönem (yıl) – Yaş (doğumdan itibaren geçen süre)

1.4. Dikey Kesit Veriler için Önerilen Düzenler

Dikey kesit çalışmalar; çeşitli sayıda değişkenler, dalgalar ve örneklemlerden oluşurlar. Dalga sayısı, aynı örneklem içindeki birimlerin farklı zaman dilimleri içerisinde kaç defa gözlendiği şeklinde tanımlanmaktadır. Belirli örneklerin ve dalgaların seçimi ile bunların kombinasyonu dikey kesit düzenler olarak adlandırılır.[12]

Dikey kesit çalışmalarda verilerin toplanması, toplanan verilerin konuları, konulara ilişkin uygulanacak analizlerin belirlenmesi ve sonuçların yorumlanması için çeşitli düzenler oluşturulmuştur. Dikey kesit çalışmalarında çalışma düzeninin doğru seçilmesi çalışmanın daha iyi yürütülmesi ve daha kolay bir şekilde istenilen sonuçlara ulaşılması anlamına gelmektedir. Çalışmanın düzeni, yönetilen son dalgaya kadar sabittir. Düzen aşamasında yapılacak hataların düzeltilmesi oldukça zor ve masraflı olacaktır. Bu nedenle çalışmanın başında uygulanacak düzenin seçilmesi için acele edilmemeli ve

[11] A. g. e., Menard, pp:6-9.
[12] A. g. e., Catrien, p:19.

en uygun düzenin seçilmesinden sonra analize başlanmalıdır. İlk aşamada araştırmacı çalışmasındaki dalga sayısını bulmak zorundadır.

Dikey kesit veriler için önerilen düzenlerin bazıları aşağıda belirtildiği gibidir.

1.4.1. Trend çalışması

Trend çalışmasında ("tekrarlı yatay kesit çalışması" olarak da adlandırılır) iki veya daha fazla yatay kesit çalışması bir arada incelenmektedir. Trend çalışmasının şematik gösterimi tablo 1.1'de verilmiştir.

Tablo 1.1. Trend Çalışmasının Şematik Gösterimi

Yaş Grubu	Örneklem	Zaman	Değişkenler
A_1	S_1	t_1	$X_1, X_2, X_3, ..., X_M$
A_1	S_2	t_2	$X_1, X_2, X_3, ..., X_M$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_1	S_T	t_T	$X_1, X_2, X_3, ..., X_M$

Bu tarz çalışmalarda araştırmacılar bireylerin dalgalar arasındaki değişimleriyle ilgilenmemektedirler. Bu yüzden trend çalışmaları raslantısal sıralamaları veya bireylerin gelişimini incelemek için uygun değildir. Yatay kesit çalışmalarının asıl olumlu yanı bütün düzeylerdeki değişimin tek tek incelenebilmesidir. Buna bağlı olarak trend çalışmalarına, örnek birimleri düzeyleri için yatay kesit ama araştırma birimleri düzeyleri için dikey kesit çalışmaların birleşimi denilebilmektedir.

1.4.2. Zaman serileri düzeni

Zaman serileri analizinde tekrarlı gözlemler aynı örneklem kümesi içinden seçilmektedir. Bu gözlemlerin eşit zaman aralıklarında seçilmeleri şart değildir. Zaman serileri düzeninin şematik gösterimi tablo 1.2'de verilmiştir.

Tablo 1.2. Zaman Serisi Çalışmasının Şematik Gösterimi

Yaş Grubu	Örneklem	Zaman	Değişkenler
A_1	S_1	t_1	$X_1, X_2, X_3, ..., X_M$
A_2	S_1	t_2	$X_1, X_2, X_3, ..., X_M$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_T	S_1	t_T	$X_1, X_2, X_3, ..., X_M$

Zaman serileri düzeninde bireyin kendi içinde değişimi gözlenebilmektedir. Çünkü zaman içerisinde sürekli aynı bireyler gözlenmiştir. Zaman serisi düzenleri oldukça esnek ve kullanışlıdır. Sonraki maddede anlatılacak olan panel çalışmaları, birçok gözlem ve değişken içeren ve kısıtlı sayıda ölçüme sahip olan zaman serisi düzeninin çeşitlemeleri olarak açıklanacaktır.

1.4.3. Panel çalışması

Zaman serileri düzeni genellikle panel çalışmalarla karıştırılır. Uygulamada zaman serileri çalışmaları, birçok dalga için düzenlenmiş düzenler olarak karşımıza çıkar. Panel çalışmalarında ise dalga sayısı sınırlıdır.

Panel çalışmalarda, araştırmaya katılan bireylerle aynı anket formu kullanılarak arka arkaya birçok kez görüşülür. Panel çalışmalarının önemli olmasının çok önemli iki nedeni vardır. Bunlardan birincisi panel çalışmanın bilgileri, mikro düzeydeki örneklem birimlerindeki değişimi kapsamasıdır. Bazı istatistiksel modeller kullanılarak değişimin miktarı çalışmadaki diğer değişkenlerle ilişkilendirilebilmektedir. Böylece bir panel düzeni ile araştırmacılar zamanın bir noktasındaki ilişkinin yerine zaman içerisindeki tüm ilişkileri gözlemleyebilirler. İkinci neden veri toplamanın maliyeti ile ilgilidir. Beş dalgaya sahip bir panel çalışmanın maliyeti beş ayrı yatay kesit çalışmanın maliyetinden daha düşüktür.[13]

[13] T. W. Taris, “A Primer in Longitudinal Data Analysis”, Sage Publications, London, 1999, pp: 6-8.

Fazla sayıda kuşak içeren dikey kesit panel düzenlerinde her çeşit dikey kesit analizi kullanılabilir. Tek kuşak panel düzenlerinde kuşaklar arası karşılaştırma yapılamadığı için çoklu kuşak düzenleri tercih edilir ve böylece yaş, dönem ve kuşak etkileri analizleri yapılabilir, değişimsel ve tarihsel değişimleri açıklanabilir ve doğrusal panel analizleri uygulanabilir.[14]

1.4.4. Geçmişle ilgili (retrospective) çalışma

Geçmişe yönelik düzenlerde çalışmaya t_T anında başlanır ve geçmişe doğru veri toplamaya devam edilir.

Olası dikey kesit çalışmalarının (panel veya aracılık çalışmaları gibi) en önemli olumsuz yanı, zamana bağlı analizlerin en az iki dalga olması durumunda uygulanabilmesidir. Eğer elde edilen dalgalar arasında da uzun yıllar varsa analizlerin uygulanması ve sonuçların yorumlanması çok daha zor olacaktır.

Bu düzende veriler, gözlemlere geçmişle ilgili sorular sorularak toplanır. Geçmişe ait sorular içeren bir anket, gözlemin geçmişe ait bilgilerine ulaşmada en kolay yoldur.

Bu düzen uygulanırken bazı sorunlarla karşılaşılabilir. Bunların en başında bu çalışmada kullanılacak örneklemin rasgele olmamasıdır.

Bu düzenlerde iki tür yanıt hataları vardır. Bunlar “hafıza hataları” ve “açıklama hataları”dır. Açıklama hatalarını Schwarz (1990), kişinin anketi uygulayan veya verileri toplayan kişilere tam olarak bilgi vermemesi olarak açıklamıştır. Ancak açıklama hatalarına sadece geçmişle ilgili çalışmalarda değil, birçok çalışmada ve düzende de rastlanabilir.

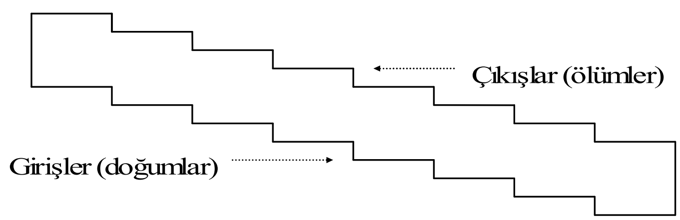
---

[14] A. g. e., Menard, p:30.

Hafıza hataları, bilginin hatırlanması durumunda karşılaşılan hatalardır. İki şekilde karşılaşılabilir: Yanıtlayıcı konuyla ilgili bilginin bir kısmını hatırlamamaktadır veya geçmişe ait toplanmış veriler değiştirilmiştir. Dikey kesit analizinde bu tarz bir düzen kullanılmak isteniyorsa, veri toplanması aşamasında bu tarz hataların yapılmamasına dikkat edilmesi gerekmektedir.[15]

1.4.5. Toplam kitle (total population) düzenleri

Toplam kitle düzenlerinde örneklem birim sayısı, çalışmanın her bir döneminde tekrar irdelenir. Çünkü bir dönemden diğer döneme geçerken bireylerin bir kısmı ölmüş veya yeni bireyler çalışmaya eklenmiş olabilir. Eğer dönemler arasında fazla süre farkı yoksa veya oldukça fazla sayıda dönem varsa bu durum göz ardı edilebilir.



Şekil 1.1. Toplam Kitle Düzeninin İşleyiş Süreci

Toplam kitle düzenlerinde kayıp gözlemlerden veya ölçüm hatalarından kaynaklanan problemlerle karşılaşılabilir. Çünkü bu düzen, dönem trendini ölçmek veya anlamak için kullanılır, ancak yaş ve kuşak etkilerini açıklayamayacak kadar da yetersizdir.[16]

[15] A. g. e., Taris, p:10.
[16] A. g. e., Menard, p:26.

1.5. Dikey Kesit Veri Kümesini Oluşturmak

Dikey kesit veri analizine başlamadan önce, veriler analize uygun hale gelecek şekilde düzenlenmelidir. Yatay kesit çalışmalarda verileri analize hazır hale getirmek oldukça kolaydır. Tek ihtiyaç olan şey her bir bireyin kendine ait kayıtlarının olduğu standart bir veri kümesidir. Dikey kesit veri analizinde ise veriler iki şekilde düzenlenir ve analize hazır hale getirilir: kişi-düzey veri kümesi ve kişi-dönem veri kümesi.

- *Kişi-düzey (person-level) veri kümesinde*, her kişinin bir kaydı ve ölçülmüş durumların verilerini içeren birden çok değişken vardır.
- *Kişi-dönem (person-period) veri kümesinde,* her kişinin her bir ölçüm durumu için birçok kaydı vardır.

1.5.1. Kişi-düzey veri kümeleri

Çok değişkenli biçim olarak da bilinen kişi-düzey veri kümelerinin yapısı, yatay kesit veri kümelerinin yapısına benzemektedir. Kişi-düzey veri kümelerinin en belirgin özelliği, toplanan verilerin dalga sayısı ne olursa olsun her bir kişiye ait sadece bir satırın (veya bir kaydın) olmasıdır. Her bir gözlemin tekrarlanmış ölçümleri ilave değişkenler olarak gösterilirler.

Kişi-düzey veri kümesinde örneklemdeki kişi sayısı kadar kayıt vardır. Eğer böyle bir çalışmaya yeni dalgalar eklenirse, çalışmaya yeni gözlem değil, yeni değişkenler eklenmiş olur. Zaman değerleri “a, b, c, d, ...”, ilgilenilen değişken “X” ve kişilerin sıra numaraları “no” ile ifade edilirse, kişi-düzey veri kümesi aşağıdaki gibi gösterilir.

Tablo 1.3. Kişi-Düzey Veri Kümesinin Gösterimi

No	X_a	X_b	X_c	X_d	...
1	X_{a1}	X_{b1}	X_{c1}	X_{d1}	...
2	X_{a2}	X_{b2}	X_{c2}	X_{d2}	...
3	X_{a3}	X_{b3}	X_{c3}	X_{d3}	...
4	X_{a4}	X_{b4}	X_{c4}	X_{d4}	...
5	X_{a5}	X_{b5}	X_{c5}	X_{d5}	...
⋮	⋮	⋮	⋮	⋮	⋮

Kişi-düzey veri kümesinin en önemli avantajı her bir kişinin deneysel (emprical) büyüme kayıtlarının daha kolay açıklanmasına ve yorumlanmasına yardımcı olmasıdır.

Kişi-düzey veri kümeleri dikey kesit çalışmalarda çok fazla tercih edilmemektedir. Bunun nedeni dört başlıkta özetlenebilir.

(1) Veriler için çok fazla açıklayıcı olmayan bilgi içerir.

(2) Açık ve belirgin bir “zaman” değişkeni içermez.

(3) Dalgaların sayısı veya genişliği gözlemler arasında değişiklik gösteriyorsa etkisiz ve kullanışsızdır.

(4) Zaman değişimlerinde ortaya çıkan değişikliklerin kestirilmesi kolay değildir.

1.5.2. Kişi-dönem veri kümeleri

Kişi-dönem veri kümelerinde (tek değişkenli biçim olarak da bilinir) her bir gözlemin her dönem için birçok kaydı vardır.

Kişi-dönem veri kümesinde, her bir kişi-dönem kombinasyonu için birçok kayıt vardır. Bu nitelikte bir çalışmaya yeni dalgalar eklenirse, çalışmaya yeni değişkenler değil, yeni kayıtlar eklenmiş olur. Aynı değerler kullanılarak kişi-dönem veri kümesi aşağıdaki gibi gösterilebilir (zaman değerleri T ile ifade edilmiştir).

Tablo 1.4. Kişi-Dönem Veri Kümesinin Gösterimi

No	T	X
1	a	X_{a1}
1	b	X_{b1}
1	c	X_{c1}
1	d	X_{d1}
⋮	⋮	⋮
2	a	X_{a2}
2	b	X_{b2}
2	c	X_{c2}
2	d	X_{d2}
⋮	⋮	⋮

Uygulamada daha çok kişi-dönem veri kümeleri kullanılmaktadır.

Tüm kişi-dönem veri kümeleri dört çeşit değişken içerirler: (1) gözlem numarasının bulunduğu değişkenler, (2) zaman göstergesi içeren değişken, (3) sonuç değişkenleri, (4) açıklayıcı değişkenler.[17]

Dikey kesit veri analizi için, ister kişi-düzey veri kümesi biçiminde, ister kişi-dönem veri kümesi biçiminde olsun hazırlanan verilerin dengede olması (aynı sayıda birim ve dalga içermesi) şart değildir. Diğer bir ifadeyle, dikey kesit veri analizinin uygulanabilmesi için her bir kişinin aynı sayıda dalgası olmak zorunda değildir.

[17] J. D. Singer, J. B. Willett, "Applied Longitudinal Data Analysis, Modeling Change and Event Occurrence", Oxford University Press, New York, 2003, pp:20-23.

1.6. Dikey Kesit Veriler için Analiz Yöntemleri

Dikey kesit verilerin düzenlenme işlemi tamamlandıktan sonra eldeki bilgiler yardımıyla uygulanacak analiz yönteminin belirlenmesi gerekmektedir.

Dikey kesit veri analizi iki aşamada incelenir. Düzey-1 olarak adlandırılan ilk aşamada zaman içerisinde, gözlemler arası değişimler hesaplanmaktadır. Bu aşamada her bir gözlem için, diğer gözlemlerle karşılaştırılmak suretiyle bireysel gelişimi izlenebilir. Düzey-1 analizlerinin amacı her bir kişinin bireysel gelişiminin şeklini belirlemektir.

Düzey-2 olarak bilinen ikinci aşamada gözlemin kendi içindeki değişimi hesaplanabilmektedir. Bu aşamada uygulanan analizlerin amacı, gözlemdeki değişimin heterojenliğini bulmak ve her bir kişinin bireysel gelişiminin şekli ile kestiriciler arasındaki ilişkinin belirlenmesini sağlamaktır.

Bunun yanı sıra dikey kesit veri analizlerinin seçiminde dikkat edilmesi gereken başka özellikler de vardır. Örneğin gizli değişkenin varlığı, verilerin ölçüm durumu (derecesi), verilerin dağılımı ve analiz için varsayılan dağılım, değişkenlerin bölünmesi durumu, gözlem sayısı, incelenen zaman-dalga sayısı ve uzunluğu gibi...

Dikey kesit verilere uygulamalarda çoklu karşılık getirme analizi, doğrusal olmayan genelleştirilmiş kanonik analiz, N=1 veya N>1 için optimal ölçeklemeyle doğrusal dinamik sistem analizi, çok değişkenli varyans analizi, yapısal denklem modelleri, çok düzeyli analizler, log-doğrusal analizler ve markov modelleri uygulanabilmektedir. Analiz teknikleri ve bu analizler için gereken özellikler tablo 1.5'de gösterilmiştir.

Tablo 1.5. Dikey Kesit Analiz Tekniklerinin Özellikleri

Dikey Kesit Analiz Tekniği	Gizli Değişken	Ölçüm Derecesi	Varsayılan Dağılım	Değişkenlerin Bölünmesi (Partitioning)	Gözlem Sayısı	İncelenen Zaman Sayısı
Çoklu Karşılık Getirme Analizi	evet	kategorik	yok	hayır	önemsiz	önemsiz
Doğrusal Olmayan Genelleştirilmiş Kanonik Analiz	evet	kategorik	yok	evet	önemsiz	önemsiz
N=1 için Optimal Ölçeklemeyle Doğrusal Dinamik Sistem Analizi	evet	kategorik	yok	evet	1	geniş
N>1 için Optimal Ölçeklemeyle Doğrusal Dinamik Sistem Analizi	evet	kategorik	yok	evet	önemsiz	orta veya geniş
Çok Değişkenli Varyans Analizi	hayır	sürekli	normallik	evet	fazla	orta
Yapısal Denklem Modeli	evet	sürekli	normallik	evet	geniş	az
Çok Düzeyli Analiz	hayır	sürekli	normallik	evet	geniş	orta
Log-Doğrusal Analiz	hayır	sürekli	çok terimli	hayır	geniş	az
Markov Modelleri	evet	kategorik	çok terimli	evet	çok geniş	az

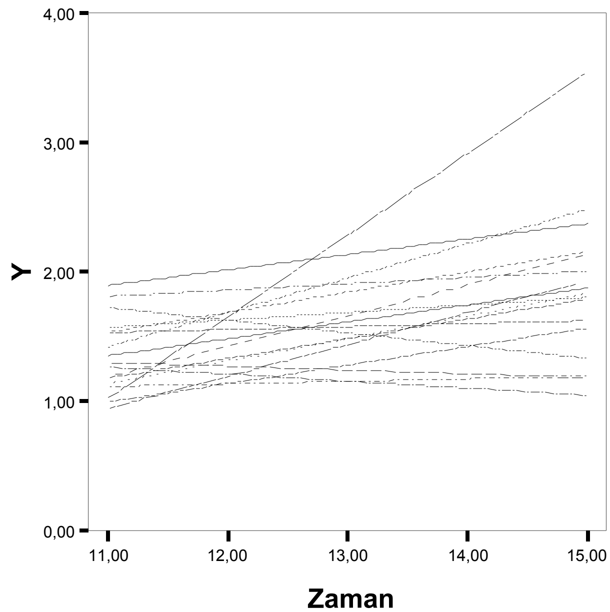
İKİNCİ BÖLÜM
DİKEY KESİT VERİLERİN İNCELENMESİ

Diğer tüm istatistiksel modeller gibi dikey kesit verileri açıklayıcı ve pekiştirici analizler olarak iki aşamada incelenirler. Açıklayıcı analizler, verilerin açıklandığı ve gerek görsel gerekse anlamsal olarak gösterilebildiği ilk aşamadır. Pekiştirici analizler ise belli sonuçlara ulaşılan, verilerin ağırlıklarının belirlendiği, yorumlarının ve hipotezlerin kurularak testlerin yapıldığı aşamadır.

Açıklayıcı analizler dikey kesit verilere çeşitli yöntemlerle uygulanabilirler. Çalışmaya ait açıklayıcı değişkenler saçılım grafiğinde, açıklanması ve yorumlanması açısından basit bir şekilde gösterilebilir. Ancak veri kümesi genişledikçe gösterim daha karmaşık ve zor hale gelmektedir. Bunun yanı sıra düzleştirme (smoothing) teknikleri de uygulanarak açıklayıcı değişkenin önemli değişimleri görsel olarak daha iyi gösterilebilmektedir. Bunun için düzleştirme eğrileri (smoothing lines), kernel kestiricileri ve sağlam (robust) yöntemlerin bilinmesi gerekmektedir. Her bir bireyin tekrarlı gözlemleri arasındaki birleşimin açıklandığı yöntemler de açıklayıcı analizler arasına girmektedir. Eğer eşit zaman aralıklı bir veri kümesi varsa korelasyonlar, eşit olmayan zaman aralığına sahip bir veri kümesi varsa variogramlar birleşimin açıklanmasına yardımcı olur. Analizde kullanılacak veri kümesi kategorik dikey kesit veriler içeriyorsa, açıklayıcı analizlerde lorelogramlar, grafiksel bir yöntem olarak kullanılabilir.

2.1. Dikey Kesit Verilerin Grafiksel Gösterimi

Dikey kesit veriler grafikler ile iki şekilde açıklanabilir. Şekil 2.1. bireylerin belirlenen süreç içerisinde gelişimlerinin izlenmesini sağlamaktadır. Ancak, gözlem sayısı arttıkça grafik karmaşık bir hale gelecek ve bireylerin tek tek gözlenmesi ve gelişimleriyle ilgili yorum yapılması daha zorlaşacaktır.



Şekil 2.1. Bireylerin Gelişimini Gösteren Açıklayıcı Grafik

Gözlem sayısının fazla olması durumunda yorumlama işleminin daha kolay olabilmesi için, her bir gözlem standartlaştırılarak ikinci bir grafik çizilebilir. Standartlaştırma işlemi her bir gözlemden ortalamanın ($\overline{y}_j$) çıkartılması ve standart sapmaya (s_j) bölünmesi ile gerçekleştirilir.

$$y_{ij}^* = \frac{(y_{ij} - \overline{y}_j)}{s_j} \qquad (2.1)$$

Böylece sonuçlar “0” değeri etrafına yerleştirilecek ve her bir gözlemin artışı veya azalışı daha rahat gözlenip genel anlamda yorumlama yapılabilecektir.[18]

[18] A. g. e., Diggle, p: 41.

2.2. Dikey Kesit Verilerinin Düzleştime Eğrilerine Uygunluğu

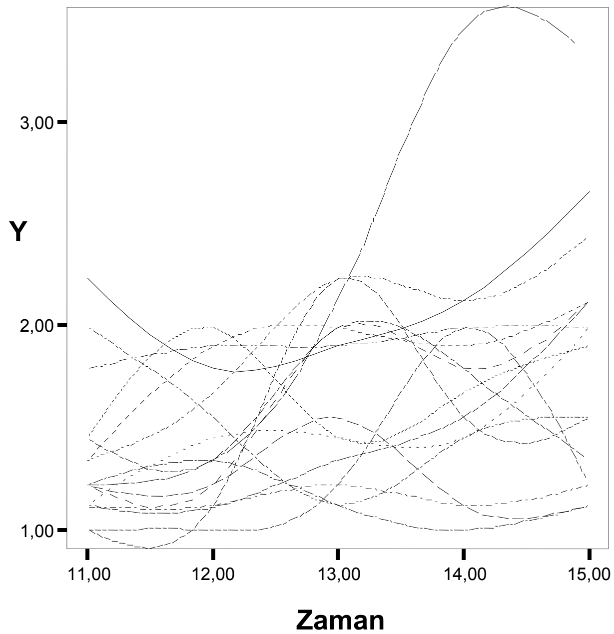
Gözlemlerin zaman içerisindeki gelişimlerini incelerken, her bir gözlemin gelişme eğilimi için ayrı parametrik modeller oluşturularak, gözlem değerlerinin bu modellere yerleştirilmesi ile eğrisel bir gösterim de sağlanabilir.

Modellerin kullanılacağı bu gösterim biçimi üç adımda elde edilir.

1. Veri kümesindeki her bir gözlem için incelenen zaman sürecini içeren bir regresyon modeli kestirilir. Doğrusal değişimi gösterecek bir model olması beklenen bu regresyon modelinde, kişi-dönem veri kümesinde düzenlenmiş olan zaman değişkeni ile incelenen gözlemin bağlı olduğu değişkene regresyon analizi uygulanır. Burada tekrar belirtilmesi gereken en önemli konu, her bir gözleme ait ayrı bir regresyon denkleminin bulunması gerektiğidir.
2. Elde edilen regresyon modellerinden tanımlayıcı istatistikler hesaplanır ve ayrı bir veri kümesi halinde düzenlenir. Doğrusal değişimi gösteren modelde hesaplanan sabit ve katsayılar, gözlemler için ayrı ayrı gelişim durumlarını göstermektedir. Ayrıca R^2 ve artıklara ilişkin varyans değerleri gözlemlerin modeller için uyum iyiliklerini göstermektedir.
3. Her bir gözleme ait regresyon doğruları koordinat ekseni üzerine yerleştirilir. Gözlemlerin var olan değerleri, regresyon modelinden elde edilecek kestirilmiş değerler ile yer değiştirir ve işaretlenir. 2.2. nolu şekil regresyon modelleri yardımıyla çizilmiş bir grafiğe örnek olarak gösterilebilir.[19]

---

[19] A. g. e., Singer, p:29.



Şekil 2.2. Düzleştirme Eğrileriyle Bireylerin Gelişimini Gösteren Açıklayıcı Grafik

Verilerin her zaman doğrusal bir regresyon modeline uyması beklenmeyebilir. Bu durumda katsayıların kestirilmesi için farklı yöntemler kullanılabilir. Kernel ve spline kestiricileri diğer yöntemlere örnek olarak gösterilebilirler. Veriler için elde edilen düzleştirme eğrileri belirtilen yöntemlerden hangisiyle elde edilmiş olursa olsun yaklaşık olarak aynı sonucu verirler.

Sadece y_i gözlemin t_i zamanında ele alındığı varsayılsın. Buradaki genel sorun böyle belirlenmiş verilerden $\mu(t_i)$ bilinmeyen ortalamayı kestirebilmektir. $\mu(t_i)$,

$$Y_i = \mu(t_i) + \varepsilon_i \tag{2.2}$$

modelinden elde edilir. Burada ε_i değerleri bağımsız ve sıfır ortalamaya sahip hatalardır.

Kernel kestirim yönteminde, her bir gözlem için $\hat{\mu}(t)$ değerleri hesaplanır. $\hat{\mu}(t)$, y_i değerlerinin genellikle 0 veya 1 değerleri kullanılarak ağırlıklandırılmış ağırlıklı ortalama değeridir. En bilinen ağırlık fonksiyonu olan Gaussian kernel, K(u) = exp (-0,5u^2) olarak bilinir. Bunlara bağlı olarak kernel kestiricisi

$$\hat{\mu}(t) = \sum_{i=1}^{m} w(t, t_i, h) y_i \Big/ \sum_{i=1}^{m} w(t, t_i, h) \qquad (2.3)$$

olarak gösterilir. Burada $w(t, t_i, h) = K\{(t - t_i)/h\}$ olarak açıklanır. *h* ise kernel'in bant genişliği değerleridir. *h*'nin büyük değerleri daha düzgün eğrilerin elde edilmesini sağlamaktadır.

En çok kullanılan ikinci parametrik olmayan eğri düzleştirici (smoothing) splinelardır. Kübik düzleştirici spline fonksiyonu *s(t)*;

$$J(\lambda) = \sum_{i=1}^{m} \{y_i - s(t_i)\}^2 + \lambda \int \{s''(t)\}^2 dt \qquad (2.4)$$

kriterini minimize etmektedir. Burada $s''(t)$, *s(t)*'nin ikinci dereceden türevidir. Yukarıdaki fonksiyonun ilk terimi, *s(t)* fonksiyonunun y_i gözlemlerine bağlılığını ölçmektedir. Fonksiyondaki integral çarpıklığı, sabit terim λ ise düzgünlüğün derecesini vermektedir. λ'nın küçük olması eğrinin düzgünlüğünün azaldığını göstermektedir.

2.3. Dikey Kesit Analizlerinde Örneklem Büyüklüğü ve Testin Gücü

Çalışmanın düzeni hazırlanırken karşılaşılabilecek ilk sorun çalışmanın örneklem büyüklüğünün belirlenmesi ve testin gücünün hesaplanmasıdır.

Çalışmanın büyüklüğü, örneklem büyüklüğü olarak da ifade edilen gözlem sayısı ile ilgilidir. Ancak dikey kesit çalışmalarda büyüklük kavramı biraz daha karmaşıktır.

Dikey kesit çalışmalardan elde edilmiş çok değişkenli durumlar söz konusu olduğunda örneklem büyüklüğü ve testin gücü tanımları daha karmaşık olacaktır ve matris dönüşümleri ile iteratif çözümler gerektirecektir.

Zaman içerisinde gerçekleşen iki denemenin incelendiği varsayılsın. Her bir deneme N sayıda gözlem içersin ve bu gözlemlerden rasgele n birim çekilsin (birimler birbirinden eşit uzaklıkta olmak zorunda değildir). Çalışmada iki grup, ortalamalarındaki değişimlerine göre karşılaştırılır. Ortalamadaki değişim doğrusal trendle açıklansın. Denemelerin etkileri, değişimin oranları veya eğilimleri arasındaki fark olarak açıklansın ve δ ile ifade edilsin. Bu durumda sıfır hipotezi denemeler arasında fark olmaması olarak kurulur ve H_0: $\delta=0$ olarak ifade edilir.

Örneklem büyüklüğü ve testin gücünden bahsetmeden önce H_0: $\delta=0$ testi için hesaplanmış olan birinci ve ikinci tür hataların bilinmesi gerekir. Birinci tür hata testin anlamlılığı olarak da bilinir ve α ile gösterilir. Birinci tür hata aşağıdaki gibi ifade edilir.

$$\alpha = P(H_0 \text{ red} \mid H_0 \text{ doğru}) \tag{2.5}$$

İkinci tür hata γ ile gösterilebilir ve aşağıdaki gibi açıklanır.

$$\gamma = P(H_0 \text{ kabul} \mid H_0 \text{ yanlış}) \tag{2.6}$$

Bu ifadelere bağlı olarak testin gücü de;

$$\text{güç} = 1 - \gamma = P(\, H_0 \text{ red} \mid H_0 \text{ yanlış}) \tag{2.7}$$

olarak açıklanır.

İki denemeye ait bir dikey kesit çalışmasında örneklem büyüklüğü iki adımda kestirilir. İlk adımda her bir gözleme uygun basit parametrik bir eğrinin var olduğu varsayılır (zaman değişkenin de doğrusal trendin olması gibi). İkinci aşamada bireylere ait elde edilmiş parametreler, bireylerin bulunduğu farklı gruplara ilişkin kovaryanslarla ilişkilendirilir.

Adım 1: Bu adımda bireylerin aynı kovaryans kümesine sahip olduğu ama her bir deneme için ayrı regresyon katsayıları olduğu bilinmektedir. Katsayıların belirtildiği model aşağıdaki gibidir.

$$Y_{ij} = \beta_{1i} + \beta_{2i} t_j + e_{ij} \tag{2.8}$$

Modelde e_{ij} sıfır ortalamalı ve σ_e^2 varyanslı normal dağılıma sahip bağımsız hata terimleridir ($e_{ij} \sim N(0,\ \sigma_e^2)$).

Adım 2: İkinci adımda bireylerin etkilerinin, $\beta_i = (\beta_{1i}, \beta_{2i})'$, rasgele olduğu varsayılır. İkinci aşamada modellenecek olan β_{1i} ve β_{2i}'ye ait ortalama ve kovaryanslar, anakütleye ait parametrelerdir. Her bir deneme grubu için β_i, bireyler arası kovaryansların fonksiyonu gibi modellenir. Deneme gruplarına bağlı olarak β_i'nin ortalaması

$$E\,(\beta_{1i}) = \beta_1 + \beta_2 Grup_i \tag{2.9}$$

$$E\,(\beta_{2i}) = \beta_3 + \beta_4 Grup_i \tag{2.10}$$

şeklinde gösterilir. Bu eşitliklerde eğer i'inci birey incelenen denemeye ait ise $Grup_i$ = 1, ve diğer durumlarda $Grup_i$ = 0 değerini alacaktır. Bu modelde β_3 zaman içindeki ortalamada meydana gelen değişimin ortalama eğimi veya sabit oranıdır. $\beta_3 + \beta_4$ ise ikinci deneme grubuna ait ortalama eğimi göstermektedir.

β_i için belirlenen kovaryans matrisi ise

$$\text{Cov}\ (\beta_i) = \mathbf{G} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} \tag{2.11}$$

olarak gösterilir. Burada g_{11} = Var (β_{1i}), g_{22} = Var (β_{2i}), ve g_{12} = g_{21} = Cov (β_{1i}, β_{2i}) olarak açıklanır.

Bu iki aşamada gösterilmiş olan formüller örneklem büyüklüğü ve testin gücü hesaplarının kolay kullanılabilmesi için açıklanmıştır. Eğer her bir gözlem t_1, ..., t_n zamanlarında gözlenmişse ve her iki deneme grubu için N gözlem varsa (toplam 2N gözlem), yukarıdaki açıklamalar da dikkate alınmak koşulu ile, örneklem büyüklüğü ve testin gücü için yapılacak hesaplamalar yatay kesit veri analizine benzerlik gösterecektir. $\hat{\beta}_{2i}$, i'inci gözlem için eğimin sıradan en küçük kareler (OLS) kestiricisi olsun. $\hat{\beta}_{2i}$ için varyans değeri,

$$\sigma^2 = Var(\hat{\beta}_{2i}) = \sigma_e^2 \left\{ \sum_{j=1}^{n} \left(t_j - \bar{t}\right)^2 \right\}^{-1} + g_{22} \tag{2.12}$$

şeklinde ifade edilir. Burada $\bar{t}$ değeri aşağıdaki gibi gösterilir.

$$\bar{t} = \frac{1}{n} \sum_{j=1}^{n} t_j \tag{2.13}$$

Böylece $\hat{\beta}_{2i}$'nin varyans değeri iki bileşenden oluşturulmuştur. Gözlemler içi varyans, $\sigma_e^2 \left\{ \sum_{j=1}^{n} \left(t_j - \bar{t} \right)^2 \right\}^{-1}$, ve gözlemler arası varyans, g_{22} = Var ($\hat{\beta}_{2i}$). İki deneme grubu için ortalama eğimin eşitliği testinde ise;

$$Z = \frac{\overline{\beta}_2^{(T)} - \overline{\beta}_2^{(C)}}{\sigma\sqrt{\frac{1}{N} + \frac{1}{N}}} = \frac{\overline{\beta}_2^{(T)} - \overline{\beta}_2^{(C)}}{\sigma\sqrt{\frac{2}{N}}} \tag{2.14}$$

eşitliği kullanılır. Burada $\overline{\beta}_2^{(T)}$ ve $\overline{\beta}_2^{(C)}$, $\hat{\beta}_{2i}$'de kullanılan deneme ve kontrol gruplarının örneklem ortalamaları ve σ^2 =Var ($\hat{\beta}_{2i}$) olacaktır.

Yatay kesit veri analizinde kullanılan örneklem büyüklüğü formülüne benzer olarak dikey kesit veri analizi için kullanılacak formül

$$N = \frac{\left\{ Z_{(1-\alpha/2} + Z_{(1-\gamma)} \right\}^2 2\sigma^2}{\delta^2} \tag{2.15}$$

olacak ve burada σ^2, (2.12) formülünde tanımlandığı gibi olacaktır. δ ise deneme grubundaki eğimlerin veya oranların ortalamasındaki değişimin farkı olarak açıklanabilir.

Örneklem büyüklüğünü belirlemede kullanılan formül dikkate alındığı zaman, örneklem büyüklüğü ve testin gücünün aşağıdaki nedenlerden etkilendiğini söylenebilir.

(i) Çalışmanın uzunluğu

(ii) Tekrarlanan gözlemlerin sayısı

(iii) Tekrarlanan gözlemlerin birbirlerine olan uzaklığı

Örneklem büyüklüğü formülü yardımıyla testin gücü;

$$Z_{(1-\gamma)} = \sqrt{\frac{N\delta^2}{2\sigma^2}} - Z_{(1-\sigma/2)} \qquad (2.16)$$

şeklinde gösterilir. Yatay kesit çalışmalarda 1- γ olarak gösterilen testin gücü, dikey kesit çalışmalarda $\Phi\{Z_{(1-\gamma)}\}$ olarak gösterilir. Burada $\Phi(.)$ birikimli standart normal dağılım fonksiyonunu ifade etmektedir. $Z_{(1-\gamma)}$ verilen eşitlik yardımıyla hesaplanır ve standart normal eğrisi altında kalan ve bu değerin sol tarafına doğru uzanan alan testin gücünü verir.

Bu hesaplamalar kayıp gözlemlerin olmadığı durumlarda geçerlidir. Eğer verilerde kayıp gözlemlerle karşılaşılıyorsa örneklem büyüklüğü oran yönteminin kullanılmasıyla belirlenebilir. Bu durumda kayıp olan gözlemler veri kümesinden çıkarılır ve N birimlik anakütleden, çıkarılan gözlemlerin toplam anakütleye oranı kadar birim içeren anakütle belirlenir.

2.4. Dikey Kesit Çalışmalarında Kayıp Gözlemler

Kayıp gözlemler çalışma boyunca elde edilemeyen, gözlenemeyen veya kaydedilmeden kaybedilen verilerdir. Yatay kesit çalışmalarda olduğu gibi dikey kesit çalışmalarda da zorluklara neden olan kayıp veriler, çalışma sırasında iki şekilde ele alınabilir. Kayıp gözlemler veri kümesi içinde yer alırlar ve analizler bu veri kümesine uygulanır veya kayıp gözlemler veri kümesinden çıkarılır ve geri kalan gözlemler kullanılarak analizler uygulanır. Ancak yeterli büyüklükte bir örneklem yoksa ve kayıp gözlem sayısı fazlaysa, bu gözlemlerin veri kümesinden atılması sorunlara neden olacaktır.

Kayıp gözlemler dikey kesit çalışmalar için üç önemli soruna neden olmaktadır:

(1) Dikey kesit çalışmalarda kayıp gözlemler varsa, incelenen bütün gruplarda zaman dilimi boyunca aynı sayıda tekrarlı ölçüm olmayacağı için, veri kümesi zaman içerisinde mutlaka dengesiz hale gelecektir. Böylece kayıp veriler, dengelenmiş veri kümeleri (her zaman döneminde eşit sayıda gözleme sahip olan veri kümeleri) gerektiren birçok analizin uygulanmasında sorun çıkmasına neden olacaktır. Bu dengesizlik yüzünden sorunlarla, özellikle regresyon yöntemlerinin uygulanacağı analizlerde karşılaşılmaktadır.
(2) Çalışmada kayıp gözlemlerin bulunması demek, mutlaka bilgi kaybedildiği anlamına gelmektedir. Bu durumda kayıp gözlemler etkinliğin azalmasına veya kestirilecek ortalamanın zaman içindeki değişiminden dolayı çalışmanın doğruluğunda bir azalmanın meydana gelmesine neden olacaktır. Çalışmanın doğruluğundaki azalma, veriler içinde bulunan kayıp gözlem sayısıyla doğrudan ilişkilidir. Kayıp gözlemin sayısının artması çalışmanın doğruluğu ve güvenilirliğinin daha fazla azalmasına neden olmaktadır.
(3) Belirli koşullar altında kayıp gözlemler önyargının ortaya çıkmasına ve ortalamadaki değişimler hakkında anlam çıkarmada yanlışlıklar yapılmasına neden olabilmektedir.[20]

Verilerde, herhangi bir gözlem için, incelenen zaman aralığında bir kayıp gözlem varsa, gözlemin tüm zaman birimlerine ait değerlerinin tamamının silinmesi birçok bilginin kaybolmasına neden olabileceği için uygulanmamalıdır.[21]

2.5. Kategorik Verilerde Dikey Kesit Analizleri Kavramı

Sürekli değerler alan dikey kesit veriler için, farklı zamanlarda ölçülmüş olan değişkenler arasındaki ilişki korelasyonla açıklanır. Kategorik veriler içeren değişkenler söz konusu olduğunda ilişkinin yorumlanabilmesi için

[20] G. M. Fitzmaurice, N. M. Laird, J. H. Ware, "Applied Longitudinal Analysis", Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Publication, New Jersey, 2004, pp:376,404.
[21] A. Von Eye (Edited by), "Statistical Methods in Longitudinal Research-Volume I, Principles and Structuring Change, Academic Press, London, 1990, p:40.

log-olabilirlik oranı (log-odds ratio) kullanılır. İki sonuçlu değerler içeren Y_1 ve Y_2 değişkenleri için olabilirlik oranı (risk oranı);

$$\gamma(Y_1, Y_2) = \frac{P(Y_1 = 1, Y_2 = 1) P(Y_1 = 0, Y_2 = 0)}{P(Y_1 = 1, Y_2 = 0) P(Y_1 = 0, Y_2 = 1)} \qquad (2.17)$$

şeklinde hesaplanır ve log-olabilirlik oranı ise LOR = $\log \gamma$ biçiminde olacaktır. t_1, ..., t_n zamanlarında ölçülmüş olan Y_1, ..., Y_n değerleri için hazırlanmış olan dikey kesit bir düzende, LOR ifadesi özelleşmiş ve lorelogram adını almıştır.

$$\text{LOR}(t_j, t_k) = \log \gamma(Y_j, Y_k) \qquad (2.18)$$

y_{ij} ($j = 1, ..., n_i$ ve $i = 1, ..., m$) şeklindeki dikey kesit veriden lorelogramları kestirebilmek için γ'nın hesaplanmasında bulunmuş olan teorik oranlar kullanılmaktadır.[22]

[22] A. g. e., Diggle, p:53.

ÜÇÜNCÜ BÖLÜM
DİKEY KESİT VERİLERİN ÇOK DÜZEYLİ ANALİZLERİ

Çok düzeyli analizler (multilevel analysis), hiyerarşik doğrusal regresyon modellerinin açıklanmasında kullanılan yöntemlere verilen genel bir isim olarak tanımlanabilir. Çok düzeyli analizler son 15 yıl içerisinde geliştirilmiş ve uygulaması artmıştır. Bu konuyla ilgili ilk çalışmalar eğitim istatistiklerinde görülmüştür. Sosyal ve davranış bilimlerinde hiyerarşik yapılanmış veriler oldukça sık kullanılmaktadır. Örneğin işçilerin departmanlar içinde gruplanması, sporcuların takımlar içinde gruplanması gibi. Çok düzeyli sözcüğü bu şeklide bir yapıya sahip verileri ve bu yapıdaki verilerin analizlerini açıklamakta kullanılmaktadır. Çoklu düzey noktaları, çalışmanın konusu olmuş birimin, gözlem boyunca gruplanmış verilerle olan ilişkisini göstermektedir. Farklı düzeyler, verilerden ayırt edilmiş olan farklı grupları göstermektedir.

3.1. Çok Düzeyli Veriler

Hiyerarşik veya çok düzeyli veriler, gözlenen birimlerin *kümeleri* olarak da açıklanabilir.[23] Dikey kesit çalışmalarda kümeler, bireyin farklı durumları için elde edilmiş ölçümlerin bir araya getirilmesi ile oluşturulur. Sağlık bilimlerindeki bazı çalışmalarda verilerin dikey kesit yapıya sahip olmamasına rağmen, kümelenmiş olması durumuna da rastlanır. Uygulamada verilerin birden çok düzeyde kümelendiği de görülmektedir. Çok düzeyli veri (veya hiyerarşik veri) kavramı bu tür durumlarla karşılaşıldığında kullanılmaktadır. Verileri çok düzeyli hale getirmenin en kullanışlı yanı, birimlerin küme içi ölçümlerinin, farklı kümelerdeki ölçümlerinden daha kolay biçimde elde edilmesini sağlamasıdır.

[23] A. g. e., Catrien, p:270.

Dikey kesit veriler çok düzeyli verilerin özel bir halidir.[24]

Raudenbush (1988) çalışmalarında, araştırmalardaki verilerin hiyerarşik yapısını açıklamak için kullanılan geleneksel “tek düzey” modellerin başarısızlığını göstermiştir. Bu yargı bir çok bilim adamı tarafından kabul edilmiştir. Çünkü geleneksel regresyon modelleri homojen bir topluluktan çekilen ve bağımsız gözlemler içeren rasgele örneklerden oluşan verilere uygulanmaktaydı. Çok düzeyli verilerin kümelenmiş yapısı, hiyerarşinin her düzeyinde gözlenen birimler arasında bir bağımlılığın oluşmasına neden olmaktadır. Örneğin, aynı sınıftaki öğrenciler birbirlerine farklı sınıflardaki öğrencilerden daha çok benzerler. Eğitim araştırmasındaki amaç grup özelliklerinin her bir öğrencinin karakteri üzerindeki etkilerini incelemek ise, bu tarz durumlarda geleneksel regresyon modellerinin uygulanması sorunlara neden olacaktır. Grup düzeyi değişkenlerinin etkileri, öğrencilerin, sınıfların veya okulların, analizin bir birimi olarak ele alınmasıyla incelenebilir. Sınıf düzeyleri için yapılan regresyon analizinde (genellikle öğrenci seviye değişkenlerini içerir) tüm sınıf seviyeleri bir araya getirilir. Bu analizlerden elde edilen regresyon katsayıları öğrenci düzeyleri regresyonundan elde edilen denk katsayılardan farklılaşma eğilimi gösterirler. Bu tutarsızlığa “toplama eğilim” adı verilmektedir. Bu tutarsızlık, analizden farklı düzeylerin seçilmesinin, belirli değişkenlerin etkilerinin farklı yorumlanmasına neden olması olarak açıklanabilir.

Verilerin hiyerarşik yapıya sahip olması durumunun görmezden gelinmesi, analizin sonunda elde edilecek kestirimlerin yanlış olmasına neden olacaktır. İç içe geçmiş (nested) veriye en küçük kareler yöntemi uygulandığı zaman standart hata kestirimleri, verilerdeki gruplaşmanın etkileri olarak açıklanabilecek olan (ko)varyans bileşenlerini içereceğinden beklenen sonuçları vermeyecektir.

[24] A. g. e., Fitzmaurice, p:441.

Eğer verilerdeki hiyerarşi hesaplanmazsa toplama eğilim ve kestirimlerin yanlış olması problemleri, standart hipotez testlerinin uygulanabilir olması durumunu da ortadan kaldıracaktır.

Çok düzeyli analizler araştırmacılara hiyerarşik verilerin ele alınabilmesi için hem model formülasyonlarına hem de hipotez testlerine bağlı kalarak oldukça esnek ve güçlü araçlar sunmaktadır. Bu analizlerde değişkenler farklı düzeylerdeki kümelenmelerin değişik etkilerini gösterme esnekliğine sahiptirler. Birimler arasındaki etkilerin farklı seviyelerde değişik sonuçlara neden olacağı söylenebilir. Çok düzeyli analizlerde bu değişim ve farklı seviyelerdeki birimlerin birleşmesinden kaynaklanan etkileşim modellenebilir. Ayrıca hata yapılarının çeşitli yollarla açıkça belirtilmesi sonucunda güvenilirlik problemleri de çözülmüş olur. Çok düzeyli modeller, verilerdeki gruplaşmadan kaynaklanan (ko)varyans etkilerinin ortaya çıkarılması için de kullanılabilir. Sonuç olarak, etki ile ilgili gruplar arasındaki farklar, her grup içerisinde farklı dağılımlarla modellenir ve sadece ortalamaların farkı olarak gösterilmez. Bu durum da, çok düzeyli model kullanıldığı zaman hiyerarşik bir dağılımın modellendiği sonucunun çıkarılmasını sağlar.

Çok düzeyli modeller bir çok farklı şekilde isimlendirilmişlerdir. Rasgele katsayı regresyon modelleri (random coefficient regression models) (De Leeuw&Kreft, 1986; Prosser, Rasbash,& Goldstein, 1991), çok düzeyli karışık etkiler modelleri (multilevel mixed effects models) (Goldstein, 1986), rasgele parametre modelleri (random parameter models) (Aitkin&Longford, 1986), varyans bileşenleri modelleri (variance components models) (Aitkin&Longford, 1986), çok düzeyli doğrusal modeller (multilevel linear models) (Goldstein, 1995) ve hiyerarşik doğrusal modeller (hierarchical linear models) (Bryk&Raudenbush, 1992) bu isimlerden en bilinenleridir. Küçük farklılıkların olmasına rağmen bütün modeller hiyerarşik doğrusal regresyon modellerinin farklı versiyonlarıdır.[25]

---

[25] A.g.e., Catrien, p:273.

Çok düzeyli modellerin önemli bir uzantısı dikey kesit verilerin analizi için iki veya daha fazla düzeyli modellerin uygulamasıdır. Tekrarlanmış ölçümlerin analizinde çok düzeyli modellerin uygulanması daha basittir. Dikey kesit verilerin çözümündeki nokta, verilerin hiyerarşik bir yapıya sahip olduklarını bilerek hareket edilmesidir. İncelenen gözlemler için zaman noktaları, bireyler tarafından kapsanırlar. Böylece zamanın farklı noktalarındaki ölçümler için bireyler, iç içe geçmiş olarak kabul edilirler. Burada bireyler hiyerarşinin en üstünde ve zaman noktaları ise en altında varsayılırlar.

Çok düzeyli analizleri açıklamadan önce dikey kesit verilerin çok düzeyli modellenmesi açıklanacaktır.

3.2. Doğrusal Çok Düzeyli Modeller

Çok düzeyli modelleme yaklaşımları aynı temellere dayanmaktadır. Farklı düzeylerdeki rasgele etkilerin incelenmesi için veriler kümelenir. Kümelenen verilerdeki değişimi hesaplayabilmek için çok düzeyli modellerin kullanılması gerekmektedir.

Çok düzeyli modelde istenen bilgi, en düşük düzeydeki (veya düzey 1deki) birimlerden elde edilebilmekteyken, ortak değişkenli bilgiler her düzeyde ölçülebilir. Bu çalışmada açıklanacak olan çok düzeyli doğrusal modeller süreklilik söz konusu olduğu zaman uygulanmaktadır. İlk olarak iki düzeyli verilerin modellenmesi açıklanacak ve daha sonra verilerin daha geniş seviyeli kümelenebilmeleri için (üç veya daha fazla seviyeli düzeyler için) genelleştirilecektir.

3.2.1. İki düzeyli doğrusal modeller

İki düzeyli verilerin modellenmesinde, i indisi düzey1 birimlerini ve j indisi düzey2 birimlerini göstermektedir. Örneklemde düzey2'ye ait n_2 tane birim olduğu varsayılsın. Kümelerin (düzey2 birimleri) her biri ($j = 1, ..., n_2$) n_{1j} tane düzey1 birimlerinden oluşturulmuştur.

Y_{ij} i'inci düzey1 biriminin, j'ninci düzey2 kümesindeki cevabını göstermek üzere her Y_{ij}, $1 \times p$ boyutlu eşlik değişkeni X_{ij}'nin vektörü olduğunda, kurulacak doğrusal modelin ortalaması aşağıdaki gibi gösterilir.

$$E(Y_{ij}) = X_{ij}\beta \tag{3.1}$$

Bireyleri etkileyen birden fazla durumun bulunduğu bir deneme söz konusu olduğu zaman ortalama değer aşağıdaki gibi genelleştirilebilir.

$$E(Y_{ij}) = \beta_1 + \beta_2\, Grup_{ij} \tag{3.2}$$

Burada $Grup_{ij}$, j'ninci merkezdeki i'inci kişinin ilgilenilen sonucunu göstermektedir. Çok değişkenli modelde, Y_{ij}'nin kendi ortalaması etrafındaki değişimi de hesaplanabilir. Y_{ij} için düzenlenmiş çok düzeyli bir model için düzey1 ve düzey2 birimlerinin içindeki regresyon parametrelerinin alt kümeleri içerisinde değişim olduğu varsayılmaktadır. Y_{ij} için iki düzeyli doğrusal model;

$$Y_{ij} = X_{ij}\beta + Z_{ij}b_j + e_{ij} \tag{3.3}$$

şeklinde gösterilir. Bu modelde Z_{ij}, düzey2'deki rasgele etkilerin gösterildiği ve X_{ij}'de ayrılmış olan sütunların bir alt kümesi olarak elde edilen matristir. b_j katsayısı düzey2 birimlerinde değişkenlik gösterirken, tüm düzey1 birimlerinde sabittir. Düzey2 birimleri içindeki rasgele etkilerin, sıfır ortalama ve G kovaryansıyla ($Cov(b_j) = G$) bağımsız olduğu varsayılır. e_{ij} değerinin de düzey1 birimleri için sıfır ortalama ve σ^2 varyansla ($Var(e_{ij}) = \sigma^2$) bağımsız olduğu varsayılır. Ayrıca e_{ij} değerlerinin b_j değerlerinden bağımsız olduğu da varsayılır ($Cov\ (e_{ij},\ b_j) = 0$). Bu durum düzey1 birimlerinin, verilen düzey2 rasgele etkilerinden koşullu bağımsızlığa sahip olduğu varsayımını ortaya çıkarmaktadır.

Regresyon parametresi olan β'lar sabit etkilerdir ve ortalamadaki etkileri tanımlamakta kullanılır.

$$E(Y_{ij}) = X_{ij}\beta \qquad (3.4)$$

ifadesinde ortalama değer, hem düzey1 hem de düzey2 birimlerinin etkilerini taşımaktadır. Bunun yanı sıra koşullu ortalama değeri hesaplanarak tek bir düzeyin ortalama üzerindeki etkisi de bulunabilir.

$$E\,(Y_{ij}|b_j) = X_{ij}\beta + Z_{ij}b_j \qquad (3.5)$$

koşullu ortalamasında sadece düzey1 birimlerinin etkileri görülmektedir.

Bireyleri etkileyen birden fazla durumun bulunduğu deneme tekrar dikkate alınırsa, veriler için iki düzeyli model aşağıdaki gibi gösterilebilir.

$$Y_{ij} = \beta_1 + \beta_2\, Grup_{ij} + b_{1j} + e_{ij} \qquad (3.6)$$

modelinde b_{1j} rasgele merkez (yer) etkisini gösterecektir. Bu katsayı sabit ve bu merkeze ait tüm gözlemler tarafından ortak olarak kullanılmaktadır. Model kurulurken merkezlerin tümü için denemenin etkisinin aynı olduğu varsayılmaktadır. Bu varsayım da denemenin etkisinin merkezler arasındaki değişiminin daha rahat gözlenmesine yardımcı olmaktadır.

$$\begin{aligned} Y_{ij} &= \beta_1 + \beta_2\, Grup_{ij} + b_{1j} + b_{2j}\, Grup_{ij} + e_{ij} \\ &= (\beta_1 + b_{1j}) + (\beta_2 + b_{2j})\, Grup_{ij} + e_{ij} \qquad (3.7) \end{aligned}$$

Yukarıdaki modelde denemelerin etkilerinin, merkezler arasındaki büyüklük değişimleri gösterilebilmektedir. Denemenin ortalama etkisi, β_2 ile gösterilmektedir.

Modellerin kurulduğu durum, gözlemlerin (düzey1) merkezlerde (düzey2) rasgele bulunduğu durumu içermektedir. Diğer bir ifadeyle, gözlemler (düzey1), merkezler (düzey2) içerisinde iç içe geçmiştir, ancak denemeler merkezler ile çaprazlanmıştır. Çünkü her bir merkezdeki denemeler için gözlemler, rasgele olarak düzenlenmiştir. Bu düzenden daha farklı olarak oluşturulabilecek diğer farklı düzende ise, gözlemler (düzey1) tek bir merkezde (düzey2) iç içe geçmiş ve merkezler denemeler için rasgele seçilmişlerdir. Böylece herhangi bir merkezden alınan tüm gözlemlere aynı deneme uygulanmış olacaktır. Bu durumda merkezler denemelerle çaprazlanmış değil, iç içe geçmiş olacaktır. Genel anlamda yukarıda açıklanmış olan model, farklı merkezler arasında deneme etkisinin çok fazla değişiklik göstermediği durumlarda, belirtilen tarzda bir yapıyı da göstermekte kullanılabilir. Çünkü her merkez sadece bir deneme grubuna atanmıştır. Ayrıca deneme grubu değişkeni, $Grup_{ij}$, sabit j için i'de değişkenlik göstermemektedir. Bu nedenden dolayı bu değişken, $Grup_j$ ile değiştirilebilir. Ancak merkezlerin denemelerde iç içe geçmesi, denemenin etkisinin kestirilmesinin etkinliğinde negatif etki yaratacaktır.

Merkezlerin denemelerle çaprazlandığı düzen, merkez ağırlıklı olmayan modele göre daha etkindir. Buradaki mantık, aynı konu üzerinde tekrarlı ölçümler yaparak değişimi ölçen dikey kesit çalışmaların, yatay kesit düzenlerden daha etkin olması prensibiyle aynıdır.

Sonuç olarak (3.6) nolu formülde verilmiş olan iki düzeyli model, her biri bir hiyerarşi düzeyini temsil eden iki ayrı model şeklinde de yazılabilir. İki düzeyli model, sadece düzey1 modelinin terimleri olarak açıklanacaksa;

$$Y_{ij} = Z_{ij}\,\beta_j + e_{ij} \qquad (3.8)$$

şeklinde yazılabilir. Burada e_{ij}, sıfır ortalama ve σ^2 varyansla ($Var(e_{ij})=\sigma^2$) düzey1 birimlerinden bağımsızdır. Düzey2 modeli,

$$\beta_j = A_j\beta + b_j, \qquad (3.9)$$

olarak özetlenebilir. Burada da b_j değerlerinin sıfır ortalama ve G kovaryansıyla ($Cov(b_j) = G$), düzey2 birimlerinden bağımsız olduğu varsayılır. Eğer ikinci model birinci modelde yerine konulursa;

$$\begin{aligned} Y_{ij} &= Z_{ij}\,(A_j\beta + b_j) + e_{ij} \\ &= (Z_{ij}A_j)\beta + Z_{ij}b_j + e_{ij} \qquad (3.10) \\ &= X_{ij}\beta + Z_{ij}b_j + e_{ij} \qquad (3.11) \end{aligned}$$

elde edilir. Bu modelde de X_{ij}; $Z_{ij}A_j$ yerine kullanılmaktadır. Çok düzeyli modeli bütün bir halde yazmak yerine, hiyerarşinin her bir düzeyi için ayrı bir model olarak yazmak, hangi düzey için model oluşturulmuşsa, o düzeydeki değişkenler hakkında daha anlaşılır ve daha rahat açıklanabilir sonuçların elde edilmesini sağlamaktadır.

(3.6) nolu formül ile açıklanmış olan iki düzeyli doğrusal modeller, düzey 1 birimlerinin, düzey2 birimlerinin rasgele etkileriyle birleşerek kümelenmesini hesaplamak için kullanılmaktadır. Bu model, yanıtlardaki değişimin iki ana kaynağını göstermektedir:

(i) düzey2 birimlerindeki değişim

(ii) düzey1 birimlerindeki değişim (düzey2 birimleri ile).

Bu iki kaynaktaki değişkenliğin göreli büyüklüğü, verilerdeki kümelenmeyi açıklamaktadır. Düzey2 rasgele etkilerinin varyansı, düzey1'in (düzey2 ile) değişkenliğine göre daha büyükse, verilerin kümelenmesinin derecesi de daha büyük olacaktır.

3.2.2. Üç düzeyli doğrusal modeller

Verilerin hiyerarşisinde birçok düzey söz konusu olabilir. Üç düzeyli veriler söz konusu olduğunda, verilerin kümelenmesinin düzey1, düzey2 ve düzey3 birimlerindeki değişimle meydana geldiği varsayılmaktadır. Üç düzeyli modelin temeli, hiyerarşinin yüksek düzeylerindeki rasgele etkilerin hesaplanması ile elde edilmiş olan yanıtlardaki değişkenliğin gözlenmesidir. Bu durumda model, yanıtlardaki değişimin üç kaynağını göstermektedir:

(i) düzey3 birimlerindeki değişim
(ii) düzey2 birimlerindeki değişim (düzey3 birimleri ile)
(iii) düzey1 birimlerindeki değişim (düzey3 birimleri ile iç içe geçmiş olan düzey2 birimleriyle)

Üç düzeyli veri yapısı için, i indisi düzey1 birimlerini, j indisi düzey2 birimlerini ve k indisi düzey3 birimlerini göstersin. Örneklemde de n_3 tane düzey3 biriminin olduğu varsayılsın. Kümelerin her biri ($k = 1, \ldots, n_3$), n_{2k} düzey2 kümesinden ve n_{1jk} düzey1 birimlerinden oluşmuş olsun.

Y_{ijk}, j'inci düzey2 kümesindeki i'inci düzey1 birimlerinin, k'ıncı düzey3 kümesindeki yanıtlarını ifade etsin. Örneğin çok sayıda kliniğin incelendiği bir çalışmada, Y_{ijk} k'ıncı klinikteki j'inci hasta için i'inci denemenin sonucunu göstermektedir. Ayrıca Y_{ijk}'nın 1×p boyutlu ve X_{ijk} olarak da ifade edilen değişken vektöründen oluştuğu söylenebilir. Bu değişkenler farklı seviyeler için tanımlanmış olmalıdır. Y_{ijk} için üç düzeyli model;

$$Y_{ijk} = X_{ijk}\beta + Z_{ijk}^{(3)} b_k^{(3)} + Z_{ijk}^{(2)} b_{jk}^{(2)} + e_{ijk} \qquad (3.12)$$

şeklinde gösterilmektedir. Burada $Z_{ijk}^{(3)}$, X_{ijk}'nın belirlenmiş sütunlarının alt kümesi olarak elde edilmiş olan ve düzey3'deki rasgele etkilerin gösterildiği düzen matrisidir. $Z_{ijk}^{(2)}$ ise düzey2'deki (X_{ijk}'nın belirlenmiş sütunlarından alt kümesi olarak elde edilmiş olan) rasgele etkilerin gösterildiği düzen matrisidir. Bu notasyonda $b_k^{(3)}$ ve $b_{jk}^{(2)}$ ifadelerindeki üstsimgeler rasgele etkilerdeki değişimin hangi düzeylerde meydana geldiğini göstermektedir. $b_k^{(3)}$ için elde edilmiş olan $Z_{ijk}^{(3)}$ düzen matrisi, düzey2 ve düzey1 birimlerini göstermektedir.

Üç düzeyli modellerde değişkenliğin farklı kaynakları için öne sürülen iki varsayımı vardır.

(i) $b_k^{(3)}$ rasgele etkilerinin düzey3 birimlerinden, sıfır ortalama ve $G^{(3)}$ kovaryansla ($Cov(b_k^{(3)})=G^{(3)}$) bağımsız olduğu varsayılır. Benzer şekilde $b_{jk}^{(2)}$, sıfır ortalama ve $G^{(2)}$ kovaryansla ($Cov(b_{jk}^{(2)})=G^{(2)}$) düzey2 birimlerinden bağımsız olduğu varsayılır. Rasgele etkiler verilen düzey için ilişkili olabilirler ancak düzeyler arasında ilişkiye sahip değildirler.

(ii) Düzey1 rasgele bileşenlerinin, e_{ijk}'ların, düzey1 birimlerinden sıfır ortalama σ^2 varyansla ($Var(e_{ijk})= \sigma^2$) bağımsız olduğu varsayılır. Ayrıca e_{ijk}'lar için $b_k^{(3)}$ ve $b_{jk}^{(2)}$ rasgele etkilerinden de bağımsız oldukları varsayılmaktadır.

Aynı düzeydeki rasgele etkiler, kendi aralarında değil, aynı düzeydeki birimlerle ilişkilidir. Farklı seviyelerdeki rasgele etkilerin birbirinden bağımsız olduğu varsayımı dikkate alınmalıdır. Prensipte e_{ijk} ifadesinin yerine $Z_{ijk}^{(1)}b_{ijk}^{(1)}$ konulabilir. Burada $b_{ijk}^{(1)}$ sıfır ortalama ve $Cov(b_{ijk}^{(1)})=G^{(1)}$ değerlerine sahiptir. Bu durum düzey1 değişkenliğinde heterojenlik durumunun ortaya çıkmasına neden olacaktır.

(3.12) nolu formülde regresyon parametresi, β, değişkenlerin ortalama (düzey1, düzey2 ve düzey3 üzerindeki ortalamalar) üzerindeki etkilerini göstermektedir ve aşağıdaki gibi gösterilir.

$$E(Y_{ijk}) = X_{ijk}\beta \qquad (3.13)$$

(3.12) nolu formül ile verilmiş olan üç düzeyli model ayrıca değişkenlerin koşullu ortalama üzerindeki etkilerini de açıklamakta kullanılabilir.

$$E(Y_{ijk} \mid b_k^{(3)}) = X_{ijk}\beta + Z_{ijk}^{(3)}b_k^{(3)} \qquad (3.14)$$

ifadesinde yanıtların ortalaması sadece düzey1 ve düzey2 birimlerinin etkisi altındadır. Eğer koşullu ortalamada sadece düzey1 birimlerinin etkileri gözlenmek isteniyorsa formül,

$$E(Y_{ijk} \mid b_k^{(3)}, b_{jk}^{(2)}) = X_{ijk}\beta + Z_{ijk}^{(3)}b_k^{(3)} + Z_{ijk}^{(2)}b_{jk}^{(2)} \qquad (3.15)$$

şeklini alacaktır.

3.2.3. Parametre (katsayı) kestirimi

Çok düzeyli modellerin parametreleri, sabit etkili regresyon parametreleri ve her düzeydeki rasgele etkilerin varyanslarıdır. Herhangi bir seviyedeki rasgele etkilerin kestirimleri elde edilebilir. Çok düzeyli doğrusal modeller için, rasgele bileşenlerin çok değişkenli normal dağılıma sahip olduğu varsayılır. Örneğin üç düzeyli modelde, $b_k \sim N(0, G^{(3)})$, $b_{jk}^{(2)} \sim N(0, G^{(2)})$ ve $e_{ijk} \sim N(0, \sigma^2)$ olduğu varsayılmaktadır. Bu dağılım varsayımlarına bağlı olarak, çok düzeyli model parametrelerinin en çok olabilirlik kestirimleri oldukça kullanışlıdır.

β'nın en çok olabilirlik kestirimi, genelleştirilmiş en küçük kareler kestirim yöntemi ile elde edilir.

Genelleştirilmiş en küçük kareler kestirim yöntemi artıklar için daha karmaşık varsayımların bulunduğu istatistiksel modellere uygunluğu sağlayan ve en küçük kareler kestirim yöntemini yardımıyla elde edilmiş olan bir kestirim yöntemidir. Genelleştirilmiş en küçük kareler kestirim yöntemi, artık kareleri toplamını en küçükleyen parametre kestirimlerini araştırmaktadır. Ancak en küçük kareler yönteminin artıklar için bağımsızlık ve sabit varyanslılık koşulu sağlanmadığında uygulanmasının doğru olmamasına karşılık, genelleştirilmiş en küçük kareler yöntemi aynı karma çok düzeyli modellerde olduğu gibi artıkların ilişkisine ve değişen varyanslı bir yapıya sahip olmalarından etkilenmez.

Eğer karma yapıdaki çok düzeyli model genelleştirilmiş en küçük kareler yöntemi ile çözülecekse, sıradan en küçük kareler yöntemi (ordinary least squares) iki kere uygulanır. İlk olarak düzey1 analizlerinde, kişi-düzey veri kümeleri, kişi-özel yığın (numarasına göre) gruplarına bölünür ve zaman değişkeninin gözlem içi regresyonlarına ayrı ayrı eklenir. Daha sonra düzey2 analizlerinde, en son elde edilen birey büyüme parametre kestirimlerine regresyon uygulanır. Karma çok düzeyli modellerin yapısı ve şekli, çok aşamalı analizler yerine, kişi-düzey veri kümesini bozmadan, karma modelin yapısal kısmındaki kestiricilerin elde edilmesini de sağlamaktadır. Bu durum, veri kümesini kişi-özel yığın gruplarına bölmeden sabit etkilerin ($\gamma_{00}, \gamma_{10}, \gamma_{01}, \gamma_{11}$) kestirilmesini sağlamaktadır.

Sıradan en küçük kareler yaklaşımının kişi-düzey veri kümesinde kullanılması uygun olmamaktadır. Kişi-düzey veri kümesiyle elde edilmiş olan bir karma modelde sabit etkilerin kestirimi, genelleştirilmiş en küçük kareler kestirim yöntemini gerektirmektedir. Bu durum açıklanması zor bir durumu da beraberinde getirmektedir. Gerçekte, karma modelde sabit etkilerin kestirimi kişi-dönem veri kümesinin regresyon analizi ile elde edilir ve bunun için genelleştirilmiş en küçük kareler yöntemine gerek duyulur. Ama bu yöntem uygulanırken gerçek hata kovaryans matrisinin şekli ve içeriğinin bilinmesi gerekmektedir (özellikle artık terimlerine ilişkin değişen varyanslılık ve ilişkinin derecesinin bilinmesi gerekir). Anakütleye ait değerlere ulaşılamayacağı için örneklem değerlerinden bilgiler elde edilmeye çalışılır. Sonuç olarak kişi-düzey veri kümesinden elde edilen karma çok düzeyli modellerin analizlerin uygulanmasında anakütleye ait bilmediğimiz veya ulaşamadığımız bazı bilgilerin bilinmesi gerekmektedir.

Üç düzeyli modellerde kullanılan genelleştirilmiş en küçük kareler kestirimi;

$$\hat{\beta} = \left\{ \sum_{k=1}^{n_3} \left(X_k' V_k^{-1} X_k \right) \right\}^{-1} \sum_{k=1}^{n_3} \left(X_k' V_k^{-1} Y_k \right) \qquad (3.16)$$

olacaktır. Burada Y_k, $\sum_{j=1}^{n_{2k}} n_{1jk}$ uzunluğunda, k'ıncı kümedeki tüm ikinci ve birinci düzey birimleri için elde edilen yanıtlar yığınından oluşturulmuş bir sütun vektördür. Benzer şekilde X_k, $\sum_{j=1}^{n_{2k}} n_{1jk}$ ×p boyutlu, k'ıncı kümede bulunan tüm ikinci ve birinci düzey birimlerini içeren değişkenlerden oluşturulmuş bir matristir. V_k ise, k'ıncı kümedeki düzey1 ve düzey2 birimlerindeki gözlemler arası kovaryans olarak açıklanır ve rasgele etkileri olan kovaryans yapısına sahiptir. Bu ifade $G^{(3)}$, $G^{(2)}$ ve σ^2'nin fonksiyonu olarak gösterilir.

$V_k(G^{(3)}, G^{(2)}, \sigma^2)$'nin sınırlandırılmış en çok olabilirlik kestirimi, β'nın kestirilmesiyle aynı yöntemle elde edilir. $G^{(3)}$, $G^{(2)}$ ve σ^2'nin sınırlandırılmış en çok olabilirlik kestirimleri, sınırlandırılmış log-olabilirlik değerlerinin maksimize edilmesiyle elde edilir.[26]

3.3. Genelleştirilmiş Çok Düzeyli Doğrusal Modeller

Bir önceki bölümde, farklı düzeylerdeki rasgele etkilerin incelenebilmesi için kümelerin oluşturulduğu ve verilerin sürekli olması durumunda kullanılabilecek çok düzeyli doğrusal modeller incelenmişti. Bu bölümde ise çok düzeyli modellemenin kesikli verilere nasıl uygulanacağı açıklanacaktır. Bu modeller çok düzeyli genelleştirilmiş doğrusal modeller olarak da belirtilir. Burada açıklanacak modellerin bir önceki bölümde açıklanmış olan modellerden farkı düzey1'deki gözlemlerin dağılımına ilişkin varsayımlardır. Düzey1 gözlemlerinin normal dağılıma sahip olması zorunluluğu yerine,

---

[26] A. g. e. Fitzmaurice, p:454.

gözlemlerin üstel dağılım ailesine (Bernoulli, Poisson vb.) uygunluğu beklenmektedir.

3.3.1. İki düzeyli genelleştirilmiş doğrusal modeller

İki düzeyli genelleştirilmiş doğrusal modellerde kullanılacak notasyonda Y_{ij}, j'inci düzey2 kümesindeki i'inci düzey1 biriminin sonucunu göstersin. Y_{ij}'ler $1\times p$ boyutlu X_{ij} değişkenlerinin satır vektörüdür. Kesikli (ve sürekli) veriler için iki düzeyli modellerin formülleştirilmesinde aşağıdaki varsayımlar dikkate alınır.

1. Rasgele etkiler b_j vektörüne bağlı olarak, her Y_{ij} için koşullu dağılımın üstel ve Var(Y_{ij} | b_j) $= \upsilon\{E(Y_{ij}|b_j)\}\phi$ olduğu varsayılır. Burada $\upsilon(.)$, bilinen varyans fonksiyonu, $E(Y_{ij}|b_j)$ koşullu ortalama fonksiyonu ve ϕ ise ölçek veya dağılım parametresidir. Ayrıca Y_{ij}'lerin, b_j rasgele etkilerinden bağımsız olduğu varsayılmaktadır.
2. Y_{ij} koşullu ortalamasının, aşağıdaki doğrusal kestirici ile rasgele etkilere bağlı olduğu varsayılmaktadır.

$$\eta_{ij} = X_{ij}\beta + Z_{ij}b_j \tag{3.17}$$

$$g\{E(Y_{ij}|b_j)\} = \eta_{ij} = X_{ij}\beta + Z_{ij}b_j \tag{3.18}$$

Burada g(.) link fonksiyonu olarak da açıklanabilmektedir.

3. Rasgele etkilerin olasılık dağılımlarına sahip oldukları varsayılmaktadır.

Bu üç durum sağlandığı zaman iki düzeyli genelleştirilmiş doğrusal modeller uygulanabilmektedir.

3.3.2. Üç düzeyli genelleştirilmiş doğrusal modeller

Üç düzeyli genelleştirilmiş doğrusal modeller, üç boyutlu verilerde düzey2 ve 3'deki rasgele etkilerin değişkenliğinin ölçülmesinde kullanılmaktadır. Üç düzeyli doğrusal modellerde kullanılan notasyon aynen kullanılabilir. Burada Y_{ijk}, k'ıncı düzey3 kümesindeki j'inci düzey2 kümesine ait i'inci düzey1 biriminin sonucunu göstermektedir. Ayrıca daha önceden de açıklandığı gibi Y_{ijk}, $1\times p$ boyutlu farklı seviyelerden elde edilen X_{ijk} değişkenlerinin satır vektörüdür. Üç düzeyli genelleştirilmiş doğrusal modeller için belirlenmiş varsayımlar aşağıdaki gibidir.

1. $b_k^{(3)}$ ve $b_{jk}^{(2)}$ rasgele etkiler vektörlerine bağlı olarak, her Y_{ijk}'nın koşullu dağılımı üstel yapı gösterir ve $Var(Y_{ijk} \mid b_k^{(3)}, b_{jk}^{(2)}) = \upsilon\left\{E\left(Y_{ijk}\middle|b_k^{(3)}, b_{jk}^{(2)}\right)\right\}\phi$ olacaktır. Burada $\upsilon(.)$ varyans fonksiyonunu, $E\left(Y_{ijk}\middle|b_k^{(3)}, b_{jk}^{(2)}\right)$ koşullu ortalama fonksiyonunu ve ϕ ise ölçek veya dağılım parametresidir. Ayrıca Y_{ijk}'ların, $b_k^{(3)}$ ve $b_{jk}^{(2)}$ rasgele etkilerinden bağımsız olduğu varsayılmaktadır.
2. Y_{ijk} koşullu ortalamasının, aşağıdaki doğrusal kestirici ile rasgele etkilere bağlı olduğu varsayılmaktadır.

$$\eta_{ijk} = X_{ijk}\beta + Z_{ijk}^{(3)}b_k^{(3)} + Z_{ijk}^{(2)}b_{jk}^{(2)} \tag{3.19}$$

$$g\left\{E\left(Y_{ijk}\middle|b_k^{(3)}, b_{jk}^{(2)}\right)\right\} = \eta_{ijk} = X_{ijk}\beta + Z_{ijk}^{(3)}b_k^{(3)} + Z_{ijk}^{(2)}b_{jk}^{(2)} \tag{3.20}$$

Burada g(.) link fonksiyonu olarak da açıklanabilmektedir.

3. Rasgele etkilerin sıfır ortalama ve $Cov(b_k^{(3)}) = G^{(3)}$ ve $Cov(b_{jk}^{(2)} = G^{(2)}$ kovaryans matrisleriyle çok değişkenli normal dağılıma sahiptir. Ayrıca farklı seviyelerdeki rasgele etkilerin birbirlerinden bağımsız olduğu varsayılmaktadır.

3.3.3. Parametre (katsayı) kestirimi

Üç düzeyli veriler için β, $G^{(2)}$ ve $G^{(3)}$ parametrelerinin çıkarsamaları marjinal olabilirlik fonksiyonuna dayanmaktadır. Marjinal olabilirlik de, düzey3 birimleri için olasılık yoğunluk fonksiyonları, $f(y_k)$, yardımıyla açıklanabilir. Özel bir durum olarak da, marjinal log-olabilirlik fonksiyonu aşağıdaki toplamla gösterilebilir.

$$\sum_{k=1}^{n_3} \log f(y_k) \tag{3.21}$$

Bu durumda k'ıncı düzey3 biriminin olabilirlik fonksiyonu,

$$f(y_k) = \int \prod_{j=1}^{n_{2k}} \left\{ \int \prod_{i=1}^{n_{1jk}} f\left(y_{ijk} \middle| b_k^{(3)}, b_{jk}^{(3)}\right) f\left(b_{jk}^{(2)}\right) db_{jk}^{(2)} \right\} f\left(b_k^{(3)}\right) db_k^{(3)} \tag{3.22}$$

olarak gösterilir. Burada $f(b_{jk}^{(2)})$ ve $f(b_k^{(3)})$, ayrı ayrı düzey2 ve 3'teki rasgele etkiler için çok değişkenli normal dağılım fonksiyonlarını göstermektedir.[27]

3.4. Dikey Kesit Verilerinin Çok Düzeyli Analizi

İstatistiksel modeller insanların davranışlarının matematiksel gösterimleridir. Bu modeller incelenen topluluktaki hedef kitlenin belli özelliklerinin açıklanmasına yardımcı olur. Ancak istatistiksel modeller yardımıyla elde edilen sonuçlar örneklem için değil, tüm ana kütle için yorumlanır.

İstatistiksel modeller, parametreleri (katsayılar, varyanslar, eğimler vb.) kullanarak anakütleye ait niceliklerin açıklanmasını ve yorumlanmasını sağlar.

---

[27] A. g. e., Fitzmaurice, p:460.

Örnek verilerinden elde edilen istatistiksel modelde, anakütle parametrelerinin bilinmeyen değerleri kestirilir. Kestirim yöntemlerinin bir çoğu, modelle veriler arasındaki uygunluğu gösterebilmek için uyum iyiliğinin (R^2 değeri veya artık varyansı gibi) ölçülmesini gerektirmektedir. Eğer model verilere tam bir uygunluk göstermişse, kestirilmiş parametreler, anakütle hakkındaki sonuçların yorumlanmasında kullanılabilmektedir. Daha sonraki aşamalarda hipotez testleri ve güven aralıkları kullanılarak örneklemden anakütle hakkında yorum yapılması sağlanabilmektedir.

Basit regresyon modelleri, yatay kesit veriler için kullanılmaktadır. Ancak dikey kesit veriler söz konusu olduğu zaman değişim sürecinin gösterilebilmesi için daha farklı istatistiksel modellere ihtiyaç vardır. Kullanılacak modeller belirtilen iki sorunun cevabını içermelidir: gözlem içi değişimler hakkındaki düzey1 soruları ve değişimde gözlemler arası farkı açıklayan düzey2 soruları. Bu durumda değişimin açıklanacağı model iki düzeyin bileşenlerini içermek zorundadır:

(1) gözlemlerin zaman içerisindeki değişimini açıklayan düzey1 alt modelleri

(2) değişimlerin gözlemler arası farkını açıklayan düzey2 alt modelleri

Bu iki bileşenin birlikte ele alınması çok düzeyli istatistiksel model olarak bilinir.

3.4.1. Gözlemdeki değişimin açıklanmasında kullanılan düzey1 alt modeli

Çok düzeyli analizlerde, bireyler birinci düzeyde incelenir ve bireylerin gruplanması ikinci düzeyde dikkate alınır. Grupların bir araya getirilmesi söz konusu olduğunda ise üçüncü bir düzeye ihtiyaç vardır. Düzey sayısının artması bu şekilde devamlılık gösterir. Çok düzeyli modelin düzey1 bileşeni (bireysel büyüme modeli olarak da bilinir), belirlenen zaman dönemi içerisinde, anakütlenin her bir üyesindeki değişimi göstermek için kullanılır.

Burada dikkat edilmesi gereken önemli nokta, gözlenen veri kümesinin modelin oluşturulduğu anakütleden elde edilmiş olmasının gerekliliğidir.

Her biri N_j kadar gözlem içeren J tane grubun olduğu varsayılsın. Gözlem içi değişimin gösterileceği model doğrusal regresyon modeli olarak açıklanırsa aşağıdaki gibi gösterilecektir.

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij} \qquad (3.23)$$

Y_{ij} ve X_{ij}'deki indis değerleri j'inci gruptaki i'inci bireyi göstermektedir. β_{0j} bağımlı değişkenin beklenen değerini gösterecektir. Benzer şekilde β_{1j} eğimi, ilgili bağımsız değişkenin her bir birimindeki ortalama değişimi açıklayacaktır. Bu modelin genelleştirilmiş şekli;

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + ...+ \beta_{Qj}X_{Qij} + e_{ij} \qquad (3.24)$$

olarak gösterilir. Burada X_{qij} (q=1, ..., Q), bağımsız değişkenler kümesindeki bireylerin karakterlerini göstermektedir. Bryk ve Raudenbush bu modeli, organizasyon-düzey modelleri olarak da adlandırmışlardır. β_{qj} regresyon katsayısı, dağılım etkisini göstermektedir. Çünkü Y_{ij} sonuçlarının j "organizasyonunda", kişi karakterinin fonksiyonu gibi gösterilmesini sağlamaktadır.[28]

İki düzeyli uygulamalarda birinci düzeyde, gözlemlerdeki değişim incelenirken zaman faktörü de modele eklendiğinde model çokterimli (polynomial) fonksiyon haline gelecek ve aşağıdaki gibi gösterilecektir.

$$Y_{ti} = \beta_{0i} + \beta_{1i}T_{ti} + \beta_{2i}T_{ti}^2 + \ldots + \beta_{pi}T_{ti}^p + e_{ti} \qquad (3.25)$$

---

[28] A. g. e., Catrien, p:277.

Modelde Y_{ti}, i'inci bireyin t zamanındaki bağımlı değişken değerini göstermektedir. β'lar çok terimli fonksiyonun p dereceli katsayıları ve e_{ti}'ler ise rasgele hata terimleri olarak açıklanır. T_{ti} değişkeni, i'inci bireyin t zamanında zamana bağlı değişken değerini (yaş, hafta sayısı vb.) göstermektedir. $\beta_{gi}T_{ti}^{g}$, (g=1, ..., p) ifadesinin güçleri, bu değişkenin dönüşümlerini göstermektedir (karesel, kübik veya daha yüksek sıralı çok terimli büyüme fonksiyonu gibi).[29]

Zamana bağlı olan modelin uygun matris gösterimi;

$$y_i = T\beta_i + e_i \tag{3.26}$$

şeklinde olacaktır. y_i, Tx1 boyutlu i'inci olay için yanıt vektörü, T, $T\times r$ $(r\leq T)$ boyutlu sabit değişkenler ve bu değişkenlerin dönüşümleri matrisi, β_i, $r\times 1$ boyutlu i'inci kişiye ait birey parametreleri ve e_i ise $T\times 1$ boyutlu rasgele hatalar bileşenleri vektörüdür. T, aşağıdaki gibi gösterilebilir.

$$T = \begin{bmatrix} 1 & T_1 & T_1^2 & \dots & T_1^p \\ 1 & T_2 & T_2^2 & \dots & T_2^p \\ . & . & . & \dots & . \\ 1 & T_T & T_T^2 & \dots & T_T^p \end{bmatrix} \tag{3.27}$$

Bu formülasyonda tüm bireylerin aynı T ölçüme sahip olmaları gerekmemektedir. Açıklayıcı değişkenlerden ve bunların güçlerinden bağımsız olarak (yaş veya hafta sayısından bağımsız olarak), T matrisi standart çok terimli vektörlerin kümesini de içerebilir.

---

[29] A. g. e. Catrien, p:284

3.4.2. Gözlemler arası değişimin açıklanmasında kullanılan düzey2 alt modeli

Düzey2 alt modelleri, bireyler arasında zamana bağlı meydana gelen değişiklikleri açıklamak için kullanılır. Düzey1 doğrusal değişim modeli kullanıldığı zaman gözlemler sadece modeldeki sabit ve eğimle farklılık gösterirler. Bu durum açıklayıcı değişkenler ve değişim arasındaki ilişki hakkında yorum yapılabilmesini sağlar.

Tüm istatistiksel modeller gibi, düzey2 alt modelleri de anakütleye ilişkin sonuçların çıkarılmasında kullanılır.[30] Verilere çok düzeyli denemeler uygulanması demek β_{0j} ve β_{1j} katsayılarının gruplar boyunca değişkenlik göstermesi demektir. Katsayıların değişiklik göstermesiyle verilerin hiyerarşik yapısı modelde bir arada görülebilecektir. Bu etki de sadece düzey2'de gözlenebilmektedir. En kolay düzey2 veya gözlemler arası model, sadece basit rasgele değişimin modellendiği yapılardır. Bu yapı aşağıdaki gibi gösterilebilir.

$$\beta_{0j} = \gamma_{00} + u_{0j} \tag{3.28}$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \tag{3.29}$$

Burada γ_{00}, ilk açıklayıcı değişkenin bağımlı değişken üzerindeki etkisini, γ_{10} diğer açıklayıcı değişkenin bağımlı değişken üzerindeki etkisini ve u_{0j} ile u_{1j} ise j'inci grup boyunca β_{oj} ve β_{1j} değerlerinin γ_{00} ve γ_{10} etrafındaki artış veya azalışını gösteren rasgele terimlerdir. Belirtilen bu iki model gruplar arası model olarak da adlandırılabilir.

Gözlemdeki değişimin açıklanması için kullanılan modelle, gözlemler arasındaki değişimi açıklamakta kullanılan model birleştirildiğinde "rasgele

---

[30] A. g. e., Singer, p:60.

katsayılar modeli" elde edilir. Gözlemler arasındaki ilişkiyi gösteren model, gruplar arası modellere dönüştürülerek genelleştirilebilir. Örneğin Z'nin, j grubundaki birim sayısını ifade ettiği düşünülürse, β_{oj} ve β_{1j}'deki değişimin açıklanmasında kullanılacak modeller aşağıdaki gibi gösterilebilir.

$$B_{oj} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \qquad (3.30)$$

$$B_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \qquad (3.31)$$

Bu modellerde γ_{01}, birim sayısının bağımlı değişkenin ortalaması üzerindeki ortalama etkisini, γ_{11} ise birim sayısındaki farklar için hesaplanan β_{1j} eğimindeki ortalama artışını göstermektedir. Sonuç olarak son yazılan model daha çok bilgi sağlayacağından araştırmalarda daha sık tercih edilmektedir. Düzey2 modellerinin genelleştirilmiş biçimi aşağıdaki gibi gösterilebilir.

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}Z_j + u_{1j} \\ \beta_{2j} &= \gamma_{20} + u_{2j} \\ \cdots &= \cdots \\ \beta_{Qj} &= \gamma_{Q0} + u_{Qj} \end{aligned} \qquad (3.32)$$

İki düzeyli uygulamalarda ikinci düzeyde gözlemler arası değişim incelenirken zaman faktörü de modele eklendiğinde kurulan düzey2 modelleri değişkenlik göstermeyecektir.

Düzey2 modellerin matris olarak gösterimi,

$$\beta_i = Z_i\gamma + u_i \qquad (3.33)$$

şeklindedir. Bu eşitlikte Z_i, r×q boyutlu sabit bileşenlerden oluşan düzen matrisi, γ, q×1 boyutlu sabit katsayılar vektörü ve u_i ise r×1 boyutlu rasgele hata bileşenleri vektörüdür.

3.4.3. Karma doğrusal model

Gözlemdeki değişimin açıklanması için kullanılan düzey1 alt modeli ile gözlemler arası değişimin açıklanmasında kullanılan düzey2 alt modeli bir arada, tek bir model olarak kullanılabilir. Bunun için kullanılacak model;

$$Y_{ij} = \underbrace{\gamma_{00} + \gamma_{01}Z_j + \gamma_{10}X_{ij} + \gamma_{11}Z_jX_{ij}}_{sabit} + \underbrace{u_{0j} + u_{1j}X_{ij} + e_{ij}}_{rasgele} \tag{3.34}$$

olarak gösterilir. Modelde ayrılan ilk kısım sabit etkileri, ikinci kısım ise rasgele etkileri göstermektedir. Bu model yardımıyla modeli etkileyecek sabit ve rasgele etkiler aynı anda gözlenebilmektedir. Sabit kısım, γ katsayıları ile kestirilecek değişkenleri göstermektedir. Rasgele kısım, düzey1 ve düzey2 hata terimlerini içermektedir. Sonuç olarak model parametreleri, sabit parametreler kümesi ve rasgele parametreler kümesi olarak iki gruba ayrılabilirler.

Rasgele parametreleri düzey1 ve düzey2 hata terimlerine bağlı olarak (ko)varyans bileşenlerinden oluşurlar. Eğer;

$$R_{ij} = Y_{ij} - \hat{Y}_{ij} = Y_{ij} - (\gamma_{00} + \gamma_{01}Z_j + \gamma_{10}X_{ij} + \gamma_{11}Z_jX_{ij}) \tag{3.35}$$

için varyans değeri

$$Var(R_{ij}) = Var(u_{0j} + u_{1j}X_{ij} + e_{ij}) = \sigma_0^2 + 2\sigma_{01}X_{ij} + \sigma_1^2X_{ij}^2 + \sigma_e^2 \tag{3.36}$$

olarak gösterilirse, (ko)varyans bileşenleri açıkça görülebilir. Burada $\sigma_0^2 = Var(u_{0j})$, $\sigma_1^2 = Var(u_{1j})$, $\sigma_{01} = Kov(u_{0j}, u_{1j})$ ve $\sigma_e^2 = Var(e_{ij})$ olarak gösterilecektir. Bir önceki bölümde anlatılan modeller dikkate alınırsa; σ_0^2, X_{ij} ve Z_j ile açıklanamayan β_{0j} katsayısının varyansını, σ_1^2'nin β_{1j} eğiminin varyansı ve σ_{01} ise X_{ij} ve Z_j değişkenlerine ait eğim ve sabit arasındaki kovaryansı göstermektedir. σ_e^2 ise artıkların varyansı olarak kullanılmıştır.

(3.34) nolu modelde kullanılan $\gamma_{11}Z_jX_{ij}$ terimi düzey1 ve düzey2 arasındaki etkileşimi göstermektedir. Bu nedenle bu terime "yatay düzey etkileşim terimi" adı da verilmektedir. Bazı araştırmacılar için bu terim çok düzeyli modellerin tercih edilebilir olmasının en önemli sebebidir. Ayrıca buna bağlı olarak düzey1 regresyon katsayıları, ikinci düzeydeki regresyon analizleri için kullanılacak sonuç değişkenleri olarak kullanılacaktır.

Karma modelin matris olarak gösterimi,

$$y_i = \underbrace{TZ_i\gamma}_{sabit} + \underbrace{Tu_i + e_i}_{rasgele} \tag{3.37}$$

şeklinde olacaktır. Burada e_i vektörü $e_i \sim N(0, \Sigma_{(1)})$ olarak bağımsız dağılacaklardır. $\Sigma_{(1)} = \sigma^2 I$ olacaktır. Düzey2 rasgele terimleri u_i'ler ise bağımsız olarak $u_i \sim N(0, \Sigma_{(2)})$ şeklinde dağılacaktır. Düzey1 rasgele terimleri olan e_i'ler, düzey2 rasgele terimleri u_i'lerden bağımsız olacaktır.

Zamanın söz konusu olduğu durumlarda kurulacak karma model aşağıdaki gibi gösterilebilir:

$$Y_{ij} = \pi_{0i} + \pi_{1i}Zaman_{ij} + (\zeta_{0i} + \zeta_{1i}Zaman_{ij} + \varepsilon_{ij}) \tag{3.38}$$

Her bir kişinin her durumdaki ölçümü için j tane karma artıkları vardır. Düzey1 ve düzey2 artıklarını birbirine bağlayan karma artıkların yapısı, dikey kesit veri analizlerinde karşılaşılması beklenen değişen varyans ve otokorelasyon durumlarının ortaya çıkmasını sağlamaktadır.

j'inci ölçümde elde edilecek olan karma modelinin artık terimlerinin varyansı;

$$\sigma^2_{artik_j} = \sigma^2_0 + \sigma^2_1 Zaman^2_j + 2\sigma_{01} Zaman_j + \sigma^2_\varepsilon \qquad (3.39)$$

eşitliği yardımıyla hesaplanacaktır. j ve j' durumundaki karma modelin artıklarına ilişkin otokorelasyon ise;

$$\rho_{artik_j artik_{j'}} = \frac{\sigma^2_0 + \sigma_{01}(Zaman_j + Zaman_{j'}) + \sigma^2_1 Zaman_j Zaman_{j'}}{\sqrt{\sigma^2_{artik_j}\sigma^2_{artik_{j'}}}} \qquad (3.40)$$

şeklinde hesaplanır.

3.4.3.1. <u>değişimin açıklanma oranının sayısal gösterimi</u>

Çoklu regresyon analizinde modelin açıklanma oranı R^2 (veya düzeltilmiş R^2) istatistiği kullanılarak gösterilir. Çok düzeyli modellerde ise bu istatistik değerinin tanımı biraz daha farklılık göstermektedir. Çünkü değişim bir çok varyans bileşeninden ($\sigma^2_\varepsilon, \sigma^2_0, \sigma^2_1$) etkilenmektedir. Bu nedenden dolayı çok düzeyli modellerde modelin açıklanma oranı yapay-R' istatistiği kullanılarak açıklanır.

Koşulsuz ortalamalar modeli σ^2_ε'nin ve koşulsuz büyüme modeli de σ^2_0 ve σ^2_1'nin kestirilmesini sağlamaktadır. Bu nedenden dolayı her iki model de kendine ait yapay R^2 istatistiklerinin hesaplanmasını gerekli kılmaktadır.

Koşulsuz ortalama ve koşulsuz büyüme modelleri ile hesaplanacak olan birey içi artık varyansına (σ_{ε}^2) ilişkin yapay R^2 istatistiği;

$$Yapay R_{\varepsilon}^2 = \frac{\hat{\sigma}_{\varepsilon(kosulsuz\ ortalamalar\ modeli)}^2 - \sigma_{\varepsilon(kosulsuz\ buyume\ modeli)}^2}{\hat{\sigma}_{\varepsilon(kosulsuz\ ortalamalar\ modeli)}^2} \qquad (3.41)$$

eşitliği yardımı ile hesaplanır.

Düzey2 artık varyansları yardımıyla aynı istatistik değeri benzer bir eşitlikle hesaplanabilir. Her düzey2 artık varyans bileşeninin kendine ait yapay R^2 değeri vardır. Bu durumda yapay R^2 değeri;

$$Yapay R_{\zeta}^2 = \frac{\hat{\sigma}_{\zeta(kosulsuz\ buyume\ modeli)}^2 - \sigma_{\zeta(daha\ sonra\ olusturulmus\ model)}^2}{\hat{\sigma}_{\zeta(kosulsuz\ buyume\ modeli)}^2} \qquad (3.42)$$

ile hesaplanır.

Daima pozitif (veya sıfır) değerler alabilen bilinen R^2 istatistiğinden farklı olarak, yapay R^2 istatistiği negatif değerler alabilir. Çok düzeyli modellerde modele değişken katılması varyans bileşenlerini küçülteceğinden dolayı yapay R^2 istatistik değerini arttıracaktır.[31]

3.4.3.2. <u>varsayımlar</u>

Karma doğrusal model olarak gösterilen eşitlikte genellikle aşağıda belirtilecek olan varsayımlar kullanılır.

(1) Açıklayıcı değişkenler X_{ij} ve Z_j, sabit değişkenler olarak varsayılır.

(2) j grubundaki düzey1'e ait rasgele e_{ij} terimlerinin bağımsız ve sıfır ortalama, σ_e^2 varyansla normal dağıldığı varsayılır. Bu varsayım da

[31] A. g. e., Singer, p:104.

düzey1 kovaryans matrisi $\Sigma_{(1)}$'in $\sigma_e^2\mathbf{I}$ ifadesine eşit olması sonucunu doğurur.

(3) Düzey2 rasgele terimleri u_{0j} ve u_{1j}'nin sıfır ortalama, $\Sigma_{(2)}$ kovaryans matrisiyle birleşik normal dağılıma sahip olduğu varsayılır.

(4) Farklı seviyelerdeki rasgele terimler birlerinden bağımsız olarak dağılırlar.

3.4.4. Parametre (katsayı) kestirimi

Çok düzeyli modelin kestirilmesi, karma modeldeki parametrelerin kestirimi ile gerçekleştirilir. Bu durumda karma modelin kestirimi, iki küme parametrenin (sabit ve rasgele parametreler) kestirimi ile gerçekleşecek ancak parametre kümelerinden birinin kestirimi diğer kümeye de bağlıdır. (Ko)Varyans bileşenlerinin bilinmesi sabit etkilerin kestirimini kolaylaştırmaktadır. Ancak (ko)varyans bileşenleri bilinmiyorsa bu parametreler verilerden kestirilmelidir.

Model parametrelerinin iki kümesinin kestirimlerini elde etmek için bazı sayısal yöntemler uygulanmaktadır. Uygulanan tüm sayısal yöntemler en çok olabilirlik yöntemine dayanmaktadır. En çok olabilirlik kestiriminin en önemli özelliği tutarlılık ve etkinlik koşullarının sağlanması gerekliliğidir. Dikey kesit modellerin incelendiği çalışmalarda kullanılmak üzere tam bilgili en çok olabilirlik (full information maximum likelihood) ve kısıtlı en çok olabilirlik (restricted maximum likelihood) yöntemleri geliştirilmiştir. Bu yöntemlerin en büyük sakıncası çok karmaşık bir işleyişe sahip olması ve iteratif işlemlerin uygulanmasının gerekliliğidir.

En çok olabilirlik kestirim yöntemleri; tam (full) ve kısıtlı (restricted) olmak üzere iki grupta incelenmektedir. Bu iki yöntem birbirinden, parametre kestirimlerine etki eden ve hipotez testlerinde gereken stratejilerin oluşturulmasında kullanılacak olabilirlik fonksiyonunun oluşturulması aşamasında birbirinden farklılıklar göstermektedir.

Bilinen olabilirlik yönteminde, verilerin fonksiyonu ve varsayımları tüm bilinmeyen parametreleri (hem sabit etkileri (γ'lar) hem de varyans bileşenlerini ($\sigma_\varepsilon^2, \sigma_0^2, \sigma_1^2, \sigma_{01}$)) içermektedir.

Tam en çok olabilirlik yöntemi ile sabit etkilerin (γ'lar) tam en çok olabilirlik kestirimlerini içeren varyans bileşenleri ($\hat{\sigma}_\varepsilon^2, \hat{\sigma}_0^2, \hat{\sigma}_1^2, \hat{\sigma}_{01}$) kestirilir. Bunun anlamı varyans bileşenleri kestirildiği zaman sabit etkilerin gerçek değerleri bilinmediği halde biliniyormuş gibi işleme katılmaktadır.

Tam en çok olabilirlik yönteminin eksik ve kullanılamayacağı durumlarda alternatif olarak kullanılması için kısıtlı en çok olabilirlik yöntemi geliştirilmiştir. Ancak hem tam hem de kısıtlı en çok olabilirlik yöntemleri çok düzeyli modellerin analizlerinde kullanıldığı zaman arka arkaya ve yoğun iterasyon işlemleri gerektirmektedir.

Farkın daha iyi görülebilmesi için yatay kesit veriler için oluşturulmuş doğrusal regresyon modeline tam en çok olabilirlik yönteminin uygulanması açıklanacaktır. n birimli ve p tane bağımsız değişkene sahip Y bağımlı değişkeni için kurulacak regresyon modeli;

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_p X_{pi} + \varepsilon_i \qquad (3.43)$$

olarak gösterilsin. Burada i indisleri gözlemleri ve ε_i bağımsız ve sıfır ortalama ile homoskedastik σ_ε^2 varyansa sahip normal dağılımlı hata terimleridir. Bu değerlere bağlı olarak hata terimleri;

$$\varepsilon_i = Y_i - (\beta_0 + \beta_1 X_{1i} + ... + \beta_p X_{pi}) \qquad (3.44)$$

olarak hesaplanacaktır. Bilinmeyen artık varyansı σ_ε^2'nin tam en çok olabilirlik kestirimi hata terimlerinin kareleri toplamının örnek sayısına bölünmesi ile hesaplanmaktadır ve aşağıdaki gibi gösterilir.

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^{n} \varepsilon_i^2}{n} \tag{3.45}$$

Burada regresyon katsayılarının anakütle değerlerinin bilindiği varsayılır ve yeniden kestirilmelerine gerek duyulmaz. Ancak uygulamada regresyon parametreleri için anakütlenin değerlerini bulmak mümkün değildir. Bunun yerine örneklemdeki verileri kullanarak kestirim yapılır ve eşitlik

$$\hat{\varepsilon}_i = Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \ldots + \hat{\beta}_p X_{pi}) \tag{3.46}$$

Bu değer (3.45) nolu formülde yerine konursa eşitlik;

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n} \tag{3.47}$$

olarak gösterilir. Bu paydanın kullanılması, bu parametrenin kestirilmesinde örneklemin gerçek serbestlik derecelerinin kullanıldığı varsayılır. Ancak artıkların hesaplanmasında (p+1) tane regresyon parametresi kestirildiği için serbestlik derecesi olarak (p+1) kullanılır. Artık varyansının yansız kestirimleri, (3.47) eşitliğindeki paydanın azalmasıyla artacaktır ve artık varyansları aşağıdaki gibi hesaplanacaktır.

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-(p+1)} \tag{3.48}$$

Artık varyanslarının kestiriminde kullanılan (3.45) ve (3.47) nolu eşitliklerdeki farklılık çok düzeyli modellerde tam ve kısıtlı en çok olabilirlik kestirim yöntemleri arasındaki farklılığa da benzemektedir. Kısıtlı en çok olabilirlik yönteminde olduğu gibi (3.47) nolu eşitlik, artık varyanslarının (varyans bileşenlerinin) kestiriminden önce regresyon parametrelerinin kestirimini gerektirmektedir. Tam en çok olabilirlik yöntemi ise (3.45) nolu eşitliğe benzemektedir.

Varyans bileşenlerinin kısıtlı en çok olabilirlik kestirimleri örnek artıklarının olabilirliklerinin maksimize edilmesi ile hesaplanır. Uzun bir iteratif süreç de bulunmaktadır. İlk olarak γ sabit etkileri, genellikle en küçük kareler veya genelleştirilmiş en küçük kareler yöntemleri kullanılarak kestirilir. Daha sonra bilinen regresyon analizi gibi her durumdaki her bir kişi için bir artık kestirmek için $\hat{\gamma}$'lar kullanılır. Düzey1 ve düzey2 artıkları için bilinen varsayımlar altında (bağımsızlık, sabit varyanslılık ve normallik), gözlenen verilerin artıklar veya bilinmeyen varyans bileşenleri cinsinde olabilirlikleri yazılır. Kısıtlanmış olabilirliğin logaritması alınır ve sabit etkilerin (γ'ların) bilindiği varsayımı altında, varyans bileşenlerinin kısıtlı en çok olabilirlik kestirimlerinin maksimizasyonu sağlanmış olur.

Modelin farklı kısımlarına bu iki yöntem uygulanarak uyum iyiliği istatistikleri de hesaplanabilir. Tam en çok olabilirlik yöntemi ile tüm modele uyum incelenebilirken, kısıtlı en çok olabilirlik yöntemi ile sadece stokastik kısma (rasgele etkiler) uyum incelenebilmektedir. Bu durumda tam en çok olabilirlik yöntemi ile elde edilen uyum iyiliği istatistiği herhangi bir parametre için hipotezlerde kullanılabilirken kısıtlı en çok olabilirlik yöntemi ile elde edilen uyum iyiliği istatistikleri sadece varyans bileşenlerine ilişkin hipotezlerde kullanılabilmektedir.

En çok olabilirlik yönteminin uygulandığı hesaplama süreci, EM algoritmasını ve Fisher skorlama yöntemini de içermektedir. Goldstein (1986) çok düzeyli modellerde etkin parametre kestirimlerinin elde edilmesi için iteratif genelleştirilmiş küçük kareler (iterative generalized least squares) yöntemini geliştirmiştir. Goldstein, rasgele terimlerin çok değişkenli dağılım göstermesi sonucunda iteratif genelleştirilmiş küçük kareler kestirimlerinin tam bilgili en çok olabilirlik kestirimleriyle aynı olduğunu göstermiştir.

(Ko)Varyans bileşenlerinin kestirilmesinde kullanılan diğer bir yöntem Bayes kestirimidir. Seltzer (1990) bu uygulamayı hiyerarşik doğrusal modeller için geliştirmiştir. Bayes kestirim yönteminin olumsuz yanı, çok geniş veri kümesine ihtiyaç duyulması ve oluşacak modelin çok karmaşık olmasıdır.

3.5. Modellerin Sapma İstatistikleri Kullanılarak Karşılaştırılması

Oluşturulan birçok modelin uygun olanının seçilmesinde sapma istatistikleri kullanılabilir. Açıklanacak olan yöntemlerin kullanılmasının avantajları;

1. kullanılan istatistiklerin aynı anda birçok parametrenin incelenmesini sağlaması
2. I. Tür hatanın (H_0 doğru iken reddedilmesi olasılığı) hesaplanmasına imkan sağlaması

olarak özetlenebilir.

3.5.1. Sapma istatistiği (the deviance statistic)

Sapma istatistiği, en çok olabilirlik yönteminin prensipleri ile açıklanabilir. En çok olabilirlik kestirimleri, örneklemin tamamında bulunan verilerin bileşik olabilirliklerinin logaritmalarının alınması ile elde edilen log-olabilirlik fonksiyonun en büyüklenmesi ile elde edilmekteydi. Böylece modele ve varsayımlara bağlı olarak elde edilen log-olabilirlik fonksiyonu örnek verilerini ve tüm bilinmeyen parametreleri (γ'lar ve σ'lar) içermektedir.

Bilgisayar programları verileri ve parametrelerin kestirimlerini bir araya getiren log-olabilirlik fonksiyonunun büyüklüğünü hesaplayabilmektedirler. Programlar yardımıyla elde edilen bu sayıya örnek log-olabilirlik istatistiği (LL) denilmektedir. Genel anlamda, veri kümesine birçok model uygulanmışsa, log-olabilirlik istatistiği daha büyük olan model tercih edilmesi gereken model olacaktır.

Sapma istatistiği ise log-olabilirlik istatistiklerini iki model için karşılaştırır;

(1) elde bulunan hazır model
(2) örnek verilerine daha iyi şekilde uyan genel bir model (doygun model)

Bu durumda sapma bu iki model için farkın -2 katına eşit olacak ve

$$\text{Sapma} = -2(LL_{\text{eldeki model}} - LL_{\text{doygun model}}) \qquad (3.49)$$

şeklinde gösterilecektir.

Sapma istatistiği küçük olan modelin verilere, büyük sapma istatistiğine sahip modellerden daha fazla uyduğu ve tercih edilebilir olduğu söylenebilir.

Eşitlikte belirtildiği gibi sapma istatistiğinin hesaplanması için doygun modelin log-olabilirlik istatistiğine ihtiyaç duyulmaktadır. Çok düzeyli modellerde bu değerin elde edilmesi diğer modellere göre daha kolaydır. Çünkü doygun modeller kişi-dönem veri kümesinin yapısında olduğu gibi en iyi uyumun sağlanması için gerekli olabilecek kadar sayıda çok parametreyi içermektedir. Bu da olabilirlik fonksiyonunun alabileceği en büyük değerin 1 olacağı anlamına gelebilir. 1 değerinin logaritmasının 0 olmasından dolayı, doygun modelin de log-olabilirlik istatistiği 0 olacaktır. Bu nedenden (3.49) nolu ifadenin eşitlikten sonraki ikinci terimi çıkartılırsa, ifade;

$$\text{Sapma} = -2LL_{\text{eldeki model}} \tag{3.50}$$

şeklini alacaktır.

Standart normal teori varsayımı altında, veri kümesine uyum sağlayan bir çift iç içe geçmiş modelin sapma istatistikleri arasındaki fark bilinen bir dağılıma sahip olmaktadır. Bu da sapma istatistikleri kullanılarak ilgilenilen modellerle ilgili hipotez testlerinin kurulmasını sağlamaktadır.

Sapma istatistiklerinin iki modelde karşılaştırılması için;

(1) her bir kestirimin aynı verilerden elde edilmiş olması gerekmektedir

(2) modellerden biri diğerine iç içe geçmiş olmalıdır. Eğer iç içe geçmiş modeller söz konu değilse kişi-dönem veri kümesinde kayıp olan verilere sahip bir değişken modelden çıkartılarak model iç içe geçmiş hale getirilir.

Çok düzeyli modeller karşılaştırıldığında, sapma istatistiklerini kullanmadan önce dikkat edilmesi gereken üçüncü bir durum daha söz konusudur. Modelin sabit etkiler (γ'lar) ve varyans bileşenleri (σ'lar) olmak üzere iki çeşit parametre içermesinden dolayı tam ve değişken sayısı indirgenmiş model üç farklı yolla birbirlerinden farklılık gösterirler: sabit etkileri, varyans bileşenleri ve ya bu değerlerin birleşimi. Kestirim yöntemine bağlı olarak sadece bazı tür farklılıklar test edilebilir. Bu sınırlama kestirim yöntemlerinin uygulanmasında dikkate alınacak varsayımlardan kaynaklanmaktadır. Tam en çok olabilirlik yönteminde örnek verilerin olabilirlikleri en büyüklenirken kısıtlı en çok olabilirlik yönteminde örnek hatalarının olabilirlikleri en büyüklenmektedir. Sonuç olarak tam en çok olabilirlik yönteminde sapma istatistiği hem sabit ham de rasgele etkileri içeren ilk modele uyumu incelemekteyken kısıtlı en çok olabilirlik yönteminde sapma istatistiği modelin sadece stokastik kısmına uyumunu açıklayabilmektedir.

Tüm kısıtlamalarla birlikte kurulacak olan sıfır hipotezine bağlı olarak, tam ve değişken sayısı indirgenmiş modellerin sapma istatistik değerleri arasındaki fark asimptotik olarak, bağımsız kısıt sayısı serbestlik dereceli χ^2 dağılımına uygunluk göstermektedir.[32]

3.5.1.1. <u>sapma istatistiğine ilişkin hipotez testleri</u>

Elde edilen iki model karşılaştırılmadan önce;

(1) Modellerin aynı veri kümesinden oluşturulduğundan emin olunmalıdır

(2) İlk modelin sonraki modele iç içe geçmiş olması gerekmektedir.

(3) Daha sonra eklenmiş olan kısıtların sayısının bilinmesi gerekmektedir.

Sapma istatistiğine ilişkin hipotez testleri daha çok, her bir düzey2 alt modeline aynı anda bir veya daha çok kestiricinin eklenmesi durumunda modellerin karşılaştırılması için kullanılır. Ayrıca bu hipotezler sabit ve rasgele etkilerden kurulmuş olan iç içe geçmiş modellerin karşılaştırılmasında da kullanılmaktadır.

3.5.2. AIC ve BIC istatistikleri: iç içe geçmiş modellerin karşılaştırılmasında bilgi kriterinin kullanılması

Modellerin uyum iyiliği ölçümlerinin karşılaştırılmasında iki ad hoc kriteri (ad hoc criteria) kullanılabilmektedir: Akaike bilgi kriteri (AIC) ve Bayesyen bilgi kriteri (BIC). Sapma istatistiğinin olduğu gibi, her iki kriter de log-olabilirlik istatistiğine dayanmaktadır. Modellerin karşılaştırılması için çoğu zaman log-olabilirlik yönteminin kendisinin kullanılmasının yerine bu kriterlerin kullanılması tercih edilir. AIC, modeldeki parametre sayısına bağlı olan bir yöntemdir. Bu nedenden parametre sayısının arttırılması log-olabilirlik istatistik değerini arttıracak ve buna bağlı olarak da sapma istatistik değeri düşecektir. BIC ise sadece modeldeki parametre sayısına bağlı değil, aynı zamanda örnek büyüklüğüne de bağlıdır. Her iki kriter için elde edilen

[32] A. g. e., Singer, p:119.

sonuç -2 ile çarpılırsa, bilgi kriterinin değeri sapma istatistiğine neredeyse eşit duruma gelir. Tam en çok olabilirlik kestirim yönteminde hem sabit etkiler hem de varyans bileşenleri bilgi kriterinin hesaplanmasında kullanılırken, kısıtlı en çok olabilirlik kestirim yönteminde sadece varyans bileşenleri parametrelerinin kullanılması yeterli olmaktadır.

Bilgi kriteri genel anlamda;

Bilgi kriteri = -2 [LL – (ölçek faktörü)(modeldeki parametre sayısı)]
= Sapma + 2 (ölçek faktörü)(modeldeki parametre sayısı)]

olarak hesaplanır. AIC hesaplanırken ölçek faktörü 1 değerini alırken, BIC hesaplanırken bu değer örnek büyüklüğünün logaritmik değerinin yarısına eşit olacaktır. Her iki kriter için de, bilgi kriter değeri küçük olan modelin daha tercih edilebilir olduğunu ve verilere daha fazla uygunluk gösterdiği söylenebilir.

3.5.3. Sabit etkilere ilişkin testlerde wald istatistiğinin kullanılması

Hipotezlerin testlerinde sapma istatistikleri sonuca ulaşmak için kullanılan tek yöntem değildir. Bu bölümde bu amaçla kullanılabilecek olan ve parametre kestirimlerinin kendi standart sapmalarına bölünmesiyle elde edilen Wald istatistiği açıklanacaktır. Wald istatistiğinin en büyük avantajı, kullanılan kestirim yönteminin ne olduğu önemli olmaksızın karma hipotezlerdeki çoklu etkilerin bir arada test edilmesine olanak sağlamasıdır.

Karma sıfır hipotezi;

$$H_0 : \gamma_{00} = 0 \quad \text{ve} \quad \gamma_{10} = 0 \tag{3.51}$$

şeklinde gösterilecektir. Bu bileşik hipotez, her bir parametre için ayrı bağımsız durumların karşılaştırılması yerine tüm anakütle hakkındaki karma durumu gösterecektir.

Belirtilen sıfır hipotezi genel bilinen form olan “genel doğrusal hipotez” şeklinde de yazılması sağlanabilir. Bu gösterim şeklinde her bir modelin sabit etkileri belirlenmiş bir katsayıyla çarpılır (bir tamsayı, ondalık bir sayı, bir kesir veya sıfır olabilir) ve bu değerlerin toplamı başka bir sabite (genellikle sıfır) eşitlenir. Bu durumda (3.51) nolu eşitlik bu formda;

$$H_0 : \begin{array}{l} 1\gamma_{00} + 0\gamma_{01} + 0\gamma_{02} + 0\gamma_{10} + 0\gamma_{12} = 0 \\ 0\gamma_{00} + 0\gamma_{01} + 0\gamma_{02} + 1\gamma_{10} + 0\gamma_{12} = 0 \end{array} \qquad (3.52)$$

şeklinde yazılabilir.

Birçok yazılımda genel doğrusal hipotezlerin matris notasyonunda yazılması gerekmektedir. Bu durum da hipotezin iki farklı kısma ayrılmasını gerektirmektedir: (1) sabit değerlerin çarpım matrisi, (2) parametre vektörü. Sabitlerin çarpımlarının matrisi, **C** ile gösterilir ve (3.51) nolu eşitlik dikkate alınırsa;

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad (3.53)$$

olarak gösterilebilir. Sabit etkilerin vektörü ise $\boldsymbol{\gamma}$ ile ifade edilir ve

$$\gamma = \begin{bmatrix} \gamma_{00} & \gamma_{01} & \gamma_{02} & \gamma_{10} & \gamma_{12} \end{bmatrix} \qquad (3.54)$$

olarak gösterilebilir. Bu durumda **C** matrisi ve $\boldsymbol{\gamma}$ vektörünün traspozesi ile oluşturulacak olan genel doğrusal hipotez;

$$H_0 : \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{02} \\ \gamma_{10} \\ \gamma_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.55}$$

olarak yazılacaktır. Eğer daha genel anlamda gösterilirse, H_0 : **C**$\boldsymbol{\gamma}$’ = 0 olarak da ifade edilebilir. Verilmiş bir model için **C**’nin değerleri hipotezden hipoteze değişiklik gösterirken $\boldsymbol{\gamma}$’nin değerlerin daima aynı kalır. [33]

[33] A. g. e.Singer, 132.

DÖRDÜNCÜ BÖLÜM
BEBEKLERİN DİL GELİŞİMLERİNE İLİŞKİN UYGULAMA

Bu araştırmanın konusunu İstanbul'da 2002 yılında doğmuş olan 40 bebeğin, doğumlarını takip eden 3 yıl içerisindeki gelişimlerinin incelenmesi oluşturmaktadır. Gelişimin incelenmesinde, bebeklerin yaşlar itibariyle normal olarak gerçekleştirmeleri gereken fiziksel veya zihinsel aktivitelerin dünya standartlarına göre belirlendiği ve Türkiye standartlarına göre uyarlandığı Denver II Gelişim Tarama Testi kullanılmıştır.

Araştırmada için seçilen örneklem Mature Eğitim ve Aile Danışmanlığı Merkezi'nden hizmet alan ve doğduklarından beri gelişimleri aylar bazında incelenen 40 bebek ve ailesinden oluşmaktadır. Bu araştırma için örneklemin 12, 24 ve 36 aylık dönemlerdeki gelişimleri ve bu dönemlerde gerçekleştirmeleri gereken fiziksel ve zihinsel aktiviteleri dikkate alınmıştır. Bu dönemlerde ilgilenilen maddeler Ek-1'de gösterilmiştir.

Kullanılan Denver II Gelişimsel Tarama Testi'ne ilişkin bilgiler bir sonraki bölümde açıklanmıştır.

4.1. Denver II Gelişimsel Tarama Testi

Küçük çocukların gelişimindeki sapmaları rutin fizik muayene sırasında anlamak güçtür. Gelişimsel bozukluğun varlığı genellikle çocukta yürüyememe, konuşamama gibi belirgin sorunlar ortaya çıktığında ya da okulda başarısız olduğunda fark edilir. Denver Gelişimsel Tarama Testi (DGTT) küçük çocuklarda bulunabilecek gelişimsel sorunları yakalamada sağlık personeline yardımcı olabilmesi amacıyla ilk kez 1967 yılında yayınlanmış, ilk yayınlanışından sonra geniş kullanım alanı bulmuştur. Birçok ülkede toplumlara göre uyarlanarak ve standardize edilerek 50 milyondan fazla çocuğun taranmasında kullanılmıştır. Türkiye'de de 1980 yılında

Hacettepe Üniversitesi Tıp Fakültesi Çocuk Nörolojisi Bölümü tarafından standardize edilerek ülke çapında kullanıma sunulmuştur.

DGTT, yaygın kullanımı sonucunda elde edilen deneyimlerin ışığında 1990 yılında Frankenburg ve Dodds tarafından yeniden gözden geçirilmiştir. DGTT ile deneyimlerden ile ortaya çıkan dört ana sorun: daha fazla dil maddesine gerek duyulması, 1967'deki normların 1990'da geçerliliği, bazı maddelerin uygulanmasında ve yorumlanmasındaki güçlükler ve bazı maddelerin bazı çocuk grupları için uygun olamayabileceği idi. Bu tarz sorunların ortadan kaldırılması için Denver II oluşturulmuştur.

Denver II, 0-6 yaş arasındaki görünürde sağlıklı olan çocuklara uygulanmak üzere düzenlenmiştir. Çocuğun yaşına uygun birtakım becerilerini değerlendiren bu test çocukları gelişimsel problemler açısından taramada, kuşkulu durumları objektif bir ölçütle doğrulamada ve gelişimsel açıdan risk altındaki çocukları (örneğin perinatal sorunlar geçirmiş bebekleri) izlemede değerlidir.

Denver II bir zekâ testi değildir. Gelecekteki zihinsel veya uyumsal yeteneği tahmin etmede kullanılmaz. Bunun yerine, çocuğun birtakım işlevlerdeki becerisini yaşıtları ile karşılaştırır. Öğrenme güçlüğü, konuşma bozukluğu, duygusal bozukluk gibi tanılar vermek üzere yapılmamıştır ve hiçbir zaman fizik muayene veya tanısal değerlendirme yerine kullanılmamalıdır.
Denver II, aşağıdaki gelişimsel alanları taramak üzere test formu üzerinde dört bölümde toplanmış 116 maddeden oluşmaktadır:

1. Kişisel-Sosyal: insanlarla anlaşma, kendi bireysel gereksinimlerini karşılayabilme.
2. İnce motor: el-göz koordinasyonu, kişilik cisimleri kullanabilme, problem çözme.
3. Dil: işitme, anlama ve dili kullanma.

4. Kaba motor: oturma, yürüme, zıplama ve genel olarak büyük kasların hareketi. [34]

4.2. Gözlemler Arası İlişkinin İncelenmesi ve Tanımlayıcı İstatistikler

Dalgalar boyunca gözlemler arasındaki ilişkinin incelenmesi, incelenen gözlemlerin zaman içindeki gelişim ve değişimlerinin yorumlanmasına yardımcı olacaktır. Korelasyonlar yardımıyla incelenen değişkenlerin zaman içerisinde ne kadar değiştikleri büyüklük olarak belirlenemez ancak değişimin yönü hakkında fikir sağlanabilir. Tablo 4.1, 12, 24 ve 36 aylık zaman dönemlerinde bebeklerin kişisel sosyal, ince motor, dil ve kaba motor gelişimleri arasındaki korelasyonları göstermektedir.

Korelasyonlar tablosu dikkate alındığında; 12 aylık bebeğin kişisel sosyal gelişimi, 24 ve 36. aylarda pozitif korelasyona sahip olmaktadır. Bu durum da bebeğin ilerleyen dönemlerde kişisel ve sosyal açıdan pozitif yönde geliştiğini göstermektedir. 12 aylık bebeğin ince motor gelişimi ile 24 aylık bebeğin ince motor gelişimleri arasındaki korelasyon pozitif yönlü olduğundan bebekler zaman içerisinde ince motor değişkenleri açıdan pozitif yönde gelişmişlerdir denilebilmektedir. Ancak 36 ay döneminde korelasyonun hesaplanamadığı görülmüştür. Bunun sebebinin bu dönemde bütün bebeklerin sorumlu oldukları maddelerin tamamından geçmiş olmaları ve bu döneme ait verilerin değişken olma özelliği taşımaması olduğu söylenebilir. 12 aylık bebeğin dil gelişimi 24 aylık döneme gelindiğinde pozitif yönde geliştiği ancak 36 aylıkken gelişiminin 12 aylıktaki durumuna göre negatif yönde olduğu görülmektedir. Bebeğin gelişiminin 36 aylıkken, 12 aylığa kıyasla yavaşladığı söylenebilir. Son olarak 12 aylık bebeğin kaba motor gelişimi zaman geçtikçe ve bebek büyüdükçe pozitif yönde gelişmektedir denilebilmektedir.

---

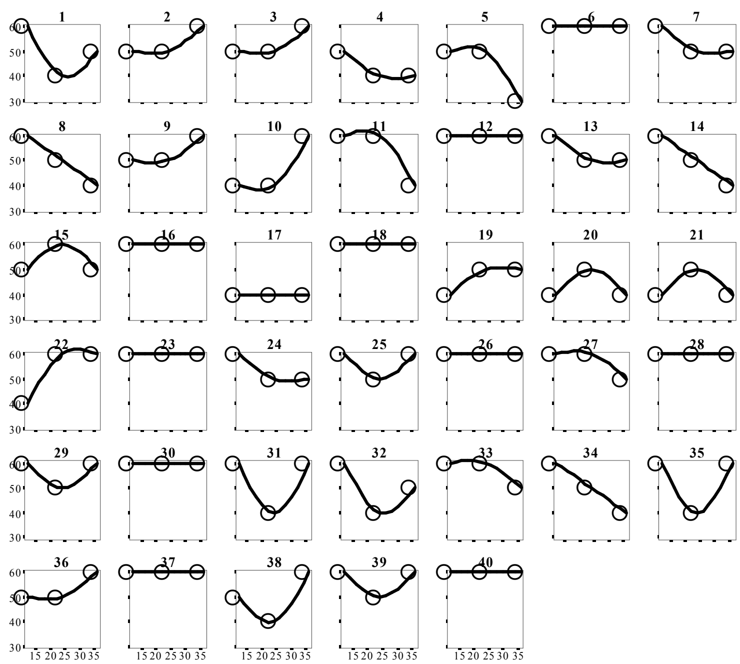
[34] B. Anlar, K. Yalaz, “Denver II Gelişimsel Tarama Testi, Türk Çocuklarına Uyarlanması ve Standardizasyonu”, Hacettepe Üniversitesi Yayınları, Ankara, 2000, s:3.

Korelasyonlar

		12 aylik bebegin kisisel sosyal gelisimi	12 aylik bebegin ince motor gelisimi	12 aylik bebegin dil gelisimi	12 aylik bebegin kaba motor gelisimi	24 aylik bebegin kisisel sosyal gelisimi	24 aylik bebegin ince motor gelisimi	24 aylik bebegin dil gelisimi	24 aylik bebegin kaba motor gelisimi	36 aylik bebegin kisisel sosyal gelisimi	36 aylik bebegin dil gelisimi	36 aylik bebegin kaba motor gelisimi
12 aylik bebegin kisisel sosyal gelisimi	**Pearson Korelasyon**	1,000	,379	-,090	,439	,297	,242	-,127	,027	,232	,275	,425
	Sig. (2-yönlü)	,	,016	,580	,005	,063	,132	,436	,870	,150	,086	,006
	N	40	40	40	40	40	40	40	40	40	40	40
12 aylik bebegin ince motor gelisimi	**Pearson Korelasyon**	,379	1,000	,058	,158	,198	,211	,420	-,197	,372	,221	,182
	Sig. (2-yönlü)	,016	,	,723	,329	,221	,191	,007	,224	,018	,171	,262
	N	40	40	40	40	40	40	40	40	40	40	40
12 aylik bebegin dil gelisimi	**Pearson Korelasyon**	-,090	,058	1,000	-,043	,426	,014	,418	-,116	-,061	-,055	-,075
	Sig. (2-yönlü)	,580	,723	,	,790	,006	,933	,007	,475	,707	,738	,646
	N	40	40	40	40	40	40	40	40	40	40	40
12 aylik bebegin kaba motor gelisimi	**Pearson Korelasyon**	,439	,158	-,043	1,000	,190	,071	,025	,010	,149	-,033	,125
	Sig. (2-yönlü)	,005	,329	,790	,	,240	,663	,876	,953	,358	,839	,441
	N	40	40	40	40	40	40	40	40	40	40	40
24 aylik bebegin kisisel sosyal gelisimi	**Pearson Korelasyon**	,297	,198	,426	,190	1,000	,144	,244	,103	,229	,256	,066
	Sig. (2-yönlü)	,063	,221	,006	,240	,	,374	,129	,526	,155	,111	,684
	N	40	40	40	40	40	40	40	40	40	40	40
24 aylik bebegin ince motor gelisimi	**Pearson Korelasyon**	,242	,211	,014	,071	,144	1,000	-,016	,043	,159	-,094	,338
	Sig. (2-yönlü)	,132	,191	,933	,663	,374	,	,922	,794	,327	,562	,033
	N	40	40	40	40	40	40	40	40	40	40	40
24 aylik bebegin dil gelisimi	**Pearson Korelasyon**	-,127	,420	,418	,025	,244	-,016	1,000	-,136	,251	,192	-,088
	Sig. (2-yönlü)	,436	,007	,007	,876	,129	,922	,	,402	,118	,236	,591
	N	40	40	40	40	40	40	40	40	40	40	40
24 aylik bebegin kaba motor gelisimi	**Pearson Korelasyon**	,027	-,197	-,116	,010	,103	,043	-,136	1,000	,241	,153	,155
	Sig. (2-yönlü)	,870	,224	,475	,953	,526	,794	,402	,	,135	,345	,339
	N	40	40	40	40	40	40	40	40	40	40	40
36 aylik bebegin kisisel sosyal gelisimi	**Pearson Korelasyon**	,232	,372	-,061	,149	,229	,159	,251	,241	1,000	,304	,458
	Sig. (2-yönlü)	,150	,018	,707	,358	,155	,327	,118	,135	,	,056	,003
	N	40	40	40	40	40	40	40	40	40	40	40
36 aylik bebegin dil gelisimi	**Pearson Korelasyon**	,275	,221	-,055	-,033	,256	-,094	,192	,153	,304	1,000	-,057
	Sig. (2-yönlü)	,086	,171	,738	,839	,111	,562	,236	,345	,056	,	,726
	N	40	40	40	40	40	40	40	40	40	40	40
36 aylik bebegin kaba motor gelisimi	**Pearson Korelasyon**	,425	,182	-,075	,125	,066	,338	-,088	,155	,458	-,057	1,000
	Sig. (2-yönlü)	,006	,262	,646	,441	,684	,033	,591	,339	,003	,726	,
	N	40	40	40	40	40	40	40	40	40	40	40

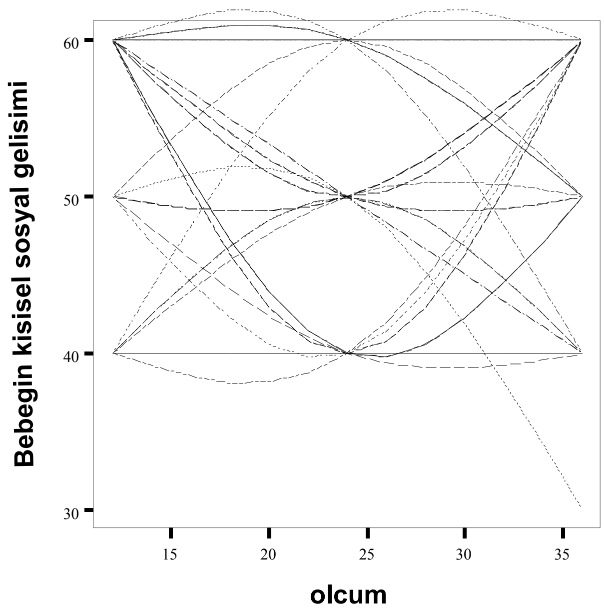
Tablo 4.1. Değişkenlerin Dönemler Arası Korelasyonları

Düzleştirme eğrileri kullanılarak incelenen bebeklerin değişkenler için zaman içindeki gelişimleri de gözlenebilir. Bunun için gözlemlerin zamana bağlı değerleri dikkate alınarak kişisel sosyal gelişim puanlarına göre düzleştirme eğrileri çizilmiş ve şekil 4.1‘de gösterilmiştir.



Şekil 4.1. Kişisel Sosyal Gelişime Ait Düzleştirme Eğrileri

Düzleştirme eğrileri yardımıyla 6 ve 40 numaralı bebeklerin kişisel sosyal gelişimlerinin zaman içerisinde sabit kaldığı, 14 ve 34 numaralı bebeklerin zaman içerisinde gelişimlerinin yavaşladığı, 31 ve 35 numaralı bebeklerin gelişiminin ise 24 aylık döneme gelindiğinde yavaşladığı ancak 36 aylık döneme gelindiğinde arttığı görülmektedir. Tüm bebeklerin kişisel sosyal gelişimlerinin bir arada izlenebildiği düzleştirme eğrisi şekil 4.2’de gösterilmiştir.



Şekil 4.2. Tüm Bebeklere Ait Kişisel Sosyal Gelişim Eğrileri

Bu grafik, 40 bebeğin kişisel sosyal gelişimlerinin düzleştirme eğrilerini bir arada göstermektedir. Buna göre bebeklerin kişisel sosyal gelişimlerinin dönemler itibariyle hızlanması veya yavaşlaması yönünde genel bir eğilimin olmadığını ve her bir bebeğin gelişiminin kendine özgü bir gelişim şekli olduğunu göstermektedir.

İnce motor, dil ve kaba motor değişkenlerinde, her bir bebek için oluşturulmuş düzleştirme eğrileri ve tüm bebeklerin gelişimlerinin gözlendiği grafikler ek-2 ve ek-3'te verilmiştir. Bebeklerin ince motor gelişimleri incelendiğinde, 2 yaş (24 aylık) döneminde gelişim hızlarında bir yavaşlama olduğu ancak 3 yaş (36 aylık) döneminde gelişim hızlarının arttığı görülmektedir. Bebeklerin dil gelişimlerin de ise 3 yıllık süreç içerisinde dikkati çeken herhangi bir artma veya azalma ivmesi görülmemektedir. Hatta neredeyse tüm bebeklerin dil gelişimlerinin eksiksiz ve zamanında gerçekleştiği söylenebilmektedir. Bebeklerin kaba motor gelişimleri dikkate alındığında ise 2 yaş (24 aylık) döneminin bir kırılma noktası olduğu ve bu dönemde bebeklerin bir kısmında gelişim hızının arttığı, bir kısmında yavaşladığı ve çok az kısmında da sabit kaldığı görülmektedir.

Bireylerin gelişim süreçlerinin açıklanabilmesi için her bir bireyin verilerinin ayrı ayrı parametrik modellere yerleştirilebilecekleri daha önceden açıklanmıştı. Ancak incelenen dönem sayısının az ve hesaplanacak olan modelde serbestlik derecesinin 1 olmasından dolayı, bebeklere ilişkin kurulacak modeller için elde edilecek olan F ve p (sig.) değerleri, modellerin anlamsız olduğunu gösterecektir. Bu nedenle uygulamaya bireylerin kendi içlerindeki gelişimlerini gösterecek modeller dahil edilmemiştir.

4.3. Dikey Kesit Verilere Çok Düzeyli Analizin Uygulanması

Bu uygulamada amaç Denver II Gelişimsel Tarama Testi ile elde edilmiş olan ve bebeğin gelişimini açıklayan bazı değişkenlerin birbirleri ile arasındaki ilişkilerinin ve gelişimi etkilediğine inanılan diğer değişkenlerin etkilerinin incelenmesidir.

40 bebeğe ait gelişimlerini gösteren dikey kesit veriler kişi-dönem veri kümesi olarak düzenlenmişti. Birden çok değişkenin ve bu değişkenleri etkileyen alt değişkenlerin ilişkilerinin incelenmesinde çok düzeyli modelleme yöntemi kullanılmıştır. Uygulama boyunca çok düzeyli modellere ilişkin analizlerin yapılmasında HLM 6.0 For Windows[35] paket programı kullanılmıştır.

Dönemlere ilişkin veriler incelendiğinde, bebeklerin kendi aralarındaki ve birbirleri arasındaki ilişkiyi ölçmeye yardımcı olacak iki düzeyli model kurulmasına karar verilmiştir. Model, bebeklerin bu dönemlerdeki dil gelişim skorları ile elde edilecek ve diğer değişkenlerle açıklanmaya çalışılacaktır.

---

[35] HLM 6 Hierarchical Linear and Nonlinear Modeling, Stephen Raudenbush, Tony Bryk, Scientific Software International, Inc., 2000.

4.3.1. Dil gelişiminin açıklanmasında tüm değişkenlerin kullanıldığı model

Bebeklerin 12, 24 ve 36 aylık dönemlerdeki dil gelişimleri, ek gıda kullanılmaksızın sadece anne sütünün alındığı süre, bebeğin bu üç dönem içerisindeki kişisel sosyal gelişimi ve ince motor gelişimi ile açıklanacaktır. Bunun yanı sıra bebeğin ek gıdaya geçmeden anne sütü aldığı süre; bebeğin cinsiyeti, annenin bebeğini doğurduğu zamanki eğitim durumu ve bebeğin doğduğu andaki ağırlığı ile açıklanacaktır. Bunların yanı sıra bebeğin kişisel sosyal gelişimi; annenin eğitim durumu ve bebeğini doğurduğu zamanki yaşı, bebeğin ince motor gelişimi ise bebeğin cinsiyeti, annesinin eğitim durumu, bebeğin doğduğu zamanki boyu ve ağırlığı değişkenleri ile açıklanacaktır. Analiz boyunca dil gelişim düzeyini açıklamak üzere kullanılacak ve düzey1, düzey2 değişkenleri olarak isimlendirilecek değişkenler tablo 4.2'de gösterilmiştir.

Tablo 4.2. Dil Gelişimini Açıklamakta Kullanılacak Olan Değişkenler

Düzey1 Değişkenleri	Düzey2 Değişkenleri
Bebeğin ek gıda almaksızın anne sütü alma süresi (SUTSURE)	• Bebeğin cinsiyeti (CINSIYET) • Annenin eğitim durumu (ANNEEG) • Bebeğin doğduğu zamanki ağırlığı (AGIRLIK)
Bebeğin kişisel sosyal gelişim skorları (KISSOS)	• Annenin eğitim durumu (ANNEEG) • Annenin doğum yaptığı zamanda yaşı (ANNEYAS)
Bebeğin ince motor gelişim skorları (INCEMO)	• Bebeğin cinsiyeti (CINSIYET) • Annenin eğitim durumu (ANNEEG) • Bebeğin doğduğu zamanki boyu (BOY) • Bebeğin doğduğu zamanki ağırlığı (AGIRLIK)

Bu değişkenler kullanılarak elde edilecek iki düzeyli model aşağıdaki gibi gösterilecektir.

Düzey1 modeli:

$DIL_{ti} = \pi_{0i} + \pi_{1i} * SUTSURE + \pi_{2i} * KISSOS + \pi_{3i} * INCEMO + e_{ti}$

Düzey2 model:

$$\pi_{0i} = r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} * (CINSIYET) + \beta_{12} * (ANNEEG) + \beta_{13} * (AGIRLIK)$$

$$\pi_{2i} = \beta_{20} + \beta_{21} * (ANNEEG) + \beta_{22} * (ANNEYAS)$$

$$\pi_{3i} = \beta_{30} + \beta_{31} * (CINSIYET) + \beta_{32} * (ANNEEG) + \beta_{33} * (BOY) + \beta_{34}(AGIRLIK)$$

Modele ilişkin sonuçlara kısıtlı en çok olabilirlik kestirim yöntemi kullanılarak 35 iteratif işlemin sonucunda ulaşılmıştır. Düzey1 ve düzey2 modelleri için hesaplanmış olan varyanslar aşağıda gösterilmiştir.

Sigma Kare (Düzey1 Varyansı - v) = 15,99379

Tau (Düzey2 Varyansı - τ) = 1453,73876

Elde edilen bu varyans değerleri modelin güvenilirliğinin (reliability) hesaplanmasında kullanılmaktadır. Bir model için güvenilirlik, parametre varyansının toplam varyansa oranı olarak açıklanmaktadır ve aşağıdaki gibi hesaplanır.

$$\text{güvenilirlik } (\hat{\beta}) = \frac{\tau}{\tau + v} \qquad (4.1)$$

Güvenilirliği ifade eden oranın 1'e yakın olması gerçek değerlerle gözlenen değerler arasında tam bir uyum olduğunu gösterir. Bu nedenden dolayı güvenilirlik değeri 1'e yakın olma durumuna göre yorumlanır[36].

Bebeklerin dil gelişimlerinin açıklanması için kurulmuş olan modelde güvenilirlik,

$$\text{güvenilirlik } (\hat{\beta}) = \frac{1453{,}73876}{1453{,}73876 + 15{,}99379} = 0{,}989$$

olarak hesaplanmıştır. Bu oranın 1'e çok yakın olması modelin hesaplanmasında kullanılan gerçek değerlerle, gözlenen değerler arasında uyumun oldukça yüksek olduğunu göstermektedir.

Model için hesaplanmış olan katsayılar ve anlamlılıkları da aşağıda belirtilmiştir.

Tablo 4.3. Tam Modelde Kullanılacak Düzey2 Modele Ait Katsayılar

Sabit Etkiler	Katsayılar	Standart Hatalar	T- Oranı	Serbestlik Derecesi	p-değeri
SUTSURE slope, P1					
INTRCPT2, B10	2.592552	11.180434	0.232	108	0.817
CINSIYET, B11	0.595277	3.223445	0.185	108	0.854
ANNEEG, B12	-4.720588	3.620386	-1.304	108	0.195
AGIRLIK, B13	8.413330	3.041809	2.766	108	0.007
KISSOS slope, P2					
INTRCPT2, B20	0.140573	0.473149	0.297	108	0.767
ANNEEG, B21	-0.158012	0.146187	-1.081	108	0.283
ANNEYAS, B22	0.006927	0.015245	0.454	108	0.650
INCEMO slope, P3					
INTRCPT2, B30	-0.243515	1.169021	-0.208	108	0.836
CINSIYET, B31	-0.011702	0.158399	-0.074	108	0.942
ANNEEG, B32	0.415969	0.176862	2.352	108	0.021
BOY, B33	0.059822	0.032032	1.868	108	0.064
AGIRLIK, B34	-0.950085	0.218351	-4.351	108	0.000
Bağımlı Değişken: DIL					

[36] S. W. Raudenbush, A. S. Bryk, Hierarchical Linear Models, Applications and Data Analysis Methods, Second Edition, Sage publications, London, 2002, p:46.

Model için katsayılar incelendiğinde süt süresi için ağırlık (p=0,007<0,05), ince motor gelişimi değişkeni için de anne eğitimi (p=0,021<0,05) ve ağırlık (p=0.000<0,05) değişkenleri dışında kalan tüm değişkenler anlamsız çıkmıştır.

Uygulamalarda sonuçların her bir düzeydeki rasgele etkilerin dağılımları hakkındaki varsayımlara uygun olması gerekmektedir. Rasgele etkilerin varsayımlarında oluşacak bir bozulum, sabit etkiler için elde edilmiş sonuçların yanlış çıkmasına neden olacaktır. Rasgele etkilere ilişkin varsayımların, sabit etkiler ile ilgili sonuçların üzerindeki duyarlılığını kontrol edebilmek için sağlam (robust) standart hatalar hesaplanmaktadır. Dikey kesit çalışmalarda bu durum ilk kez Liang ve Zeger (1986) tarafından incelenmiş ve sağlam standart hatalarının, yüksek düzeylerdeki birim sayılarının fazla olduğu zamanlarda kullanılmasının uygun olduğunu ispat etmişlerdir.

İki düzeyli bir model için sağlam standart hatalarının hesaplanması için izlenecek yol aşağıda belirtilmektedir:

1. $Y_{ij} = \gamma_0 + \gamma_1 X_{ij} + e_{ij}$ şeklinde tanımlanan bir model için matris notasyonu;

$$\begin{pmatrix} Y_{1j} \\ \vdots \\ Y_{n_j j} \end{pmatrix} = \begin{pmatrix} 1 & X_{1j} \\ \vdots & \vdots \\ 1 & X_{n_j j} \end{pmatrix} \begin{pmatrix} \gamma_0 \\ \gamma_1 \end{pmatrix} + \begin{pmatrix} e_{ij} \\ \vdots \\ e_{n_j j} \end{pmatrix} \tag{4.2}$$

şeklinde olup, bu notasyon doğrultusunda model genel olarak,

$$\mathbf{Y_j} = \mathbf{X_j}\gamma + \mathbf{e_j} \tag{4.3}$$

şeklinde tanımlanır.

2. Normallik varsayımları altında $\mathbf{e_j}$~N(0, $\sigma^2\mathbf{I_{nj}}$) en küçük kareler kestiricisi ;

$$\hat{\gamma} = \left(\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{X_j^T X_j} \right)^{-1} \sum_{j=1}^{J} \mathbf{X_j^T Y_j} \tag{4.4}$$

olarak hesaplanacaktır.

3. Bu değerlere bağlı olarak sağlam varyans kestirimi;

$$V\hat{a}r(\hat{\gamma}) = \left(\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{X_j^T X_j} \right)^{-1} \sum_{j=1}^{J} \mathbf{X_j^T} (\mathbf{Y_j} - \mathbf{X_j}\hat{\gamma}) \times (\mathbf{Y_j} - \mathbf{X_j}\hat{\gamma})^T \mathbf{X_j} \left(\sum_{j=1}^{J} \mathbf{X_j^T X_j} \right)^{-1} \tag{4.5}$$

eşitliği yardımıyla hesaplanacaktır. Buradan hesaplanacak olan standart hata değerleri, e_{ij} hata terimleri normal dağılmadığı zaman güven aralıkları ve hipotez testlerinin daha doğru sonuçlar vermesine yardımcı olacaktır. Sağlam standart hatalar, $\hat{\gamma}$ değerlerinin sıradan en küçük kareler kestirim yöntemiyle hesaplanması durumunda da kullanılabilmektedir[37].

Bu açıklamalar doğrultusunda düzey1 hata terimleri e_{ij}'ler için geçerli olan normal dağılma varsayımı incelenmiştir. Bunu test edebilmek için kullanılan Kolmogorov-Smirnov teste ilişkin sonuç tablo 4.4'de gösterilmiştir.

Testin sonucunda düzey1 hata terimlerinin normal dağılmadığı (p=0,006<0,05) belirlenmiştir. Bu nedenle modelden daha etkin sonuçlar alabilmek için sağlam standart hatalarının hesaplanması ve kullanılması uygun olacaktır.

---

[37] A. g. e., Raudenbush, p:276.

Tablo 4.4. Tam Modelin Hata Terimlerine Ait Kolmogorov-Smirnov Test Sonucu

Kolmogorov-Smirnov Testi

		Düzey1 hata terimleri
N		120
Normal Parametreleri	**Ortalama**	,04249
	Std. Sapma	3,132061
Farklar	**Mutlak**	,156
	Pozitif	,150
	Negatif	-,156
Kolmogorov-Smirnov Z		1,705
Asymp. Sig. (2-yönlü)		,006

Yapılan açıklamalara göre daha önceden elde edilmiş olan katsayılar ve bu katsayılar için hesaplanmış olan sağlam standart hataları aşağıdaki tabloda verilmektedir.

Tablo 4.5. Tam Modelde Kullanılacak Düzey2 Modele Ait Katsayılar (Sağlam Std. Hata İle)

Robust Standart Hataları ile

Sabit Etkiler	Katsayılar	Standart Hatalar	T-Oranı	Serbestlik Derecesi	p-değeri
SUTSURE slope, P1					
INTRCPT2, B10	2.592552	8.154804	0.318	108	0.751
CINSIYET, B11	0.595277	2.714838	0.219	108	0.827
ANNEEG, B12	-4.720588	4.411046	-1.070	108	0.287
AGIRLIK, B13	8.413330	2.426162	3.468	108	0.001
KISSOS slope, P2					
INTRCPT2, B20	0.140573	0.404690	0.347	108	0.729
ANNEEG, B21	-0.158012	0.124477	-1.269	108	0.207
ANNEYAS, B22	0.006927	0.012470	0.555	108	0.579
INCEMO slope, P3					
INTRCPT2, B30	-0.243515	0.940086	-0.259	108	0.796
CINSIYET, B31	-0.011702	0.205820	-0.057	108	0.955
ANNEEG, B32	0.415969	0.268240	1.551	108	0.124
BOY, B33	0.059822	0.029679	2.016	108	0.046
AGIRLIK, B34	-0.950085	0.272969	-3.481	108	0.001

Sağlam standart hataları yardımıyla elde edilen ve katsayıların anlamlılığını ölçmede bir kriter olarak kullanılan t istatistiği ve p-değerleri, standart hatalar yardımıyla hesaplanan değerlerden farklılık göstermiştir. Sağlam standart hatalarıyla incelenen modelde, bebeğin doğduğu zamanki ağırlığının, bebeğin süt emme süresindeki marjinal katkısının, diğer değişkenlere göre daha fazla olduğu söylenebilir. Bu durumda bebeğin ek gıdaya geçmeden anne sütü alma süresi, doğduğu zamanki ağırlığı ile açıklanabilmektedir. Bebeğin ince motor gelişimi incelendiğinde ise, bebeğin doğduğu zamanki ağırlığı ve daha sonra bebeğin doğduğu zamanki boy uzunluğunun değişkene ve modele marjinal katkısı diğer değişkenlerden daha fazladır denilebilir. Bu da bebeğin ince motor gelişimi olarak adlandırılan basit parmak ve vücut hareketlerinin gelişiminin, bebeğin doğduğu zamanki fiziksel özellikleri ile açıklanabildiğini göstermektedir.

Değişkenlerin anlamlı olanlarının bağımlı değişkeni açıklayabildiğini söyleyebilmek için modelin anlamlı çıkması gerekmektedir. Çok düzeyli dikey kesit veri analizlerinde de modellerin anlamlılığının belirlenebilmesi için, regresyon modellerinde olduğu gibi ANOVA kullanılmaktadır. Ancak bu analizdeki en büyük fark, test istatistiği olarak F değerinin değil, χ^2 değerinin kullanılmasıdır. Çünkü burada birbiri ardı sıra gelen birçok dönem birbirleri ile karşılaştırılmak üzere incelenmektedir.

Bu analizde hesaplanacak χ^2 değeri;

$$\chi^2 = \frac{\sum n_j (\overline{Y}_{.j} - \hat{\gamma}_{00})^2}{\hat{\sigma}^2} \tag{4.6}$$

eşitliğinden elde edilecektir. Bu değer (j-1) serbestlik dereceli (j= birim sayısı) tablo değeriyle karşılaştırılır ve modelin anlamlılığı hakkında karar verilir.

Dil gelişimini açıklamak için kurulan modelde, modelin anlamlılığını belirlemek için hesaplanan değerler aşağıdaki tabloda belirtildiği gibidir.

Tablo 4.6. Tam Model İçin Ki-Kare Değeri

Rasgele Etkiler		Standart Sapma	Varyans	s.d	Ki - Kare	p-değeri
INTRCPT1,	R0	38.12793	1453.73876	40	9613.22346	0.000
Düzey-1	E	3.99922	15.99379			

Hesaplanan ki-kare değerine (p=0,000<0,05) göre modelin anlamlı olduğu söylenebilir. Bu durumda bebeğin 12, 24 ve 36 aylık dönemlerdeki dil gelişimleri, bebeğin anne sütü alma süresi, aynı dönemlerdeki kişisel sosyal ve ince motor gelişimleri ile açıklanabilmektedir. Bu durum, bebeğin süt alma süresinin onun beslenmesinde ve gelişimde önemli bir rol oynadığının, bebeğin geliştikçe kişisel, sosyal ve ince motor gelişimlerinin hızlanmasını ve bu gelişimlerin de bebeğin iletişim kurma ihtiyacını arttırdığını, bunun da bebeğin dil gelişiminin hızlanmasını sağladığını göstermektedir.

Model için sapma istatistiği değeri de hesaplanmış ve aşağıdaki tabloda gösterilmiştir.

Tablo 4.7. Tam Model İçin Sapma İstatistiği

Model için sapma istatistik değeri	
Sapma	= 899.124277
Kestirilen parametre sayısı	= 2

Bu değer elde edilecek modellerden en uygun olanının seçilmesinde kullanılacaktır. Modellerin karşılaştırılmasında sapma istatistiklerinin yanı sıra Akaike bilgi kriterinin (AIC) de kullanılabileceği belirtilmişti. Bu model için hesaplanmış olan AIC ise;

$$\begin{aligned} AIC &= 899{,}124277 + 2\,(1)\,(2) \\ &= 903{,}124277 \end{aligned}$$

olarak gösterilir.

4.3.2. Dil gelişiminin açıklanmasında kişisel sosyal gelişimin dikkate alınmadığı model

Tüm değişkenlerin kullanıldığı model dikkate alındığında kişisel sosyal gelişimi açıklayan değişkenlerin tamamının anlamsız oldukları görülmüştür. Bu nedenden modelden kişisel sosyal gelişim değişkenin çıkartılıp, tam modelle iç içe geçmiş hale gelmiş olan yeni bir model (iç içe geçmiş model1) elde edilmiştir. Dil gelişimini açıklamak için oluşturulan yeni modelde kullanılan değişkenler tabloda gösterilmiştir.

Tablo 4.8. İç İçe Geçmiş Model1'i Açıklamakta Kullanılacak Olan Değişkenler

Düzey1 Değişkenleri	Düzey2 Değişkenleri
Bebeğin ek gıda almaksızın anne sütü alma süresi (SUTSURE)	• Bebeğin cinsiyeti (CINSIYET) • Annenin eğitim durumu (ANNEEG) • Bebeğin doğduğu zamanki ağırlığı (AGIRLIK)
Bebeğin ince motor gelişim skorları (INCEMO)	• Bebeğin cinsiyeti (CINSIYET) • Annenin eğitim durumu (ANNEEG) • Bebeğin doğduğu zamanki boyu (BOY) • Bebeğin doğduğu zamanki ağırlığı (AGIRLIK)

Bu değişkenler kullanılarak elde edilecek iki düzeyli model aşağıdaki gibi gösterilecektir.

Düzey1 modeli:

$DIL_{ti} = \pi_{0i} + \pi_{1i} * SUTSURE + \pi_{2i} * INCEMO + e_{ti}$

Düzey2 model:

$\pi_{0i} = r_{0i}$

$\pi_{1i} = \beta_{10} + \beta_{11} * (CINSIYET) + \beta_{12} * (ANNEEG) + \beta_{13} * (AGIRLIK)$

$\pi_{2i} = \beta_{20} + \beta_{21} * (CINSIYET) + \beta_{22} * (ANNEEG) + \beta_{23} * (BOY) + \beta_{24}(AGIRLIK)$

29 iteratif işlemin sonucunda düzey1 ve düzey2 modelleri için hesaplanmış olan varyanslar aşağıda gösterilmiştir.

Sigma Kare (Düzey1 Varyansı - v) = 15,30087

Tau (Düzey2 Varyansı - τ) = 1756,92536

Tam model sonucunda elde edilen değerler incelendiğinde, kişisel sosyal değişkeninin modelden atılması düzey1 varyans değerini düşürürken, düzey2 varyans değerini de arttırdığı söylenebilir. Bu durumda model için güvenilirlik;

$$\text{güvenilirlik } (\hat{\beta}) = \frac{1756{,}92536}{1756{,}92536 + 15{,}30087} = 0{,}991$$

olarak hesaplanmıştır. Modelin güvenilirliğinin %98,9'dan, %99,1'e yükseldiği söylenebilmektedir. Tam modelde de oldukça yüksek bir orana sahip olan gözlenen değerler ile gerçek değerler arasındaki uyum, modelden kişisel sosyal gelişim değişkeninin çıkartılmasıyla daha da artmıştır.

Model için hesaplanmış olan katsayılar ve anlamlılıkları da aşağıda belirtilmiştir.

Tablo 4.9. İç İçe Geçmiş Model1'de Kullanılacak Düzey2 Modele Ait Katsayılar

Sabit Etkiler	Katsayılar	Standart Hatalar	t oranı	s.d.	p-değeri
SUTSURE slope, P1					
INTRCPT2, B10	12.522575	11.066601	1.132	111	0.261
CINSIYET, B11	0.034453	3.419726	0.010	111	0.992
ANNEEG, B12	-5.409765	3.706068	-1.460	111	0.147
AGIRLIK, B13	6.536576	3.135228	2.085	111	0.039
INCEMO slope, P2					
INTRCPT2, B20	-0.628379	1.137240	-0.553	111	0.581
CINSIYET, B21	0.008986	0.155997	0.058	111	0.955
ANNEEG, B22	0.320639	0.168567	1.902	111	0.059
BOY, B23	0.056602	0.032184	1.759	111	0.081
AGIRLIK, B24	-0.762469	0.212447	-3.589	111	0.001
Bağımlı Değişken : Dil					

Model için katsayılar incelendiğinde hem süt süresi için, hem de ince motor gelişimi için ağırlık ($p<0,05$) değişkeni anlamlı çıkmıştır. Yani bebeğin ek gıdaya geçmeden süt alma süresi ile ince motor gelişimi için bebeğin doğduğu zamanki ağırlığının etkisi diğer tüm etkilerden daha fazladır.

Elde edilen yeni modelde hata terimlerinin normal dağılıma uygunlukları Kolmogorov-Smirnov testi ile incelenmiş ve sonuçlar aşağıdaki tabloda verilmiştir.

Tablo 4.10. İç İçe Geçmiş Model1 Hata Terimlerine Ait Kolmogorov- Smirnov Testi

Kolmogorov-Smirnov Testi

		Düzey1 Hata Terimleri
N		120
Normal Parametreleri	**Ortalama**	,03785
	Std. Sapma	3,117558
Farklar	**Mutlak**	,246
	Pozitif	,246
	Negatif	-,244
Kolmogorov-Smirnov Z		2,699
Asymp. Sig. (2-yönlü)		,000

Bu test yardımıyla hata terimlerinin normal dağılıma uygunluk göstermedikleri belirlenmiştir. Tüm değişkenlerin kullanıldığı modelde olduğu gibi bu modelde de standart hata değerleri yerine, sağlam standart hata değerlerinin kullanılmasının daha uygun olduğuna karar verilmiştir. Bu nedenle hesaplanmış olan sağlam standart hata değerlerini de içeren tablo aşağıda gösterilmiştir.

Sonuçlara göre bebeğin ek gıdaya geçmeden önce anne sütü alma süresi ve 12, 24 ve 36 aylık dönemdeki ince motor gelişimini açıklamak için kullanılan değişkenlerden marjinal olarak en çok katkıyı, bebeğin doğduğu zamanki ağırlığını ifade eden değişken sağlamaktadır. Katsayılar da dikkate alındığında, bebeğin doğduğu zamanki ağırlığının artması normal şartlar altında anne sütü alma süresini arttırdığını ama ince motor gelişimini yavaşlattığını göstermektedir. Bu durum bebeğin yapısal olarak iri olması, basit fiziksel hareketlerin gerçekleştirilmesini ölçen ince motor gelişimini azaltmaktadır. Bebeğin ağırlığının fazla olması hareket kapasitesini azaltmaktadır.

Tablo 4.11. İç İçe Geçmiş Model1'de Kullanılacak Düzey2 Modele Ait Katsayılar (Sağlam Std. Hatalar İle)

Robust Standar Hatalarla					
Sabit Etkiler	Katsayılar	Standart Hatalar	T-Oranı	s.d.	p-değeri
SUTSURE slope, P1					
INTRCPT2, B10	12.522575	6.415896	1.952	111	0.053
CINSIYET, B11	0.034453	2.733319	0.013	111	0.990
ANNEEG, B12	-5.409765	4.002255	-1.352	111	0.179
AGIRLIK, B13	6.536576	2.307744	2.832	111	0.006
INCEMO slope, P2					
INTRCPT2, B20	-0.628379	0.950472	-0.661	111	0.510
CINSIYET, B21	0.008986	0.214507	0.042	111	0.967
ANNEEG, B22	0.320639	0.258934	1.238	111	0.219
BOY, B23	0.056602	0.030689	1.844	111	0.067
AGIRLIK, B24	-0.762469	0.261206	-2.919	111	0.005

Tüm bu etkiler altında modelin anlamlılığına ilişkin uygulanmış olan ki-kare testine ait sonuçlar tabloda gösterilmiştir.

Tablo 4.12. İç İçe Geçmiş Model1 İçin Ki-Kare Değeri

Rasgele Etkiler		Standart Sapma	Varyans	s.d.	Ki-Kare	p-değeri
INTRCPT1,	R0	41.91569	1756.92536	40	12266.96192	0.000
level-1,	E	3.91163	15.30087			

Model için hesaplanmış olan ki-kare değeri modelin anlamlı olduğunu göstermektedir. Bebeğin 12, 24 ve 36 aylık dönemlerdeki dil gelişimini açıklamak için aynı bebeğin anne sütü alma süresi ve aynı dönemlerdeki ince motor gelişimleri değişkenleri yeterlidir.

Bu modelin diğer modellerle karşılaştırılarak kullanılabilir olma durumunu gösterecek olan sapma istatistik değeri ise;

Tablo 4.13. İç İçe Geçmiş Model1 İçin Sapma İstatistiği

Model için sapma istatistik değeri	
Sapma	= 891.355937
Kestirilen parametre sayısı	= 2

olarak hesaplanmıştır. Alternatif bir kriter olan AIC ise;

$$
\begin{aligned}
AIC &= 891,355937 + 2\ (1)\ (2) \\
&= 895,355937
\end{aligned}
$$

olarak hesaplanmıştır.

4.3.3. Dil gelişiminin açıklanmasında kişisel sosyal gelişimin ve anne eğitiminin dikkate alınmadığı model

Kişisel sosyal gelişimi açıklayan değişkenlerin tamamının ve de incelenmiş olan her iki modelde de anlamsız olduğu görülen annenin eğitim durumunu gösteren değişkenin modelden çıkartılmasının, bebeğin incelenen süreçteki dil gelişimine etkisi araştırılmak istenmiştir. Kişisel sosyal gelişim ve anne eğitimi değişkenin çıkartıldığı model bir önceki bölümde açıklanan model gibi, tam modelle iç içe geçmiş halde bulunmaktadır ve iç içe geçmiş model2 olarak adlandırılmıştır. Dil gelişimini açıklamak için oluşturulan yeni modelde kullanılan değişkenler tabloda gösterilmiştir.

Tablo 4.14. İç İçe Geçmiş Model2'yi Açıklamakta Kullanılacak Olan Değişkenler

Düzey1 Değişkenleri	Düzey2 Değişkenleri
Bebeğin ek gıda almaksızın anne sütü alma süresi (SUTSURE)	• Bebeğin cinsiyeti (CINSIYET) • Bebeğin doğduğu zamanki ağırlığı (AGIRLIK)
Bebeğin ince motor gelişim skorları (INCEMO)	• Bebeğin cinsiyeti (CINSIYET) • Bebeğin doğduğu zamanki boyu (BOY) • Bebeğin doğduğu zamanki ağırlığı (AGIRLIK)

Bu değişkenler kullanılarak elde edilecek iki düzeyli model aşağıdaki gibi gösterilecektir.

Düzey1 modeli:

$$DIL_{ti} = \pi_{0i} + \pi_{1i} * SUTSURE + \pi_{2i} * INCEMO + e_{ti}$$

Düzey2 model:

$$\pi_{0i} = r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} * (CINSIYET) + \beta_{12} * (AGIRLIK)$$

$$\pi_{2i} = \beta_{20} + \beta_{21} * (CINSIYET) + \beta_{22} * (BOY) + \beta_{23} * (AGIRLIK)$$

31 iteratif işlemin sonucunda düzey1 ve düzey2 modelleri için hesaplanmış olan varyanslar aşağıda gösterilmiştir.

Sigma Kare (Düzey1 Varyansı - v) = 15,07322

Tau (Düzey2 Varyansı - τ) = 1903,76090

Elde edilen yeni varyans değerleri yardımıyla model için güvenilirlik;

$$\text{güvenilirlik } (\hat{\beta}) = \frac{1903{,}76090}{1903{,}76090 + 15{,}07322} = 0{,}992$$

olarak hesaplanmıştır. Modelin güvenilirliğinin %99,1'den, %99,2'ye yükseldiği söylenebilmektedir. Ancak %0,1'lik bir artışın modellerin karşılaştırılmasında dikkate alınmasının doğru olmadığı söylenebilir.

Model için hesaplanmış olan katsayılar ve anlamlılıkları da aşağıda belirtilmiştir.

Tablo 4.15. İç İçe Geçmiş Model2'de Kullanılacak Düzey2 Modele Ait Katsayılar

Sabit Etkiler	Katsayılar	Standart Hatalar	T-Oranı	s.d.	p-değeri
SUTSURE slope, P1					
INTRCPT2, B10	12.262257	11.174255	1.097	113	0.275
CINSIYET, B11	-1.980641	3.358468	-0.590	113	0.556
AGIRLIK, B12	5.494537	3.178966	1.728	113	0.086
INCEMO slope, P2					
INTRCPT2, B20	-0.090060	1.118531	-0.081	113	0.936
CINSIYET, B21	0.160216	0.132191	1.212	113	0.228
BOY, B22	0.039956	0.031664	1.262	113	0.210
AGIRLIK, B23	-0.623982	0.204718	-3.048	113	0.003
Bağımlı Değişken: Dil					

Model için katsayılar incelendiğinde ince motor gelişimi için ağırlık ($p<0{,}05$) değişkeni anlamlı çıkmıştır. Yani bebeğin ince motor gelişimi olarak adlandırılan basit fiziksel hareketlerin gerçekleştirilebilmesi için bebeğin doğduğu zamanki ağırlığının etkisi diğer tüm etkilerden daha fazladır.

Elde edilen yeni modelde hata terimlerinin normal dağılıma uygunlukları Kolmogorov-Smirnov testi ile incelenmiş ve sonuçlar aşağıdaki tabloda verilmiştir.

Tablo 4.16. İç İçe Geçmiş Model2 Hata Terimlerine Ait Kolmogorov- Smirnov Testi

Kolmogorov- Smirnov Testi

		Düzey1 Hata Terimleri
N		120
Normal Parametreleri	Ortalama	,03697
	Std. Sapma	3,111335
Farklar	Mutlak	,250
	Pozitif	,250
	Negatif	-,247
Kolmogorov-Smirnov Z		2,737
Asymp. Sig. (2-yönlü)		,000

Bu test yardımıyla hata terimlerinin normal dağılıma uygunluk göstermedikleri belirlenmiştir. Diğer iki modelde de olduğu gibi bu modelde de standart hata değerleri yerine, sağlam standart hata değerlerinin kullanılmasının daha uygun olduğuna karar verilmiştir. Bu nedenle hesaplanmış olan sağlam standart hata değerlerini de içeren tablo aşağıda gösterilmiştir.

Sonuçlara göre bebeğin ek gıdaya geçmeden önce anne sütü alma süresi ve 12, 24 ve 36 aylık dönemdeki ince motor gelişimini açıklamak için kullanılan değişkenlerden marjinal olarak en çok katkıyı, bebeğin doğduğu zamanki ağırlığını ifade eden değişken sağlamaktadır.

Tablo 4.17. İç İçe Geçmiş Model2'de Kullanılacak Düzey2 Modele Ait Katsayılar (Sağlam Std. Hatalar İle)

Robust standart Hatalarla					
Sabit Etkiler	Katsayılar	Standart Hatalar	T-Oranı	s.d.	p-değeri
SUTSURE slope, P1					
INTRCPT2, B10	12.262257	7.866962	1.559	113	0.122
CINSIYET, B11	-1.980641	2.635379	-0.752	113	0.454
AGIRLIK, B12	5.494537	2.308033	2.381	113	0.019
INCEMO slope, P2					
INTRCPT2, B20	-0.090060	0.891315	-0.101	113	0.920
CINSIYET, B21	0.160216	0.162625	0.985	113	0.327
BOY, B22	0.039956	0.027885	1.433	113	0.155
AGIRLIK, B23	-0.623982	0.238271	-2.619	113	0.010

Tüm bu etkiler altında modelin anlamlılığına ilişkin uygulanmış olan ki-kare testine ait sonuçlar tabloda gösterilmiştir.

Tablo 4.18. İç İçe Geçmiş Model2 İçin Ki-Kare Değeri

Rasgele etkiler		Standart Sapma	Varyans	s.d.	Ki-Kare	p-değeri
INTRCPT1,	R0	43.63211	1903.76090	40	13927.35867	0.000
level-1,	E	3.88242	15.07322			

Model için hesaplanmış olan ki-kare değeri modelin anlamlı olduğunu göstermektedir. Bebeğin 12, 24 ve 36 aylık dönemlerdeki dil gelişimini açıklamak için aynı bebeğin anne sütü alma süresi ve aynı dönemlerdeki ince motor gelişimleri değişkenleri, kendi alt değişkenleri ile birlikte yeterlidir.

Bu modelin diğer modellerle karşılaştırılarak kullanılabilir olma durumunu gösterecek olan sapma istatistik değeri ise;

Tablo 4.19. İç İçe Geçmiş Model2 İçin Sapma İstatistiği

Model için sapma istatistik değeri	
Sapma	= 897.526212
Kestirilen parametre sayısı	= 2

olarak hesaplanmıştır. Alternatif bir kriter olan AIC ise;

$$AIC = 897{,}526212 + 2\,(1)\,(2)$$
$$= 901{,}526212$$

olarak hesaplanmıştır.

4.3.4. Modellerin birbirleriyle karşılaştırılmaları

Uygulamalar sonucunda üç tane model elde edilmiştir. Bunlardan bir tanesi eldeki tüm değişkenler kullanılarak oluşturulmuş olan tam model, bir tanesi katsayılarının anlamsız çıkması yüzünden modele hiçbir katkı sağlamadığı düşünülen kişisel sosyal gelişimini açıklayan değişkenin modelden çıkarılması ile elde edilen olan ve tam modelle iç içe geçmiş halde bulunan model ve diğeri de düzey1'de bulunan kişisel sosyal gelişim ile düzey2'de bulunan ve modelde anlamsız çıkan anne eğitimi değişkenlerinin çıkarılmasıyla elde edilen model. Elde bulunan üç modelden hangisinin bebeklerin dil gelişimlerini açıklamakta kullanılacağını belirlemek için hesaplanmış olan sapma istatistikleri kullanılacaktır.

Hesaplanmış olan sapma istatistikleri ve AIC değerleri aşağıdaki tabloda tüm modeller için bir arada gösterilmiştir.

Tablo 4.20. Modellere İlişkin Sapma İstatistikleri

Tam Model	İç İçe Geçmiş Model1	İç İçe Geçmiş Model2
Sapma İstatistiği : 899,124277	Sapma İstatistiği : 891,355937	Sapma İstatistiği : 897,526212
AIC : 903,124277	AIC : 895,355937	AIC : 901,526212

Daha önceden açıklandığı gibi, anlamlı olduğu bilinen modellere ilişkin sapma istatistikleri karşılaştırıldığında, değeri küçük olan model tercih edilen, kullanılması diğerine göre daha uygun olan model olarak açıklanır. AIC değeri için de aynı durum söz konusundur. Her üç model için hesaplanmış olan değerler incelendiğinde iç içe geçmiş modele ait olan sapma istatistiği ve AIC değeri, tam model ve iç içe geçmiş model2 için hesaplanan değerlerden daha düşük olduğu gözlenmiştir. Bu nedenle ikinci modelin (iç içe geçmiş model1) kullanılması daha uygundur. Bu da bebeklerin dil gelişimlerinin, bebeğin anne sütü alma ve aynı dönemlerdeki ince motor gelişimleri yardımıyla açıklanmasının daha uygun olduğunu göstermektedir.

Elde edilen üç model için BIC değerleri de hesaplanmış ve karşılaştırmada kullanılan diğer istatistiklerle aynı sonuçları verdiği için uygulamaya eklenmemiştir.

TARTIŞMA VE SONUÇ

Mature Eğitim ve Aile Danışmanlığı Merkezi'nden hizmet alan 40 bebeğin 12, 24 ve 36 aylık dönemlerdeki dil gelişimlerini etkileyen değişkenler incelenmiştir. Bu nedenle dil gelişimi, bebeğin ek gıdaya geçmeden önce anne sütü alma süresi, aynı dönemlerdeki kişisel sosyal gelişimi ve ince motor gelişimleri ile açıklanmaya çalışılmıştır. Bunu açıklamak için daha fazla ayrıntıya ihtiyaç duyulmuş ve bu değişkenleri ekleyebileceği düşünülen cinsiyet, bebeğin doğduğu zamanki ağırlığı ve boy uzunluğu, annenin doğum yaptığı zamanki yaşı ve annenin eğitimi değişkenleri ikinci düzey olarak modele eklenmiştir. Bebeğin ek gıdaya geçmeden anne sütü alma süresinin bebeğin doğum zamanındaki ağırlığıyla, ince motor gelişiminin ise bebeğin doğduğu zamanki boy ve ağırlığıyla açıklanabildiği model anlamlı çıkmıştır. Bu durum bebeğin anne sütü alması süresinin artması fiziksel olarak geliştiğini ve dili kullanma yeteneğini arttırdığını göstermektedir. Ayrıca bebeğin ince motor hareketleri olarak adlandırılan basit fiziksel hareketleri yerine getirmesi, söylenenleri anlayarak tepki vermesi ve sosyal ortamlarda bulunması, topluluğa uyum sağlama ve ihtiyaçlarını ifade edebilmesi gerekliliğini doğurduğundan dil gelişimini hızlandırıcı yönde etki yaptıkları söylenebilmektedir.

Model incelenirken kişisel sosyal gelişime ait düzey2 değişkenlerinin model için anlamsız oldukları belirlenmiş ve bu nedenle kişisel sosyal değişkenin modelden çıkarılarak daha uygun sonuçlara ulaşılacağı düşünülmüştür. Çok düzeyli analiz tekrar modele uygulanmış ve anne sütü alma süresiyle ince motor gelişim değerlerinin, dil gelişimini açıklamak için yeterli olduğu görülmüştür. Ancak tam modelden farklı olarak sadece bebeğin doğduğu zamanki ağırlığı süt alma süresi ve ince motor gelişimlerini etkileyen değişken olarak karşımıza çıkmaktadır. Bu durum bebeğin doğum sırasındaki ağırlığının, diğer bir deyişle fiziki yapısının bebeğin süt alma süresini ve basit fiziksel hareketleri gerçekleştirebilme kapasitesini etkilemektedir.

En son aşamada kişisel sosyal değişkenin yanı sıra tüm düzey2 değişkenleri için kullanılan ve her birinde anlamsız sonuçlar veren annenin eğitim durumu değişkeni de modelden çıkartılmıştır. Bu şekilde analiz edilen model de anlamlı çıkmıştır. Dil gelişimini açıklarken bebeğin süt alma süresinin ve ince motor gelişiminin kullanılmasının yanlış olmadığı görülmüştür. Bir önceki durumla aynı olarak yine bebeğin doğduğu zamanki ağırlığı gelişim sürecini açıklamakta kullanılan en önemli değişken olarak karşımıza çıkmaktadır.

Elde edilen bu üç modelden bebeğin 12, 24 ve 36 aylık dönemlerdeki dil gelişimini açıklamakta kullanılacak olanının belirlenmesi için sapma ve AIC istatistikleri kullanılmış ve bu değerleri en düşük çıkan modelin kullanılmasına karar verilmiştir. Sadece kişisel sosyal değişkenin çıkarıldığı model için bu istatistikler daha düşük değerlere sahip oldukları için bu modelin kullanılmasına karar verilmiştir.

Sonuç olarak, bebeğin anne sütü alarak fiziksel olarak gelişmesine ve ince motor gelişim skoruna bağlı olarak ihtiyaçlarını yerine getirmek için iletişim kurma gereksiniminin artmasından dolayı, bebeğin 12, 24 ve 36 aylık dönemlerdeki dil gelişimi hızlanmaktadır. Bunun yanı sıra bebeğin süt alma süresi ve fiziksel gelişimin bir göstergesi olan ince motor gelişimi, bebeğin doğduğu zamanki ağırlığı ile açıklanabilmektedir. Bebeğin ağırlığının fazla olması fiziksel hareketlerini engellemekte ve ince motor gelişimini yavaşlatmaktadır denilebilmektedir.

Bu çalışmada sadece iki düzey kullanılmıştır. Daha çok düzey içeren uygulamalar ve modellerde değişken seçiminin uygulandığı durumlar bundan sonraki çalışmalarda araştırılabilir.

KAYNAKLAR

ADAMSON, L.; CHOJENTA, C.; LEE, C., ***Telephone contact of existing participants in longitudinal surveys,*** Australian & New Zealand Journal of Public Health, Vol. 29 Issue 2, p188, 2p, Apr2005.

ANLAR, B., YALAZ, K., ***Denver II Gelişimsel Tarama Testi, Türk Çocuklarına Uyarlanması ve Standardizasyonu,*** Hacettepe Üniversitesi Yayınları, Ankara, 2000.

BLAAS, H. G., EIK-NES, H. and BREMNES, J. B., ***The Growth of the Human Embryo, A Longitudinal Biometric Assessment From 7 to 12 Weeks of Gestation,*** Ultrasound Obstet Gynecol, 12:346-354, 1998.

CATRIEN, C. J., BIJLEVELD and LEO, J., ***Longitudinal Data Analysis, Designs, Models and Methods,*** Sage Publications, London, 1998.

CHISWICK, B., LEE, Y. and MILLER, P., ***Longitudinal Analysis of Immigrant Occupational Mobility: A Test of the Immirant Assimilation Hypothesis,*** Discussion Paper Series, IZA DP No. 452, March, 2002.

DIGGLE, P. J., HEAGERTY, P., LIANG, K. and Zeger, S. L., ***Analysis of Longitudinal Data,*** Second Edition, Oxford University Pres, Oxford, 2002.

FARAWAY, J., ***A Graphical Method of Exploring the Mean Structure in Longitudinal Data Analysis,*** Journal of Computational and Graphical Statistics, Volume 8, Number 1, p:60-68, 1999

FITZMAURICE, G. M., LAIRD, N. M. and WARE, J. H., ***Applied Longitudinal Analysis***, Wiley Series in Probability and Statistics, John Wiley & Sons, inc., Publication, 2004.

HAGENAARS, J., ***Categorical Longitudinal Data, Log-Linear, Panel, Trend and Cohort Analysis,*** Sage Publications, London, 1990.

LIPOVETSKY, S., ***Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models,*** Technometrics, Vol. 47 Issue 2, p237, 1/2p, May2005.

McARDLE, J., CAJA, E., HAMAGAMI F. and WOODCOCK, R., ***Comparative Longitudinal Structural Analysis of the Growth and Decline of Multiple Intellectual Abilities Over the Life Span,*** Journal of Developmental Psychology, Vol. 38, No. 1, 115-142, 2002.

MENARD, S., ***Longitudinal Research,*** Series: Quantitative Applications in the Social Sciences, Sage Publications, 1991.

MULLER, H.G., ***Functional Modelling and Classification of Longitudinal Data,*** Scandinavian Journal of Statistics, , Vol. 32 Issue 2, p223, 18p, Jun2005.

PHILLIPS, S., BANDINI, L. And MUST, A., ***A Longitudinal Comparison of Body Composition by Total Body Water and Bioelectrical Impedance in Adolescent Girls,*** The Journal of Nutrition, 133,5, Academic Research Library, p.1419, May, 2003.

RAUDENBUSH, S. W., BRYK, A. S., ***Hierarchical Linear Models, Applications and Data Analysis Methods,*** Second Edition, Sage Publications, 2002.

REITMAN, F; SCHNEER, Joy A., ***The Long-Term Negative Impacts Of Managerial Career Interruptions: A LONGITUDINAL STUDY OF MEN AND WOMEN Mbas,*** Group & Organization Management, Vol. 30 Issue 3, p243, 20p, Jun 2005.

RIBAUDO, H., BERNHARD, M. and THOMPSON S., ***A Multilevel Analysis of Longitudinal Ordinal Data: Evaluation of the Level of Physical Performance of Women Receiving Adjuvant Therapy for Breast Cancer,*** Journal of Royal Statistical Society, 162, Part 3, pp.349-360, 1999.

SEONG-J. K.; TAKEHİKO F.; Kİ-YONG C.; MYUNG-WON S., ***Model-matching control applied to longitudinal and lateral automated driving***, Journal of Automobile Engineering, Vol. 219 Issue 5, p583, 16p, Proceedings of the Institution of Mechanical Engineers, 2005.

SINGER, J. D., WILLETT, J. B., ***Applied Longitudinal Data Analysis, Modeling Change and Event Occurrence,*** Oxford University Press, New York, 2003.

SNIJDERS, T., BOSKER, R., ***Multilevel Analysis, An Introduction to Basic and Advanced Multilevel Modeling,*** Sage Publications, London, 1999.

SPIKMAN, J., TIMMERMAN, M. and ZOMEREN, A., ***Recovery Versus Retest Effects in Attention After Closed Head Injury,*** Journal of Clinical and Experimental Neuropsychology, Vol.21, No.5, pp. 585-605, 1999.

SRİNİVASAN, K. K.; MAHMASSANİ, H. S., ***A Dynamic Kernel Logit Model for the Analysis of Longitudinal Discrete Choice Data: Properties and Computational Assessment***, Transportation Science, Vol. 39 Issue 2, p160, 22p, May2005.

TARIS, T. W., ***A Primer in Longitudinal Data Analysis***, Sage Publications, London, 1999.

VON EYE, A,(Edited by), ***Statistical Methods in Longitudinal Research, Volume I, Principles and Structuring Change,*** Academic Press, London, 1990.

VON EYE, A,(Edited by), ***Statistical Methods in Longitudinal Research, Volume II, Time Series and Categorical Longitudinal Data,*** Academic Press, London, 1990.

EKLER

EK – 1. Gelişim Maddeleri

Tablo 1.A. 12 Aylık Bebek İçin İlgilenilen Gelişim Maddeleri

12 Aylık Bebeğin Gelişimi İçin İlgilenilen Maddeler			
Kişisel-Sosyal	İnce Motor	Dil	Kaba Motor
• El çırpma • İsteklerini ağlamadan belirtme • Testorle top oynama • Bardaktan su içme • Ev işlerini taklit • Evde basit işlere yardım	• Üzümü baş ve diğer parmaklarla alma • Küpü elden ele geçirme • Küpleri birbirine vurma • Küpü bardağa koyma • Üzümü gösterilince çıkarma • Kendiliğinden karalama	• Sesli gülme • Tıkırtı sesine dönme • İnsan sesine dönme • Tek heceler söyleme • Özgül olmayan anne/baba • Konuşma seslerini taklit • Özgül olarak anne/baba • Anne, babadan başka iki sözcük • Anne, babadan başka dört sözcük • Bebeğin 1 kısmını gösterme • Bebeğin 2 kısmını gösterme • Bebeğin 4 kısmını gösterme • Yarı anlaşılır konuşma • Bir yönergeye uyma	• Tutunarak ayakta durma • Sıralama • Tutunarak ayağa kalkma • 2 sn ayakta durma • 10 sn ayakta durma • Düzgün yürüme

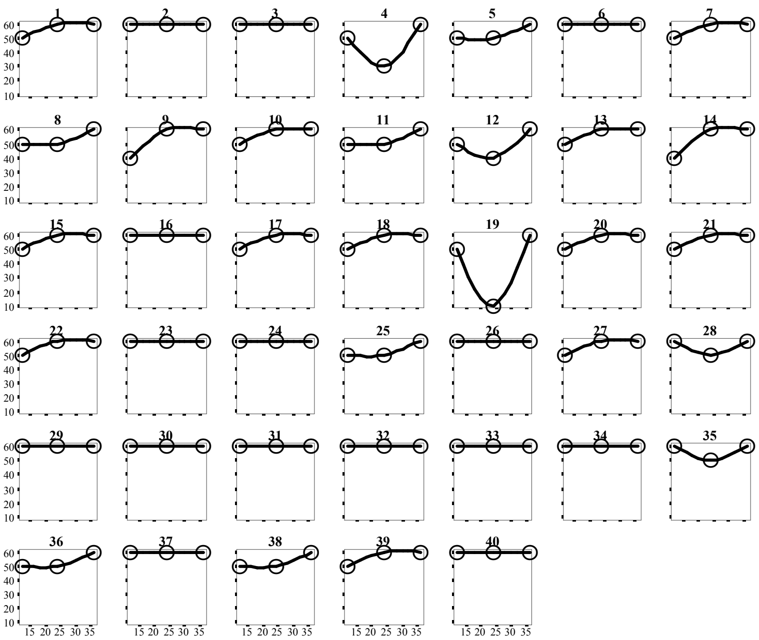
Tablo 1.B. 24 Aylık Bebek İçin İlgilenilen Gelişim Maddeleri

24 Aylık Bebeğin Gelişimi İçin İlgilenilen Maddeler			
Kişisel-Sosyal	İnce Motor	Dil	Kaba Motor
• Bardaktan su içme • Ev işlerini taklit • Evde basit işlere yardım • Kaşık çatal kullanma • El yıkama-kurulama • Giysisini giyme	• Kendiliğinden karalama • 3 küpten kule • 4 küpten kule • 5 küpten kule • 6 küpten kule • 7 küpten kule	• Özgül olarak anne/baba • Anne, babadan başka 2 sözcük • Anne, babadan başka 4 sözcük • Bebeğin 1 kısmını gösterme • Bebeğin 2 kısmını gösterme • Bebeğin 4 kısmını gösterme • Yarı anlaşılır konuşma • Bir yönergeye uyma • 1 resmi gösterme • 2 resmi gösterme • 1 resmin adını söyleme • 2 resmin adını söyleme • Yer bildiren bir terim anlama • Bir işlev bilme	• Topa vurma • Koşma • Merdiven çıkma • Yerinde zıplama • Tek ayakta 1 sn durma • Tek ayakta 2 sn durma

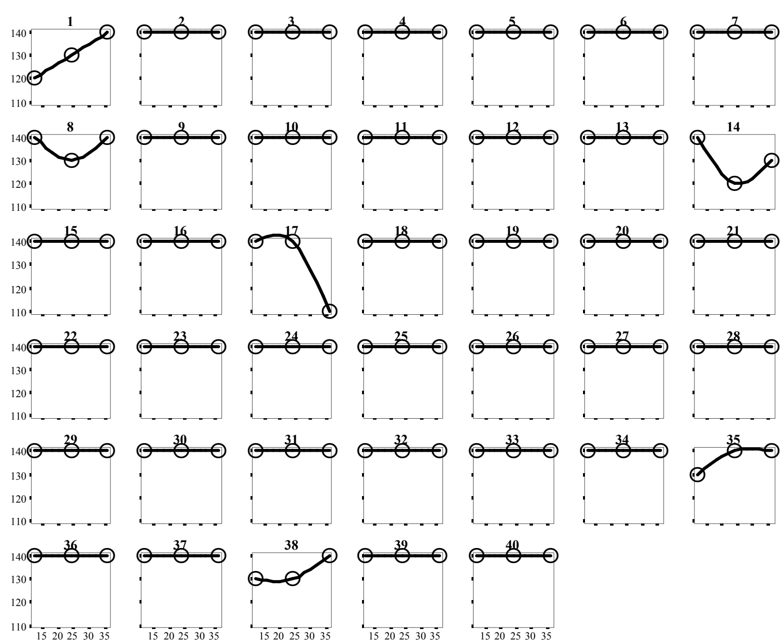
Tablo 1.C. 36 Aylık Bebek İçin İlgilenilen Gelişim Maddeleri

36 Aylık Bebeğin Gelişimi İçin İlgilenilen Maddeler			
Kişisel-Sosyal	İnce Motor	Dil	Kaba Motor
• Düğme ilikleme • Tişört giyme • Giysisini giyme • El yıkama-kurulama • Kaşık çatal kullanma • Evde basit işlere yardım	• 5 küpten kule • 6 küpten kule • 7 küpten kule • Dik çizgiyi kopya • 8 küpten kule • 0 kopya	• 1 resmi gösterme • 2 resmi gösterme • 1 resmin adını söyleme • 2 resmin adıı söyleme • Yer bildiren bir terim anlama • 4 resim gösterme • Bir işlev bilme • 3 resmin adını söyleme • Üşü, yorgun, aç-1 doğru • 1 cismin kullanımını bilme • Tam anlaşılır konuşma • Bir sözcük tanımlama • Zıt anlamlar 1/3 • 2 sözcük tanımlama	• Yerinde zıplama • Tek ayakta 1 sn durma • Tek ayakta 2 sn durma • Uzağa atlama • Tek ayakta 3 sn durma • Tek ayakta 4 sn durma

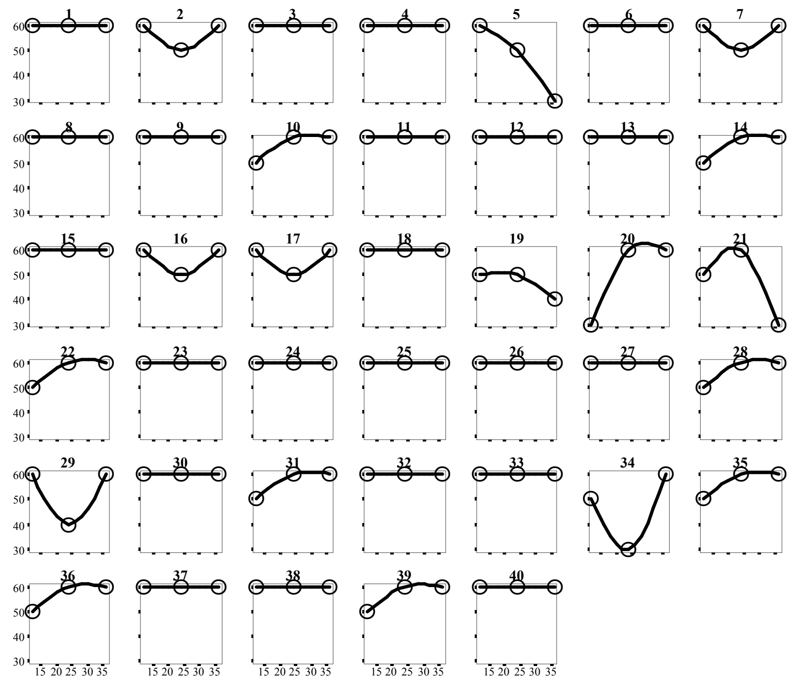
EK – 2. Düzleştirme Eğrileri



Şekil 2.A. İnce Motor Gelişimi İçin Elde Edilen Düzleştirme Eğrisi

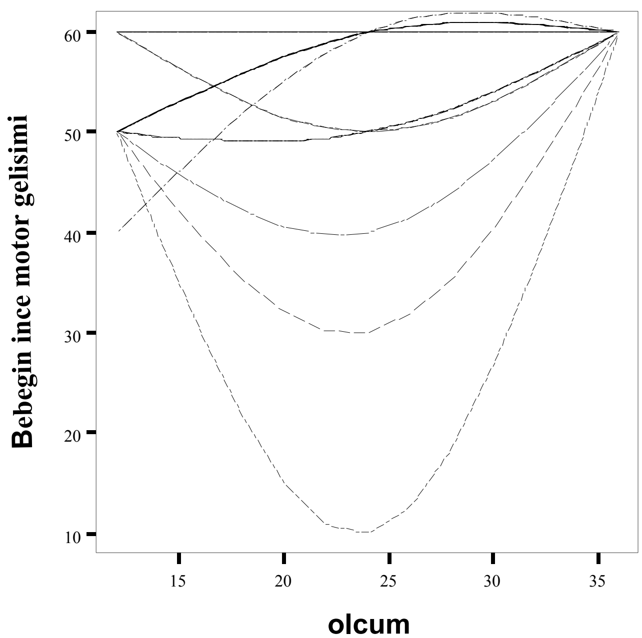


Şekil 2.B. Dil Gelişimi İçin Elde Edilen Düzleştirme Eğrisi

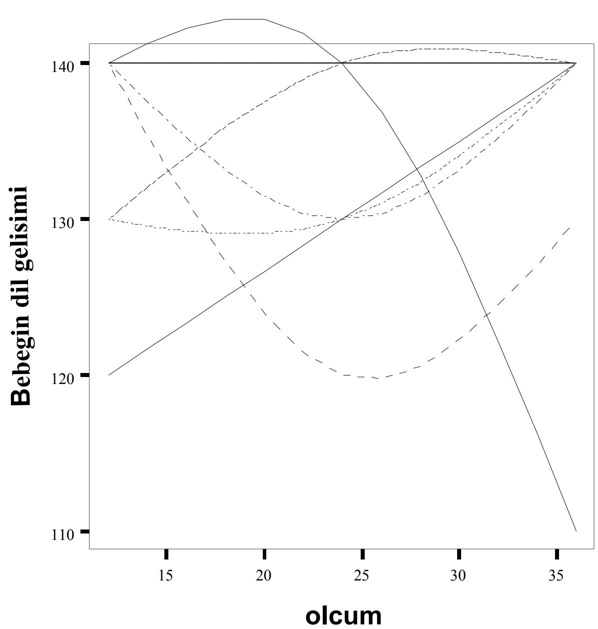


Şekil 2.C. Kaba Motor Gelişimi İçin Elde Edilen Düzleştirme Eğrisi

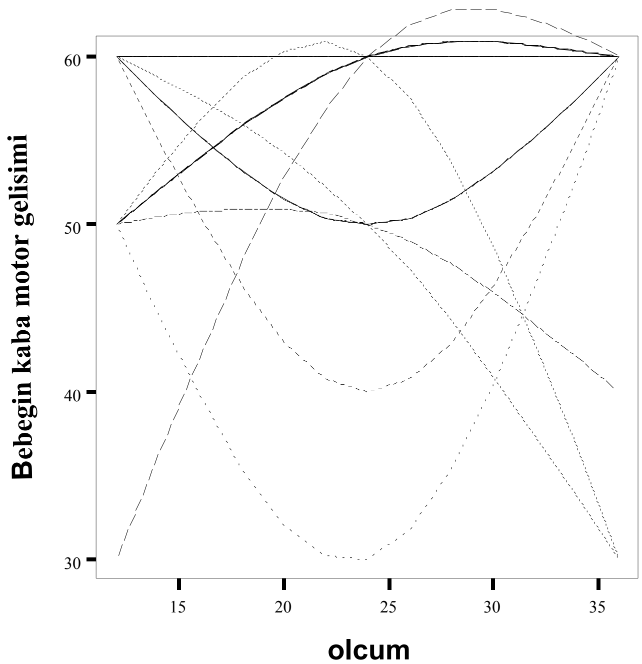
EK – 3. Tüm Bebekler İçin Düzleştirme Eğrileri



Şekil 3.A. İnce Motor Gelişimi İçin Tüm Bebeklerle Elde Edilen Düzleştirme Eğrisi



Şekil 3.B. Dil Gelişimi İçin Tüm Bebeklerle Elde Edilen Düzleştirme Eğrisi



Şekil 3.C. Kaba Motor Gelişimi İçin Tüm Bebeklerle Elde Edilen Düzleştirme Eğrisi