

Küçük Örneklerde Beta4 ve Polynomial Loglineer Öndüzgünleştirme ve Kübik Eğri Sondüzgünleştirme Metotlarının Uygunluğu

Accuracy of Beta4 Presmoothing Polynomial Loglineer Presmoothing and Cubic Spline Postsmoothing Methods for Small Samples

Şeref TAN

Gazi Üniversitesi Gazi Eğitim Fakültesi, Eğitimde Ölçme ve Değerlendirme Anabilim
Dalı, sereftan@gmail.com

ÖZ

Bu çalışmada eşit yüzdelliği eşitlemede kullanılan beta4 ve polinomial loglineer öndüzgünleştirme ve kübik eğri sondüzgünleştirme yöntemlerinin uygunluğunun gerçek bir veri seti ve küçük örnekler için karşılaştırılması sunulmuştur. Form X ve Form Y verilerine beta4 öndüzgünleştirme metodu uygulandığında düzgünleştirilmiş dağılım Form X için tüm dört momenti korumuş ve Form Y için ilk üç momenti korumuştur. Ki kare istatistiğine göre beta4 öndüzgünleştirme metodu deneysel veriye uyum göstermiştir. Polinomial loglineer öndüzgünleştirmede, hem Form X hem de Form Y verileri için ikinci dereceden polinomial loglineer modelle yapılan düzgünleştirme en iyi veri uyumunu sağlamıştır. Kübik eğri sondüzgünleştirme metodu, $S=30$ düzeyinde en uygun model uyumunu sağlamıştır. Sonuç olarak, bu çalışmada kullanılan üç düzgünleştirme metodu da etkili bulundu ve hem eşitlemenin ortalama bootstrap standart hatası hem de moment korunumu ölçütü dikkate alındığında 200-250 civarı gibi küçük örnekler için üç düzgünleştirme metodundan beta4 öndüzgünleştirme metodunun kullanımının göreceli olarak daha uygun olduğu bulunmuştur.

Anahtar Kelimeler: Beta4 öndüzgünleştirme, Polinomial loglineer öndüzgünleştirme, Kübik eğri sondüzgünleştirme, Eşit yüzdelliği eşitleme.

ABSTRACT

In this study, a comparison of beta4 and polynomial loglinear presmoothing methods and the cubic spline postsmoothing method of equipercetile equating with balanced single group design data set for small samples are presented. When the beta4 presmoothing method was applied to Form X and Form Y data, Form X maintained for all the four sample data moments and Form Y maintained for the first 3 sample data moments. According to the chi square statistics beta4 presmoothing method showed good fit with sample data. In the polynomial loglinear presmoothing, presmoothing with second degree polynomial model provided the best data fit for both Form X and FormY. Cubic spline postsmoothing method has provided the most accurate fit

for $S = .30$. As a conclusion, all of the three smoothing methods, used in the study, were effective and relatively beta4 presmoothing method was found the most appropriate method among the three smoothing methods by using both bootstrap standart errors of equating and moment preservation criterion for small samples such as 200-250.

Keywords: Beta4 presmoothing, Polynomial loglineer presmoothing, Cubic postsmoothing, Equipercentile equating

GİRİŞ

Aynı amaç için hazırlanan birden çok test formları aynı veya farklı zamanlarda kullanıldıklarında, bu formlardan alınan puanlardan faydalanarak öğrenciler hakkında bazı kararlar verilmektedir. Bu durumlarda, öğrencinin hangi test formunu aldığına bakmaksızın öğrencinin başarı konumu mutlak veya bağıl anlamda mümkün oldukça doğru olacak şekilde belirleyebilmek için test puanlarının eşitlenmesi gerekir. Fairbank (1987) test eşitlemeyi, iki veya daha fazla benzer testlerin hangi puanlarının bir öğrenci evreninde aynı yetenek düzeyine karşılık geldiğinin bulunması olarak tanımlar. Test puanlarının eşitlenmesinde aynı içerik ve istatistiksel özelliklere sahip test formlarının güçlüklerinden gelen farklılıkların düzeltilmesi amaçlanır. Aynı içerik ve istatistiksel özelliklere sahip olacak şekilde test formları hazırlansa bile pratikte bu test formlarının güçlük düzeyleri farklı olabilmektedir. Söz konusu paralel olarak hazırlanan test formlarının güçlük düzeyleri arasındaki farklılıklar istatistiksel olarak manidar bulunmasada, test formlarının güçlük düzeylerinden kaynaklanan bu farklılıklar öğrencilerin sıralamasını ve öğrenci puanlarına yönelik olarak yapılacak bağıl değerlendirmeleri etkilemektedir. Test eşitlemeyle, aynı amaç için hazırlanmış test formlarının güçlüklerinden kaynaklanan farklılıkların giderilmesi sonucunda öğrencilerin daha kolay veya daha zor bir test formunu almış olmasından dolayı bir avantaj veya dezavantaj sağlanması önlenmiş olunur.

Test formlarından elde edilen puanların dağılımı veya iki test formunu birbirine eşitlemede kullanılan eşitleme fonksiyonunu örneklemden elde ettiğimiz istatistiklerden faydalanarak kestiririz. Çoktan seçmeli testlerde test maddeleri doğru cevaplandığında 1 puan ve diğer durumlarda 0 puan olarak puanlandığında elde edilen test puanları süreklilik özelliği göstermemektedir. Örneğin Form X 20 tane çoktan seçmeli test

maddesinden oluşuyorsa testten alınabilecek puanlar 0 ile 20 arasındaki tam sayılardır. Bu durumda Form X puanları Form Y puanlarına dönüştürüldüğünde elde edilen dönüştürme eşitliği fonksiyonu veya grafiği düzgün bir grafik olmamakta veya bir eşitlikle fonksiyonun belirtilmesinde dağılım süreksiz olduğundan dolayı sorun yaşanmaktadır. Eşit yüzdelikli eşitlemede, eşitleme ilişkilerini kestirmede örneklem yüzdeler ve yüzdeler sıraları kullanıldığından, örnekleme hatasından dolayı pratik amaçlar için eşitleme genellikle yeterli kesinlikte olmaz (Kolen & Brennan, 2014, 65). Bu nedenle test puanları dağılımının veya dönüştürme fonksiyonunun evrendeki biçimi kestirilmeye çalışılmaktadır. Bu kestirimin yapılmasında ise örneklem dağılımlarının düzgünleştirilmesi yoluna gidilmektedir. López, Olson & Bansal (2011)'in da belirttiği gibi düzgünleştirme puan dağılımındaki düzensizliği azaltır. Anlaşılabileceği üzere, düzgünleştirme metotları deneysel dağılımların kestirimlerini ve evren karakteristiğinde düzgünlük özelliği olan eşit yüzdelikli ilişkileri sağlamak amacıyla geliştirilmiştir (Kolen & Brennan, 2014, 65). Ayrıca Livingston'un (1992) belirttiği gibi düzgünleştirmeye (belirli bir uygunluğun sağlanması için) gerekli olan örneklem büyüklüğü en azından yarı yarıya düşürür.

Düğünlleştirme kullanılsın veya kullanılmınsın, yapılan her eşitleme işlemi bazı eşitleme hatalarını içerir. Wang'ın (2006) belirttiği gibi eşitleme hataları randum ve sistematik hata olarak sınıflandırılabilir, Sistematik hatalar genelde eşitleme metodunda sayıtların sağlanamaması, örneklem istatistiğindeki yanlılık gibi nedenlerden oluşurken randum hata örnekleme hatasından oluşur. Bağlı değerlendirme amaçlı ölçmelerde randum ve sistematik hatanın kontrolü önemlidir. Eşit yüzdelikli eşitleme çalışmalarında kullanılan düğünlleştirme çalışmaları eşitlemedeki randum hatayı azaltmaktadır. Buna rağmen, düğünlleştirme metotları eşitlemedeki sistematik hataya yönelik hiçbir iyileşme veya düzeltme yapmaz. Hatta bazı durumlarda düğünlleştirme sonucu randum hatada azalma olmasına rağmen sistematik hatada bir artış bile gözlenebilir. Her halükarda düğünlleştirme sonucu toplam hatada (randum ve sistematik hatanın toplamında) bir azalma gözlenir.

Bu araştırmada eşit yüzdellikli eşitlemede test puanları dağılımının ve test puanlarını dönüştürmedeki iki form arasındaki ilişkinin düzgünleştirilmesine yönelik modellerin eşit yüzdellikli eşitleme için küçük kabul edilen bir gerçek veri setine uygulanması ve sonuçların karşılaştırılması amaçlanmaktadır. Anlaşılacağı üzere, eşit yüzdellikli eşitlemenin daha uygun bir şekilde yapılabilmesi için puan dağılımlarının veya eşitleme fonksiyonlarının düzgün bir fonksiyon olarak evrendeki biçimlerinin kestirilmesi düzgünleştirme metotlarıyla sağlanır. Ayrıca, test puanlarının eşitlenmesinde puan dağılımlarının veya eşit yüzdellikli eşitleme fonksiyonunun düzgünleştirilmesi yöntemleriyle ilgili Türkçe yazılmış hiçbir makaleye literatürde rastlanmaması bu çalışma için ek bir motivasyon kaynağı olmuştur.

Eşit Yüzdellikli Eşitlemede Düzgünleştirme Metotları

Farklı test formlarının eşitlenmesinde eşit yüzdellikli eşitleme kullanıldığında her iki test formunda da aynı yüzdellik sırada olan puanlar eşdeğer sayılır. Shin'ın (2011) ifade ettiği gibi, eşit yüzdellikli eşitleme bir testin eski formundaki bir yüzdellik sıraya karşılık gelen yeni bir test formundaki test puanını belirleme yöntemidir.

“Sınav alan öğrencilerin bir veya daha fazla evrenden örneklendikleri düşünülür, örnekleme hatasından dolayı ham puan dağılımlarının grafiği çizildiğinde genelde düzensiz olduğu görülür.” (Cui & Kolen, 2009, 135). Özellikle örneklem küçük oldukça örnekleme hatasından kaynaklanan randum hata artacağından ham puan dağılımlarındaki bu düzensizlikler daha fazla olur. Örnekleme hatasından kaynaklanan randum hatayı azaltmanın bir yolu da düzgünleştirme metotlarının kullanılmasıdır. Moses & Liu (2011)'nin belirttiği gibi eşitleme araştırmalarında ve pratikte, düzensiz eşitleme fonksiyonlarındansa düzgünleştirilmiş eşitleme fonksiyonlarının tipik olarak daha uygun olduğu varsayılır. Eşit yüzdellikli eşitlemede kullanılan öndüzgünleştirme ve sondüzgünleştirme olarak adlandırılan iki tür düzgünleştirme metodu vardır. Öndüzgünleştirmede test puanlarına ait dağılımlar düzgünleştirilirken sondüzgünleştirmede eşit yüzdellikli eşdeğerler düzgünleştirilir. Dağılımları düzgünleştirmede dağılımların kestirilmesindeki uygunluk düzeyinin belirlenmesi çok önemlidir. Düzgünleştirilmiş dağılımların uygunluğuyla ilişkili olan önemli bir ölçüt

moment korunumudur (Kolen & Brennan, 2014, 70). Aşağıda öndüzgünleştirme metodu olarak beta4 ve polinomial loglineer metotları ve sondüzgünleştirme olarak kübik eğri metodu Kolen & Brennan'dan (2014) özetlenerek sunulmuştur.

Lord'un Beta4 Öndüzgünleştirme Metodu

Lord (1965) gerçek puanı kestirmede kullanılan beta4 modelini geliştirmiştir. Bu metot güçlü gerçek puan metodu olarak da adlandırılır. "Beta4 yönteminin geliştirilmesinde, evrendeki gerçek puanlar doğru cevap oranı dağılımı için bir parametrik form varsayılır, $\psi(\tau)$. Ayrıca verilen gerçek puanın gözlenen puan dağılımı için bir koşullu parametrik form varsayılır, $f(x|\tau)$. O halde gözlenen puan dağılımı aşağıdaki gibi yazılır" (Kolen & Brennan, 2014, 72),

$$f(x) = \int_0^1 f(x|\tau)\psi(\tau)d\tau$$

Lord'un(1965) önerdiği beta4 metodunda, gerçek puan dağılımı, $\psi(\tau)$ 4-parametrelili beta olarak varsayılır (Kolen & Brennan, 2014, 72). Yukardaki eşitliğin kullanımı sonucu oluşan puan dağılımı, $f(x)$, model sayıtlarıyla birleşimiyle ancak tanımlanan 4-parametrelili beta bileşik binom dağılımı veya beta4 dağılımı olarak adlandırılır (Kolen & Brennan, 2014, 73).

"Lord(1965) momentler metoduyla, bu dağılımı ve ilişkili gerçek puan dağılımını kestirmek için bir yöntem önerdi. Bu kestirme yöntemi madde sayısını, örneklem dağılımının ilk 4 momentini (ortalama, standart sapma, kayışıklık, basıklık) ve Lord'un "k" olarak isimlendirdiği bir parametreyi kullanır. Lord'un "k" değeri direkt olarak α güvenilirlik katsayısından kestirilir" (Kolen & Brennan, 2014, 73). Modelin uyumu uydurulan dağılımla örneklem dağılımlarının merkezi momentleri ve grafikler karşılaştırılarak değerlendirilir. Model uyumluluğunu belirlemede istatistiksel metotlar da kullanılabilir (Kolen & Brennan, 2014, 73).

Polynomial Loglineer Öndüzgünleştirme Metodu

Polinomial loglineer metot aşağıdaki dağılım formunda bir model uydurur (Kolen & Brennan, 2014, 70).

$$\log[N_x f(x)] = w_0 + w_1 x + w_2 x^2 + \dots + w_c x^c$$

Bu eşitlikte yoğunluğun log'u C'nin en düşük polinomial derecesi olarak ifade edilir. Örneğin C=2 ise, $\log[N_x]f(x) = w_0 + w_1 x + w_2 x^2$ olur ve model 2. dereceden polinomial denklemdir. Buradaki w parametreleri en çok olabilirlik metoduyla kestirilebilir. Yukarıdaki eşitlikte logaritmanın kullanılması, loglineer modeller için toplanabilirliği sağlar (Kolen & Brennan, 2014, 71).

Polinomial loglineer öndüğünlüştürme metodunda, farklı C deęerleri için yapılan düğünlüştürmelerin uyumluluęu ki kare uyum iyilięi istatistięiyle belirlenir. Ayrıca düğünlüştürmede C ve C+1 parametrelili modeller arasındaki farklılıęın manidarlıęı da ki kare farklılık testiyle belirlenir. Çoklu anlamlılık testleri içerięinden dolayı, bu metotlar bir düğünlüştürme derecesi seçmede merkezi momentlerin incelenmesini ve önceki tecrübelerle grafiksel incelemeyle birlikte kullanılmalıdır (Kolen & Brennan, 2014, 72). Loglineer modeller süreksiz dağılımların örneklem momentleriyle tanımlanabilen üstel ailelerinin bir sınıfıdır (Moses & von Davier 2006). Holland & Thayer'ın (1987) polinomial loglineer metodu daha ayrıntı ve örnekler açıklamışlardır.

Kübię eğri Sondüğünlüştürme Metodu

Sondüğünlüştürmede, öncelikle hiçbir düğünlüştürme yapmadan test formlarından elde edilen puanlardan faydalarak eşit yüzdellięli eşitleme yapılır. Sonra eşit yüzdellięli eşitlemenin uygulanmasıyla elde edilen dönüştürülmüş puanlarda düğünlüştürme yapılır. Anlaşılacağı üzere, sondüğünlüştürmede, test puanlarının dağılımı deęil, dönüştürülen puanların düğünlüştürülmesi yapılır. "Sondüğünlüştürme metotları, eşit yüzdellięli ilişkiye bir eğri uydurur. Sondüğünlüştürme metotlarını uygularken, düğünlüştürülmüş ilişkiler gözlenen ilişkilerden çok fazla farklılaşmadan düğünlü görünelidir. Tanımlanacak metot Kolen (1984) tarafından sunulmuştur ve Reinsech (1967)'in tanımladığı kübię eğri düğünlüştürme ara deęerlerini kullanır. Ara deęerlerin

kestiriminde polinomlar yerine kübik eğri ara değerler kullanılır. Çünkü kübik eğri aradeğerler daha fazla esneklik sağlıyor görülmektedir, x_i tamsayısı için aradeğer fonksiyonu aşağıdaki gibidir.” (Kolen & Brennan, 2014, 80-82).

$$\tilde{d}_y(x) = v_{0i} + v_{1i}(x - x_i) + v_{2i}(x - x_i)^2 + v_{3i}(x - x_i)^3$$

Bu eşitlikte yer alan v_{0i}, v_{1i}, v_{2i} ve v_{3i} ağırlıkları her tam puan için kestirilmekte ve her puan için farklılaşmaktadır. Anlaşılacağı üzere her tam puan arasında tanımlanan farklı bir kübik eşitlik vardır. “Kübik eğri sondüzgünleştirme fonksiyonu, puan noktaları üzerinde, minimize etmeyi amaçlar, minimum eğrilik düzeyi ve aşağıdaki formüldeki S parametresine yönelik kısıtlamayı sağlar:

$$\frac{\sum_{i=low}^{high} \left[\frac{\hat{d}_y(x_i) - \hat{e}_y(x_i)}{\hat{s}_e[\hat{e}_y(x_i)]} \right]^2}{x_{high} - x_{low} + 1} \leq S$$

Bu eşitlikte toplam, eğrinin uydurulduğu tam puan noktalarında yapılır. $\hat{s}_e[\hat{e}_y(x_i)]$ terimi eşit yüzdellikli eşitlemenin kestirilen standart hatasıdır. Eşitlemenin standart hatası, standardize edilmiş düzgünleştirilen ve düzgünleştirilmeyen puanlar arasındaki farkları kullanır. Standart hatanın kullanımı standart hatanın küçük olduğu durumda düzgünleştirilmiş ve düzgünleştirilmemiş ilişkiler daha yakın olur ve standart hata büyük olduğunda ise daha uzak olur. S parametresi ($S \geq 0$) araştırmacı tarafından belirlenir ve düzgünleştirmenin derecesini kontrol eder. Hanson, Zeng & Colton'un (1994) belirttiği gibi S parametresinin sıfır olarak seçilmesi eşit yüzdellikli eşitleme fonksiyonunu gözlenen dağılımı hiç değiştirmeden eşitlemeyi yapar, buna rağmen S parametresinin çok büyük seçilmesi (999 gibi) eşitleme fonksiyonunun lineer bir fonksiyon olmasıyla sonuçlanır. Segall (1997) tarafından da vurgulandığı gibi eşitleme işlemlerinde uygulanacak düzgünleştirme derecesine dikkatli bir şekilde karar vermek gerekir, çünkü eşitleme diğer bir hata türü olan sistematik hatayı da içerir.

Problem Cümlesi

Eşit yüzdellikli eşitleme metoduyla eşdeğer test formlarından elde edilen puanlarının eşitlenmesinde beta4 ve polinomial loglineer öndüzgünleştirme ve kübik eğri sondüzgünleştirme metotlarının küçük örneklem için uygunluk düzeyleri nasıldır?

Alt problemler

1. Eşit yüzdellikli eşitlemede beta4 öndüzgünleştirme metodunun küçük örneklem için uygunluk düzeyi nasıldır?
2. Eşit yüzdellikli eşitlemede Eşit yüzdellikli eşitlemede polinomial loglineer öndüzgünleştirme metodunun küçük örneklem için uygunluk düzeyi nasıldır?
3. Eşit yüzdellikli eşitlemede kübik eğri sondüzgünleştirme metodunun küçük örneklem için uygunluk düzeyi nasıldır?

YÖNTEM**Araştırma Deseni**

Araştırmada test eşitlemede veri toplama deseni olarak dengelenmiş tek grup deseni kullanılmıştır. Özellikle denek sayısı az olduğunda ve test formları çok uzun olmadığında dengelenmiş tek grup deseninin kullanımı önerilmektedir. “Test eşitlemede geliştirilen yöntemlerin uygun çalışması için çoğu durumda her test formunun en az 1000 bazı teknikler için 10000 kadar öğrenciye uygulanması gerekmektedir. Öğretmenlerin sınıf ortamında uyguladıkları test formları için bu genişlikte örneklem büyüklüğüne sahip olma olasılığı sıfıra yakın veya çok çok düşük olur. Bu nedenle her iki formu da aynı öğrencilere uygulayarak test formlarının uygulandığı öğrenci sayısının random grup desenine göre iki kat fazla olması sağlanmıştır. Dengelenmiş tek grup deseninde hazırlanan Form X ve Form Y random olarak ikiye bölünen öğrenci gruplarına farklı sıralarda uygulanır. Öğrencilerin yaklaşık olarak yarısı ilk olarak Form X’i alırken ikinci olarak Form Y’yi alır. Öğrencilerin diğer yarısı ise tam tersine ilk önce Form Y’i alırken ikinci olarak Form X’i alırlar. Böylece öğrencilere farklı iki sırada her iki test formu da uygulanmış olunur.

Çalışma Grubu

Araştırmada kullanılan iki test formu toplam 233 öğrenciye uygulanmıştır. Çalışma grubu, Gazi Üniversitesi Gazi Eğitim Fakültesinde 2013-2014 öğretim yılı ikinci dönem eğitimde ölçme ve değerlendirme dersini araştırmacıdan alan 233 öğrencilerden oluşmaktadır.

Verilerin Toplanması

Araştırmanın verileri eğitimde ölçme ve değerlendirme dersinin sorumlusu olan araştırmacı tarafından paralel olarak hazırlanmış olan iki test formlarının uygulanmasıyla elde edilmiştir. Araştırmacı önceki yıllarda hazırlamış ve madde analizlerinin yapılmış olduğu soru bankasından faydalanarak, aynı 20 hedef davranışı ölçen iki paralel test formu hazırlamıştır. Anlaşılacağı gibi, araştırmacının soru bankasındaki test maddelerine ait önceki yıllardaki uygulamalardan elde edilen madde istatistiklerinden faydalanarak aynı istatistiksel özelliklere sahip olması beklenen iki eşdeğer form hazırlamıştır. Böylece test formları hem içeriksel olarak hem de istatistiksel özellikleri olarak eşdeğer olacak şekilde hazırlanmıştır. Hazırlanan test formları dengelenmiş tek grup deseni kullanılarak 233 öğrenciye uygulanmıştır. Böylece, hem Form X hem de Form Y'nin final sınavını alan 233 öğrenciye uygulanmasıyla araştırma verileri elde edilmiştir. Bu araştırmadaki örneklem test eşitleme uygulamaları için küçük örneklem olarak nitelendirilmektedir. N=250 veya daha az olduğunda eşitleme çalışmalarında örneklem büyüklüğü çok küçük olarak kabul edilir (Wang & Kolen, 1996).

Verilerin Analizi

Verilerin analizinde SPSS (ver. 22), SAS (ver. 9.1.3), Rage-rgequate_wg(v3.15) ve Equating-Error_wg(v2.0) programları kullanılmıştır. Rage-rgequate_wg(v3.15) ve Equating-Error_wg(v2.0) programları

<http://www.education.uiowa.edu/centers/casma/computer-programs> web sayfasından indirilmiştir.

BULGULAR

Test Formlarının Ortalamalarına Yönelik Sıra Etkisi

Araştırmada kullanılan dengelenmiş tek grup deseninde formların alışı sırasının aritmetik ortalamalar üzerinde fark yapıp yapmadığını belirlemek için Form X ve Form Y'nin farklı sıralarda uygulanmıştır. Elde edilen aritmetik ortalamalar arasındaki farklılıkların hem tüm bölümlerde hem de genel olarak manidar olup olmadıkları t testi ile test edilmiş ve hem bölümler bazında hemde genel olarak hiçbir farklılık manidar bulunmamıştır ($p > .05$). Bu bulgular net olarak formların uygulanışında aritmetik ortalama üzerinde bir sıra etkisinin olmadığını göstermektedir. Bu nedenle, bundan sonraki analizlerde formların uygulanış sırası dikkate alınmayacaktır.

Form X ve Form Y Puanlarının Aritmetik Ortalamaları

Her iki form da aynı gruba uygulandığından, grup kendi kendinin kontrolünü sağlamaktadır. Bu nedenle formların ortalamaları arasındaki farkın grup farklılığından kaynaklanmasının söz konusu olamayacağından, random grup desenlerde olduğu gibi, test formlarının aritmetik ortalamaları arasında elde edilecek farklılıklar tamamen test formların güçlük düzeylerine atfedilecektir. Test formlarının alışı sırası dikkate alınmadan elde edilen aritmetik ortalamalar ve bu ortalamalar arasındaki farklılığın manidarlığına yönelik t testi yapılmıştır. Form X için aritmetik ortalama 12.19 iken Form Y için 13.90 bulunmuştur. Bu iki aritmetik ortalama arasındaki fark manidar bulunmuştur ($t=9.096$ ve $p=0.000$). Benzer şekilde test formlarının uygulandığı farklı bölümlere göre elde edilen aritmetik ortalamalar arasındaki farklılıkların manidarlıkları da test edilmiş ve test formlarının aritmetik ortalamaları arasındaki fark testin uygulandığı tüm bölümlerde manidar bulunmuştur ($p < .05$). Yukarıda belirtildiği gibi, her iki form da aynı gruba uygulandığından, grup kendi kendinin kontrolünü sağlamaktadır. Bu nedenle, formların aritmetik ortalamaları arasındaki fark tamamen

test formlarının güçlük düzeylerine atfedilecektir. Tüm bölümlerde ve genelde Form Y, Form X'den daha kolay bulunmuştur. Bu bulgular, Form Y'yi alan öğrencilerin bir avantaj sağladığını gösterir. Form X ve Form Y eşdeğer olarak kullanılabilmesini sağlamak, yani formların güçlük düzeyleri arasındaki farklılığı giderebilmek için test eşitlemenin uygulanması gerekir.

Veri Seti İçin Öndüzgünleştirme Sonuçları

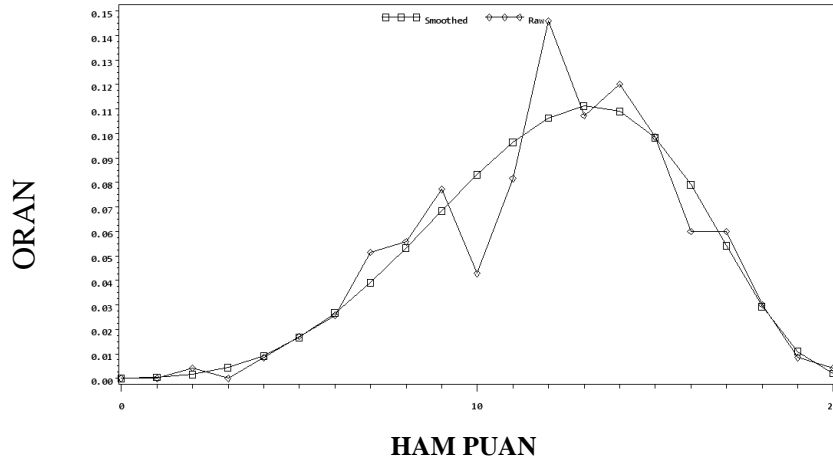
Bu çalışmada Form X puanları Form Y puanlarına eşitlendi. Tüm Form X ve Form Y puanları için gözlenen frekans oranlarıyla beta4 ve farklı C düzeylerinde polinomial loglineer öndüzgünleştirme uygulanmıştır. Öndüzgünleştirme metodunu uygulamada ilk adım ham puan dağılımlarını uyumlu yapmaktır. Aşağıda öndüzgünleştirme sonuçları sunulmuştur.

Beta4 Öndüzgünleştirme Sonuçları

Form X için Beta4 Öndüzgünleştirme Sonuçları

Şekil 1'de grafiksel analizi amacıyla, Form X için beta4 öndüzgünleştirme metoduyla düzgülendirilmiş dağılımlar düzgülendirilmemiş dağılımlarla birlikte sunulmuştur.

Form X İçin Beta 4 Öndüzgünleştirme Dağılımı

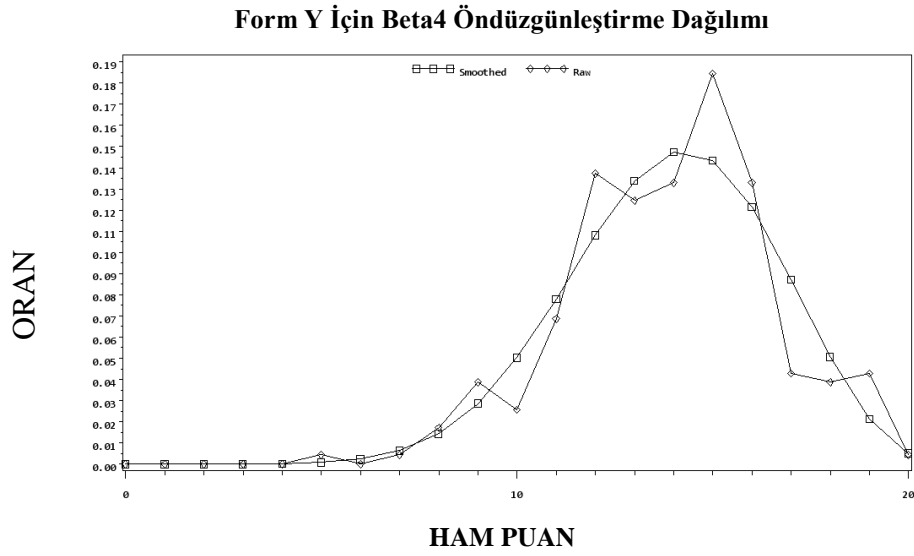


Şekil 1. Form X Puanlarına ait Dağılım İçin Beta4 Metoduyla Öndüzgünleştirme (smooted=düzgünleştirilmiş puan dağılımı ve raw=ham puan dağılımı)

Şekil 1’de görüldüğü gibi beta4 öndüzgünleştirme metodunda düzgün bir eğri elde edilmiştir. Düzgünleştirilmiş dağılımla deneysel dağılım arasında Form X’in 10 ve 12 puanları için farklılık fazla görülmektedir. Buna rağmen grafiksel analiz, beta4 öndüzgünleştirme metodunun genelde Form X puan dağılımına uyumlu olduğunu göstermektedir.

Form Y için Beta4 Öndüzgünleştirme Sonuçları

Şekil 2’de grafiksel analiz amacıyla, Form Y için beta4 öndüzgünleştirme metoduyla düzgünleştirilmiş (smoothed) dağılımlar düzgünleştirilmemiş (raw) dağılımlarla birlikte sunulmuştur.



Şekil 2. Form Y Puanlarına ait Dağılım İçin Beta4 Metoduyla Öndüzgünleştirme (smoothed=düzgünleştirilmiş puan dağılımı ve raw=ham puan dağılımı)

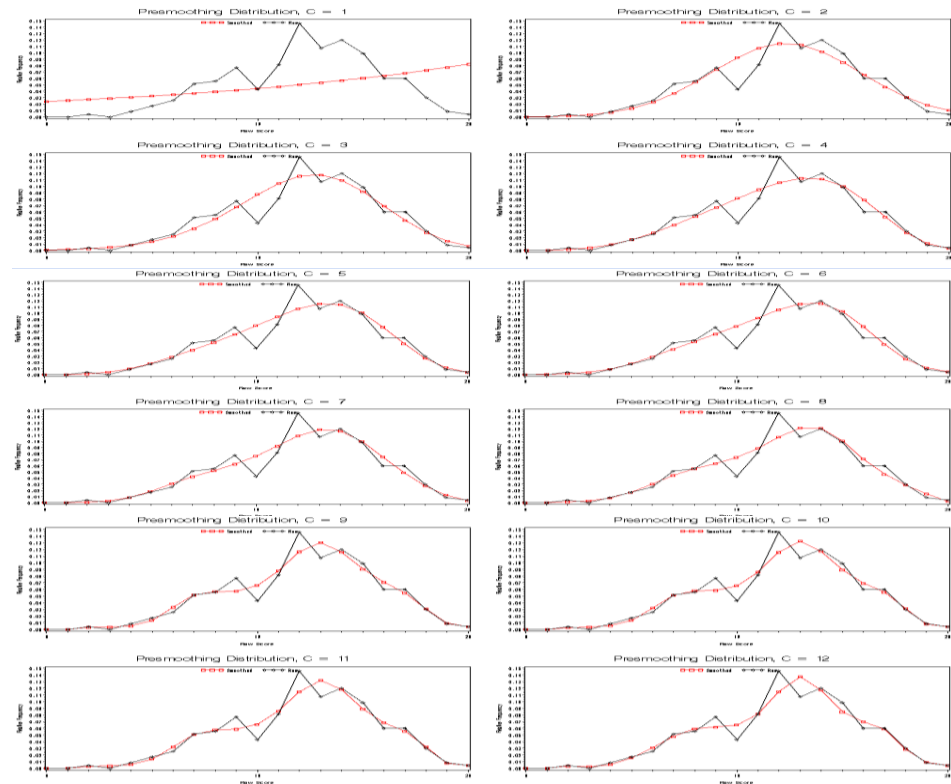
Şekil 2’de görüldüğü gibi beta4 öndüzgünleştirme metodunda düzgün bir eğri elde edilmiştir. Düzgünleştirilmiş dağılımla deneysel dağılım arasında en büyük farklılık Form Y’nin 15 ham puanı için görülmektedir. Buna rağmen grafiksel analiz beta4

öndüzgünleştirme metodunun Form Y puan dağılımına genelde uyumlu olduğunu göstermektedir.

Polinomial Loglineer Öndüzgünleştirme Sonuçları

Form X İçin Polinomial Loglineer Öndüzgünleştirme Sonuçları

Şekil 3'te grafiksel analiz amacıyla, Form X için C=1'den C=12'ye kadar farklı C düzeyleri için polinomial loglineer öndüzgünleştirme metoduyla düzgünleştirilmiş dağılımlar düzgünleştirilmemiş dağılımlarla birlikte sunulmuştur.



Şekil 3. Form X Puanlarına ait Dağılım İçin Farklı C Düzeylerinde Polinomial Loglineer Metoduyla Öndüzgünleştirme

Şekil 3'te görüldüğü gibi Form X puan dağılımı için $C=1$ düzeyi hariç diğer tüm C düzeylerinde polinomial loglineer öndüzgünleştirme metodunda deneysel dağılımla uyumlu eğriler elde edilmiştir. Ancak $C \geq 8$ için elde edilen eğriler yeterince düzgün gözükmemektedir. Bu durumda Form X için grafiksel analiz bize $C=2, C=3, C=4, C=5, C=6$ ve $C=7$ düzeylerinde polinomial log lineer öndüzgünleştirmenin uygun olduğunu göstermektedir.

Form Y İçin Polinomial Loglineer Öndüzgünleştirme

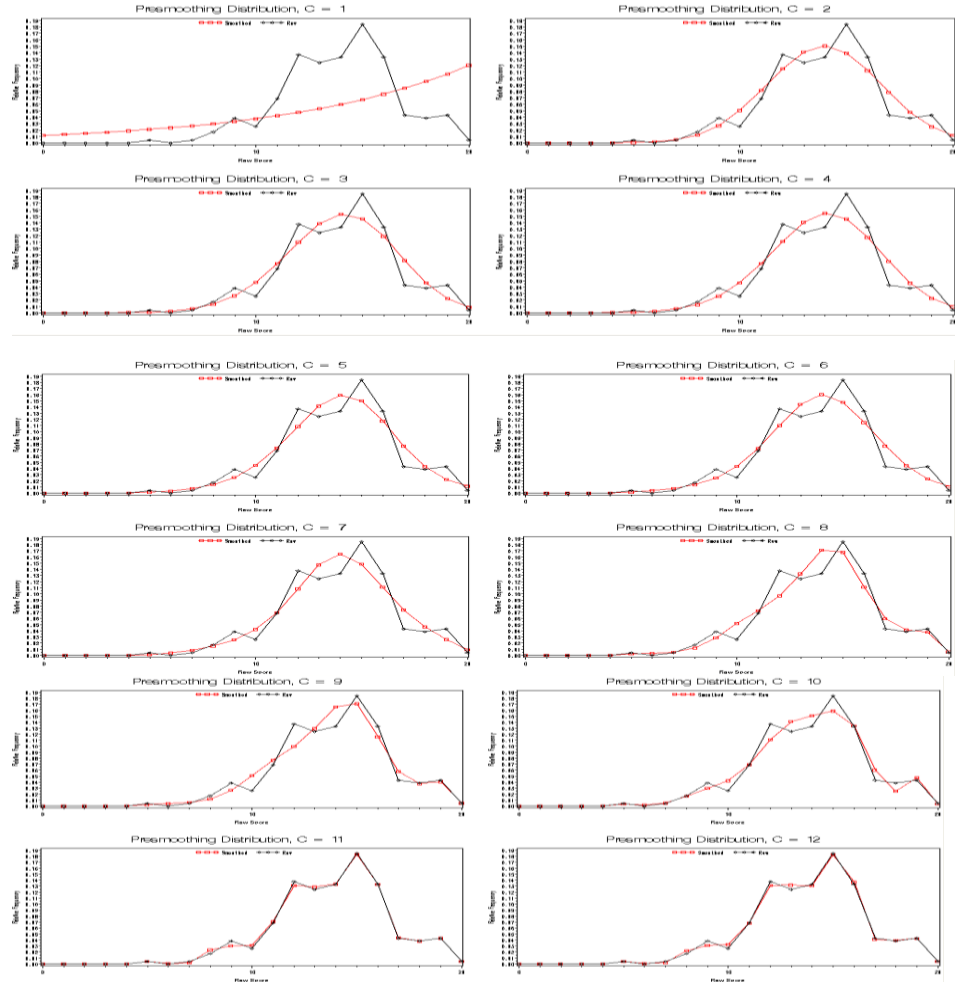
Şekil 4'te grafiksel analiz amacıyla, Form Y için $C=1$ 'den $C=12$ 'ye kadar farklı C düzeyleri için polinomial loglineer öndüzgünleştirme metoduyla düzgünleştirilmiş dağılımlar düzgünleştirilmemiş dağılımlarla birlikte sunulmuştur.

Şekil 4'te görüldüğü gibi Form Y puan dağılımı için de $C=1$ düzeyi hariç diğer tüm C düzeylerinde polinomial loglineer öndüzgünleştirme metodunda deneysel dağılımla uyumlu eğriler elde edilmiştir. Ancak $C \geq 8$ için elde edilen eğriler yeterince düzgün gözükmemektedir. Bu durumda Form Y için grafiksel analiz bize $C=2, C=3, C=4, C=5, C=6$ ve $C=7$ düzeylerinde polinomial log lineer öndüzgünleştirmenin uygun olduğunu göstermektedir.

Form X ve Form Y İçin Yapılan Öndüzgünleştirmeler için Moment Korunumları ve Uyum Testleri

Düzgünleştirme sonucu eşitlenen puanların moment korunumu incelenmiştir. Holland & Thayer'ın (2000) belirttiği gibi uydurulan bir dağılımın uygunluğunun değerlendirilmesinde hem grafiksel hem de bazı test istatistikleri faydalı olur. Öndüzgünleştirme sonucu Form X puanları eşit yüzdelliği eşitlemeyle Form Y puanlarına dönüştürüldüğünde Form Y ham puanlarının momentlerinin korunması gerekir. Livingston'un (1992) belirttiği gibi örnekler çok küçük olduğunda düzgünleştirilen dağılım, gözlenen dağılımın sadece aritmetik ortalama ve standart sapmasının korunumunu sağlayabilir. Eşit yüzdelliği eşitlemeyle dönüştürülmüş Form

Y puanlarına ait momentlerin özellikle ilk iki momentin aynen korunması veya çok yakın değerlerin elde edilmesi gerekir.



Şekil 4. Form Y Puanlarına ait Dağılım İçin Farklı C Düzeylerinde Polinomial Loglineer Metoduyla Öndüzgünleştirme

Form Y ham puanları için ilk dört moment (aritmetik ortalama, standart sapma, kayışıklık ve basıklık) sırasıyla 13.9013; 2.6047; -0.2649 ve 3.1536 bulunmuştur.

Beta4 öndüzgünleştirmeye elde edilen eşit yüzdelli eşitleme puanlarının ilk dört momenti sırasıyla, 13.9004; 2.5954; -0.2689 ve 2.7895 bulunmuştur. Bu bağlamda beta4 öndüzgünleştirmesiyle yapılan eşitlemede elde edilen puanların ilk üç momenti Form Y ham puanlarına ait momentlere çok yakın neredeyse aynıdır. Sadece dördüncü moment olan basıklık çok ciddi olmayak şekilde farklılaşmaktadır. Bu moment korunumu sonuçları da beta4 öndüzgünleştirmesiyle yapılan eşit yüzdelli eşitlemenin oldukça uygun olduğunu göstermektedir.

Polinomial loglineer dönüştürmeyle yapılan eşit yüzdelli eşitlemenin uygunluğuna bakıldığında tüm C düzeylerinde ilk moment Form Y ham puanlarındakine çok benzer veya aynı bulunmuştur. Polinomial loglineer dönüştürmede C=1 düzeyi hariç diğer tüm C düzeylerinde ikinci moment Form Y ham puanlarındakine çok benzer bulunmuştur. Polinomial loglineer dönüştürmede C=1 ve C=2 düzeyleri hariç diğer tüm C düzeylerinde üçüncü ve dördüncü momentler Form Y ham puanlarındakine çok benzer bulunmuştur. Bu bilgiler ışığında dönüştürmelerde ilk iki momentin korunmasının çok önemli olması ve üçüncü ve dördüncü momentlerdeki küçük farklılıkların dikkate alınmamasıyla C=1 düzeyi hariç diğer tüm C düzeyleri için polinomial loglineer öndüzgünleştirmeye eşitlemenin uygun olduğu sonucuna ulaşılmıştır.

Beta4 ve aşamalı olarak farklı C düzeylerinde elde edilen öndüzgünleştirme yapılan dağılımların model uyum ve model farklılıklarına yönelik ki kare manidarlık testleri yapılmıştır. Ki kare değerlerinin örneklem verisinden farklılığına yönelik hipotez testi alfa 0.05 düzeyinde beta4 öndüzgünleştirme hem X ($\chi^2 = 15.396$) hem de Y ($\chi^2 = 23.768$) test puanları için uyumlu bulunmuştur. Polinomial loglineer öndüzgünleştirmelerde hem Form X hem de Form Y için $C \geq 2$ olduğunda ki kare değerleri kritik değerden küçük olduğundan C=1 hariç tüm polinomial loglineer modellerin alfa 0.05 düzeyinde uyumlu olduğu bulunmuştur.

Moses'un (2008) belirttiği gibi loglineer modeller için birçok istatistiksel test mevcuttur. Bu çalışmada loglineer modeller için istatistiksel test olarak C derecesi bakımından aşamalı olarak sunulan polinomial loglineer modellerin farklılıklarının test

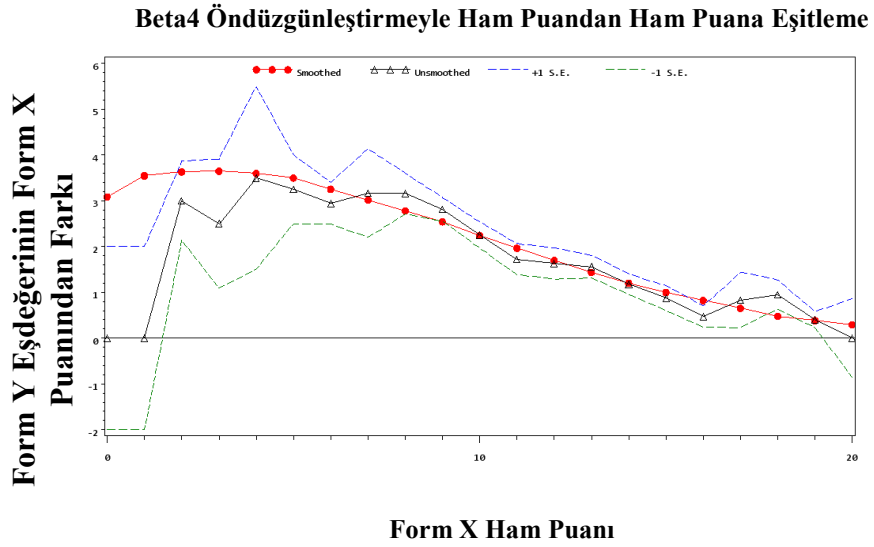
edilmesine yönelik olarak $\chi_c^2 - \chi_{c+1}^2$ fark istatistiğine bakılmıştır. Alfa 0.05 düzeyinde bir serbestlik derecesi farklılık gösteren hiyerarşik modeller için ki kare kritik değeri olan 3.84 dikkate alındığında, Form Y için C=2, C=3, C=4, C=5, C=6, C=8, C=11 manidar farklılık göstermemektedir. Form X puanları için ise $C \geq 2$ manidar değişim göstermemektedir. Manidar farklılık göstermeyen bu modellerin hepsi uygun modeller olarak düşünülür. Bu durumda en düşük dereceden polinomial model en uygun model olarak seçilir. Bu durumda grafiksel analiz sonuçlarında dikkate alınarak, polinomial loglineer öndüzgünleştirmede hem X hem de Y için en uygun modelin C=2 olan yani ikinci dereceden polinomial model olduğu belirlenmiştir.

Aşağıda uygun olan öndüzgünleştirme modelleri olan beta4 ve C=2 için polinomial loglineer öndüzgünleştirme metotlarıyla eşit yüzdellikli eşitlemede için elde edilen fark grafikleri sırasıyla Şekil 5 ve Şekil 6 sunulmuştur.

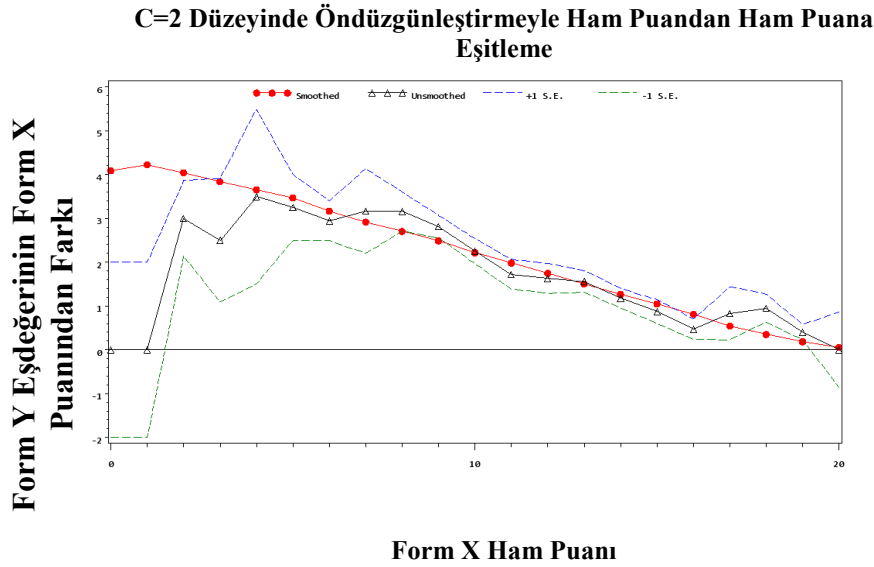
Beta4 ve Polinomial Loglineer Öndüzgünleştirme Metoduyla C=2 için Düzgünleştirilmiş ve Düzgünleştirilmemiş Eşit Yüzdellikli Eşitleme İçin Fark Grafikleri

Beta4 öndüzgünleştirmeye Form X puanları Form Y puanlarına eşit yüzdellikli eşitlemeyle eşitlendi. Şekil 5'te beta4 öndüzgünleştirmesinden sonra yapılan eşit yüzdellikli eşitleme için elde edilen fark puanları ve düzgünleştirilmemiş eşitlenmiş puanlar için ± 1 birim standart hata bandı grafiksel olarak sunulmuştur.

C=2 için polinomial loglineer öndüzgünleştirmeye Form X puanları Form Y puanlarına eşit yüzdellikli eşitlemeyle eşitlendi. Şekil 6'da C=2 için polinomial loglineer öndüzgünleştirmesinden sonra yapılan eşit yüzdellikli eşitleme için elde edilen fark puanları ve düzgünleştirilmeden eşitlenmiş puanlar için ± 1 birim standart hata bandı grafiksel olarak sunulmuştur.



Şekil 5. Beta4 Öndüzgünleştirmeye Eşit Yüzelikli Eşitleme Fark Grafiği (smoothed=düzgünleştirilmiş, unsmoothed=düzgünleştirilmemiş ve S. E.=Standart hata)



Şekil 6. C=2 için Polinomial Loglineer Öndüzgünleştirmeye Eşit Yüzelikli Eşitleme Fark Grafiği (smoothed=düzgünleştirilmiş, unsmoothed=düzgünleştirilmemiş ve S. E.=Standart hata)

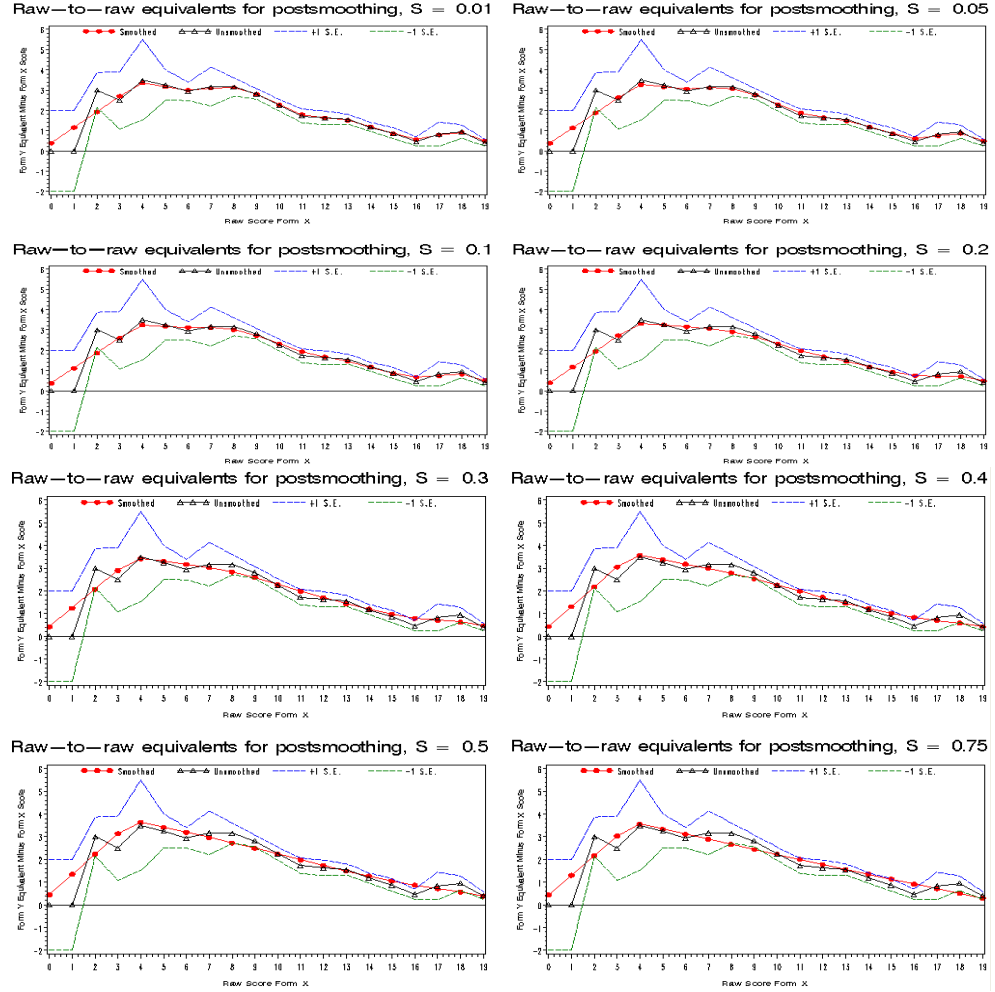
Form X ve Form Y puanlarına ait göreceli frekans dağılımları beta4 ve C=2 için polinomial loglineer öndüzgünleştirme metotlarıyla düzgünleştirildi. Daha sonra Form X puanları Form Y puanlarına eşit yüzdelikli eşitleme kullanılarak eşitleme yapıldı. Eşit yüzdelikli eşitleme ilişkileri beta4 öndüzgünleştirme için Şekil 5'te ve C=2 için polinomial loglineer öndüzgünleştirme için Şekil 6'da sunulmuştur. Kolen ve Brennan (2014)'ın belirttiği gibi eşitlemenin uygunluğu için algılanabilir bir standart olarak, düzgünleştirilmiş ilişkinin ağırlıklı olarak standart hata sınırlarında yayılması kriteri kullanılabilir. Bu nedenle Şekil 5 ve Şekil 6'da eşitleme fonksiyonu için standart hata sınırları da sunulmuştur. Şekil 5'te özdeş eşitleme, dağılımları düzgünleştirmeden yapılan eşitleme ve beta4 öndüzgünleştirmeye yapılan eşitleme ilişkileri sunulmuştur. Ayrıca, düzgünleştirilmemiş eşit yüzdelikli eşitleme için ± 1 birim standart hata sınırları sunulmaktadır. Şekil 5'te görüldüğü gibi 0, 1 ve 18 puanları için beta4 öndüzgünleştirme ile elde edilen kestirimler ± 1 birim standart hata sınırlarının dışında kalmaktadır. Şekil 5'te görüldüğü gibi Form X'in 0, 1 ve 18 puanları için eşit yüzdelikli ilişki fonksiyonu ± 1 birim standart hata bandının dışında kalmaktadır. Diğer puanlar için ise eşit yüzdelikli ilişki fonksiyonu ± 1 birim standart hata bandının içinde yer almaktadır. Uç puanlar için frekanslar az olduğundan kestirim hatasının daha fazla olması beklenen bir durumdur. Sonuç olarak, bu puanların standart hata bandının dışında kalmasına rağmen beta4 öndüzgünleştirme metodunun uyumlu olduğu sonucuna ulaşılır.

C=2 ile polinomial loglineer model için de benzer durum söz konusudur. Şekil 6'da görüldüğü gibi 1, 2, 3 ve 18 Form X puanları için elde edilen Form Y kestirimleri ± 1 birim standart hata bandının dışında yer almaktadır. Uç puanlar için frekanslar az olduğundan bu puanlar için kestirim hatasının biraz daha fazla olması normaldir.

Sonuç olarak, hem beta4 hem de polinomial loglineer öndüzgünleştirme metotlarının uygun olduğu görülmektedir.

Küçük Eđri Sondüzgünleştirme Metodu İçin Sonuçlar

Aşağıda Şekil 7’de düzgünleştirilmiş ve düzgünleştirilmemiş fark grafikleri ve düzgünleştirilmemiş eşitlenmiş puanlar için ± 1 birim standart hata bandı grafiksel olarak sunulmuştur.



Şekil 7. Küçük Eđri Sondüzgünleştirmeyle Farklı S Kriterleri İçin Düzgünleştirilmiş ve Düzgünleştirilmemiş Fark Grafikleri

Küçük eđri sondüzgünleştirme için bir istatistiksel test yoktur. Bu nedenle hangi S düzeyinde düzgünleştirmenin uygun olduğunu belirlemede grafiklerin ve momentlerin

incelenmesi önemlidir. Analizde $S=.01$ ile $S=1.00$ arasında 9 farklı S değerine göre sondüzgünleştirme yapıldı. Elde edilen eşit yüzdelikli eşitleme fonksiyonlarına ait bu grafiklerden $S=1.00$ hariç diğer sekiz tanesinin grafikleri Şekil 7’de sunulmuştur. Bu grafikler incelendiğinde, $S=.01$, $.05$, $.10$ ve $.20$ ’de düzgünleştirilmiş eşdeğerler düzgün görülmemektedir. Buna rağmen S düzeyleri için düzgünleştirilmiş eşdeğerler ± 1 birim standart hata bandının içinde. $S \geq .30$ için düzgünleştirilmiş eşdeğerler uygun görülmektedir ve tüm Form X puanları için kestirimler ± 1 birim standart hata bandının içinde. Bu durumda en az düzgünleştirme derecesi olan $S=.30$ en uygun sondüzgünleştirme modeli olarak görülmektedir.

Kübik eğri son düzgünleştirme eşit yüzdelikli eşitleme yapıldıktan sonra elde edilen dönüştürülmüş puanlar üzerinde düzeltme yapar. Dolayısıyla dönüştürülen Form X puanlarıyla Form Y puanları arasındaki ilişkiyi gösteren dağılım düzgünleştirilir.

Form Y ham puanları için ilk dört moment (aritmetik ortalama, standart sapma, kayışıklık ve basıklık) sırasıyla **13.9013**; **2.6047**; **-0.2649** ve **3.1536** olarak bulunmuştu. Bu çalışmanın verileri için grafiksel analiz sonucu en uygun düzeltme parametresi olan $S=.30$ düzeyinde sondüzgünleştirmeyle eşitlenen puanların ilk dört momenti sırasıyla, **13.9006**; **2.6304**; **-0.3499** ve **3.1954** olarak bulunmuştur. Anlaşılaacağı üzere birinci, ikinci ve dördüncü momentler hemen hemen aynıdır, üçüncü moment biraz farklıdır. Bu bulgular $S=.30$ düzeyinde sondüzgünleştirme nin moment korunumunu sağladığını göstermektedir.

Eşitlemenin Bootstrap Standart Hataları

Düzgünleştirme uygulanmadan, ilk üç momentle beta4 öndüzgünleştirme, $C=2$ düzeyinde polinomial loglineer öndüzgünleştirme ve $S=.30$ düzeyinde kübik eğri sondüzgünleştirme için her tam puan düzeyinde 500 tekrar kullanılarak elde edilen eşitlemenin bootstrap standart hataları Equating-Error_wg(v2.0) programı kullanılarak bulunmuştur. Eşitlemenin ortalama standart hataları düzgünleştirme yapılmadığında **0.37208854**; ilk 3 moment dikkate alınarak beta4 öndüzgünleştirmede **0.28237450**; $C=2$ düzeyinde polynomial log lineer eşitlemede **0.29911151** ve $S= 0.30$ düzeyinde kübik

sondüzgünleştirmede **0.33966465** bulunmuştur. Sonuç olarak, bu çalışmada kullanılan küçük örneklem veri seti için beta4 öndüzgünleştirme, polinomial loglineer öndüzgünleştirme ve kübük eğri sondüzgünleştirmeye eşit yüzdelikli eşitleme metotlarından elde edilen eşitlemenin ortalama standart hataları düzgünleştirme yapılmadan elde edilen eşitlemenin ortalama standart hatadan daha az bulunmuştur. Bu çalışmada kullanılan küçük örneklem için göreceli olarak en az ortalama standart hata beta4 öndüzgünleştirme ile yapılan eşitlemede elde edilmiştir. Buna ilave olarak beta4 ve polinomial loglineer öndüzgünleştirmeye eşitleme kübük sondüzgünleştirmeye eşitlemeden daha az ortalama standart hata üretmiştir. Bu bulgu küçük örneklerde öndüzgünleştirme yöntemlerinin sondüzgünleştirmeden daha uygun olduğunu göstermektedir.

TARTIŞMA ve SONUÇ

Bu çalışmada küçük örneklem için eşit yüzdelikli eşitlemede düzgünleştirme metotları olarak kullanılan beta4 ve polinomial loglineer öndüzgünleştirme ve kübük eğri sondüzgünleştirme metotlarının karşılaştırılması amaçlanmıştır. Livingston (1992)'un belirttiği gibi örneklem küçük oldukça örneklemin evrenden farklı olma olasılığı artacağından dolayı örneklem küçük olduğunda eşitlemede düzgünleştirmenin faydası artar. Dengelenmiş tek grup deseninin kullanıldığı çalışmada kullanılan test formlarının farklı sıralarda uygulanmaları sonucu sıra etkisinin olmadığı bulunmuştur.

Bu çalışmada Form X puanları Form Y puanlarına eşitlendi. Çalışmada öndüzgünleştirme ve sondüzgünleştirme metotları uygulanmıştır.

Öndüzgünleştirmeler sonucu elde edilen düzgünleştirilmiş dağılımlar grafiksel olarak incelenmiştir. Grafiksel analiz sonucu, beta4 öndüzgünleştirme metodunda hem Form X hem de Form Y için düzgün eğriler elde edilmiş ve beta4 öndüzgünleştirme metodunun genelde hem Form X hem de Form Y puan dağılımına uyumlu olduğu görülmüştür. Grafiksel analiz sonucu, hem Form X hem de Form Y için C=1'den C=12'ye kadar farklı C düzeyleri için polinomial loglineer öndüzgünleştirme metoduyla

düzgünleştirilmiş dağılımlar elde edilmiştir. Hem Form X hem de Form Y puan dağılımı için grafiksel analiz C=2, C=3, C=4, C=5, C=6 ve C=7 düzeylerinde polinomial loglineer öndüzgünleştirme metodunda deneysel dağılımla uyumlu eğriler elde edilmiştir.

Düzgünleştirilen dağılımların eşit yüzdelliği eşitlemeyle eşitlenmesi sonucu elde edilen eşdeğer Y puanlarının moment korunumları incelenmiştir. Beta4 öndüzgünleştirilmesiyle yapılan eşitlemede elde edilen puanların ilk 3 momentleri Form Y ham puanlarına ait momentlere çok yakın nerdeyse aynı bulunmuştur. Sadece dördüncü moment olan basıklık çok ciddi olmayak şekilde farklılaşmaktadır. Yani Form Y için beta4 öndüzgünleştirme ilk üç örneklem momentini mükemmel şekilde korumuştur. Moment korunumunun incelenmesi sonucu, Form Y için beta4 öndüzgünleştirmeyle yapılan eşit yüzdelliği eşitlemenin oldukça uygun olduğu görülmüştür. Polinomial loglineer öndüzgünleştirmeyle yapılan eşit yüzdelliği eşitleme sonucu elde edilen Y eşdeğer puanlarının Form Y puan dağılımına uygunluğu moment korunumu dikkate alınarak incelendiğinde, tüm C düzeylerinde ilk momentin Form Y ham puanlarındakine çok benzer veya aynı olduğu; C=1 düzeyi hariç diğer tüm C düzeylerinde ikinci momentin Form Y ham puanlarındakine çok benzer olduğu ve C=1 ve C=2 düzeyleri hariç diğer tüm C düzeylerinde üçüncü ve dördüncü momentlerin Form Y ham puanlarındakine çok benzer olduğu görülmüştür. Bu bilgiler ışığında dönüştürmelerde ilk iki momentin korunmasının çok önemli olması ve üçüncü ve dördüncü momentlerdeki küçük farklılıkların dikkate alınmamasıyla C=1 düzeyi hariç diğer tüm C düzeyleri için polinomial loglineer öndüzgünleştirmeyle eşitlemenin moment korunumu ölçütüne göre uygun olduğu sonucuna ulaşılmıştır.

Öndüzgünleştirme metotları için ki kare uyum indekleri belirlenmiştir. Beta4 öndüzgünleştirme hem Form X hem de Form Y puanları için uyumlu bulunmuştur. Polinomial loglineer öndüzgünleştirmeler için ki kare kritik değerleriyle hesaplanan ki kare değerlerine baktığımızda hem Form X hem de Form Y dağılımında $C \geq 2$ için ki kare değerleri kritik değerden küçük olduğundan C=1 hariç tüm polinomial loglineer modeller uyumlu bulunmuştur.

Polinomial loglineer öndüzgünleştirme farklı C derecesi bakımından aşamalı olarak sunulan modellerin farklılıklarının test edilmesine yönelik olarak $\chi^2 - \chi^2_{C+1}$ fark istatistiğine bakılmıştır. Form Y için C=2, C=3, C=4, C=5, C=6, C=8, C=11 manidar farklılık göstermemektedir. Form X puanları için ise $C \geq 2$ için manidar değişim göstermemektedir. Manidar farklılık göstermeyen bu modellerin hepsi uygun modeller olarak düşünülür. Bu durumda en düşük dereceden polinomial model uygun olarak seçilir. Bu durumda grafiksel analiz sonuçları da dikkate alınarak polinomial loglineer öndüzgünleştirmede hem Form X hem de Form Y için en uygun model C=2, yani ikinci dereceden polinomial model olarak belirlenmiştir.

Form X ve Form Y puanlarına ait göreceli frekans dağılımları beta4 ve C=2 için polinomial loglineer öndüzgünleştirme metotlarıyla düzgülendirildi. Daha sonra Form X puanları Form Y puanlarına eşit yüzdelli eşitleme kullanılarak eşitleme yapıldı.

Moses, Deng & Zhang'un (2011) belirttiği gibi eşitlemenin standart hataları düzgülendirme sonuçlarının yorumlanmasında kullanılır. 0, 1 ve 18 puanları için beta4 öndüzgünleştirme ile elde edilen kestirimler ± 1 birim standart hata sınırlarının dışında kalmaktadır. Diğer puanlar için ise eşit yüzdelli ilişki fonksiyonu ± 1 birim standart hata bandının içinde yer almaktadır. Uç puanlar için frekanslar az olduğundan kestirim hatasının daha fazla olması beklenen bir durumdur. Sonuç olarak bu puanların standart hata bandının dışında kalmasına rağmen beta4 metodunun uyumlu olduğu sonucuna ulaşılmıştır.

C=2 ile polinomial loglineer model için 1, 2, 3 ve 18 Form X puanları için elde edilen Form Y kestirimleri ± 1 birim standart hata bandının dışında yer almaktadır. Uç puanlar için frekanslar az olduğundan bu puanlar için kestirim hatasının biraz daha fazla olması normaldir. Sonuç olarak, hem beta4 hem de polinomial loglineer öndüzgünleştirme metotlarının uygun olduğu görülmektedir.

Kübik eğri öndüzgünleştirme, eşit yüzdelli eşitleme yapıldıktan sonra elde edilen dönüştürülmüş puanlar üzerinde düzeltme yapar. Dolayısıyla dönüştürülen Form X puanlarıyla Form Y puanları arasındaki ilişkiyi gösteren dağılım düzgülendirilir. Kübik

eđri sondüzgünleştirme için bir istatistiksel test yoktur. Bu nedenle hangi S düzeyinde düzgünleştirmenin uygun olduğunu belirlemede grafiklerin ve momentlerin incelenmesi önemlidir. Analizde $S=.01$ ile $S=1.00$ arasında 9 farklı S değerine göre sondüzgünleştirme yapıldı. Elde edilen eşit yüzdelikli eşitleme fonksiyonların ait grafikler Şekil 7’de sunuldu. Bu grafikler incelendiğinde, $S=.01$, $.05$, $.10$ ve $.20$ ’de düzgünleştirilmiş eşdeğerler düzgün görülmektedir. Buna rağmen S düzeyleri için düzgünleştirilmiş eşdeğerler ± 1 birim standart hata bandınının içindedir. $S \geq .30$ için düzgünleştirilmiş eşdeğerler uygun görülmektedir ve tüm Form X puanları için kestirimler ± 1 birim standart hata bandının içindedir. Bu durumda en az düzgünleştirme derecesi olan $S=.30$ en uygun sondüzgünleştirme modeli olarak görülmektedir. $S=.30$ düzeyinde kubik eđri sondüzgünleştirmede moment korunumu da sağlanmıştır.

Düzgünleştirme yapılmadan, ilk üç moment için beta4 ve $C=2$ için polinomial loglineer öndüzgünleştirme ve kubik eđri sondüzgünleştirme metotlarından elde edilen eşitlemenin bootstrap standart hataları incelenmiştir. Küçük örneklem verisi için araştırmada kullanılan düzgünleştirme yöntemleri eşit yüzdelikli eşitlemedeki randum hataları azalttığı ve beta4 ve polinomial öndüzgünleştirmenin kubik eđri sondüzgünleştirmeden göreceli olarak daha etkili olduğu bulunmuştur.

Bu çalışmanın bulguları önceki bilimsel araştırmalarla tutarlıdır. Eşit yüzdelikli eşitlemede yapılan düzgünleştirme çalışmalarına yönelik literatürde, beta4 ve loglineer öndüzgünleştirme ve kubik eđri sondüzgünleştirme metotlarının oldukça iyi çalıştığını bulan araştırmalar mevcuttur (Hanson,1991; Kolen, 1991; Hanson, Zeng, & Colton, 1994). Sonuç olarak, bu araştırmada kullanılan üç düzgünleştirme metodu da düzgünleştirmesiz eşit yüzdelikli eşitlemeye göre daha az randum hata üretmiştir. Bu nedenle bu üç düzgünleştirme metodu da etkilidir. Araştırmada kullanılan küçük örneklem verisi için göreceli etkililiğe baktığımızda, düzgünleştirme metotlarından beta4 öndüzgünleştirme polinomial loglineer ve kubik eđri sondüzgünleştirmeden daha etkili bulunmuştur.

KAYNAKLAR

- Cui, Z., & Kolen, M. J. (2009). Evaluation of two new smoothing methods in equating, The cubic b-spline presmoothing method and the direct presmoothing method. *Journal of Educational Measurement, 46*(2), 135–158.
- de Boor, C. (1978). *A practical guide to splines* (Applied Mathematical Sciences, Vol. 27). New York, Springer.
- Fairbank, B. A. (1987). The use of presmoothing and postsmoothing to increase the precision of equipercentile equating. *Applied Psychological Measurement, 11*, 245–262.
- Hanson, B. A. (1991). *Method of moments estimates for the four-parameter beta compound binomial model and the calculation of classification consistency indexes*. Iowa City, IA, ACT. (Research Report 91–5)
- Hanson, B. A., Zeng, L., & Colton, D. (1994). *A comparison of presmoothing and postsmoothing methods in equipercentile equating*. Iowa City, IA, ACT. (Research Report 94–4).
- Holland, P. W., & Thayer, D. T. (1987). *Notes on the use of log-linear models for fitting discrete probability distributions*. Princeton, NJ, Educational Testing Service. (Technical Report 87–79)
- Holland, P. W., & Thayer, D. T. (2000). Univariate and bivariate loglinear models for discrete test score distributions. *Journal of Educational and Behavioral Statistics, 25*, 133–183.
- <http://www.education.uiowa.edu/centers/casma/computer-programs>.
- Kolen, M. J. (1984). Effectiveness of analytic smoothing in equipercentile equating. *Journal of Educational Statistics, 9*, 25–44.
- Kolen M.J., Brennan R.L. (2014). *Test Equating, Scaling, and Linking Methods and Practices*. (3th edition) New York, NY, Springer.

- Kolen, M. J. (1991). Smoothing methods for estimating test score distributions. *Journal of Educational Measurement*, 28, 257–282.
- Livingston, S. A. (1992). *Small-sample equating with loglineer smothing*. Princeton, NJ, Educational Testing Service. (Research Report 142).
- López, F. Olson, A. & Bansal N. (2011). Creating composite age groups to smooth percentile rank distributions of small samples. *Journal of Psychoeducational Assessment* 29(2) 171–183.
- Moses, T., & von Davier, A. A. (2006). *A SAS macro for loglinear smoothing, Applications and implications*. Princeton, NJ, Educational Testing Service. (Research Report 06–05)
- Moses, T. (2008). *An evaluation of statistical strategies for making equating function selections*. Princeton, NJ, Educational Testing Service. (Research Report 08–60)
- Moses, T., Deng, W. & Zhang Y. (2011). Two approaches for using multiple anchors in NEAT equating, A description and demonstration. *Applied Psychological Measurement* 35(5), 362–379.
- Moses, T., & Liu, J. (2011). *Smoothing and equating methods applied to different types of test score distributions and evaluated with respect to multiple equating criteria*. Princeton, NJ, Educational Testing Service. (Research Report 11–20)
- Reinsch, C. H. (1967). Smoothing by spline functions. *Numerische Mathematik*, 10, 177–183.
- Segall, D. O. (1997). Equating the CAT-ASVAB. In W. A. Sands, *Computerized adaptive testing, From inquiry to operation* (pp. 181 - 198). Washington DC, American Psychological Association.
- Shin S. (2011). *A comparison of Van der Linden's conditional equipercentile equating method with other equating methods under the random groups design*. University of Iowa, Dissertation, <http://ir.uiowa.edu/etd/1263> (22 Ocak 2015).

- Wang, T., & Kolen, M. J. (1996). A quadratic curve equating method to equate the first three moments in equipercentile equating. *Applied Psychological Measurement*, 20, 27–43.
- Wang, T. (2006). Standard Errors of Equating for Equipercentile Equating with Log-Linear Pre-Smoothing using the Delta Method. CASMA (Research Report 14).

SUMMARY

In equipercentile equating, test scores distribution for dichotomously scored test items of examinees, due to sampling error raw score distributions usually appear irregular. The main reason for this irregularity is the sampling error. In particular, the sampling frequencies of some test raw scores have little or zero frequencies. In this case, an attempt should be performed to estimate the shape of the test score distribution of population. Therefore, some smoothing methods have been developed to estimate the population distribution of test scores. Smoothing methods can be classified as presmoothing and postsmoothing methods. "Presmoothing is defined as the process of smoothing the observed score frequency distributions prior to equating. Postsmoothing is defined as the process of smoothing the equipercentile point after equating (Fairbank, 1987, 246). In this study the beta4 presmoothing, polynomial loglinear presmoothing and cubic splines postsmoothing methods were used. Sufficient explanations about these smoothing methods are given in Kolen & Brennan (2014).

Main purpose of this study is determining accuracy of beta4 and polynomial loglinear presmoothing and cubic spline postsmoothing methods in equipercentile equating for small samples.

The single group design with counterbalancing was used to apply test forms. In this design both Form X and Form Y applied to all of the students in different orders. About the half of the students has taken Form X first then they have taken form Y. So, the other half of the students have taken Form Y first then they have taken Form X. Therefore, both test forms have been applied to the 233 students. The Form X and Form Y are paralel tests prepared to final exams of the measurement and assessment class. Both tests contain of 20 multiple choce test items with same contents and statistical properties.

It is found that the effect of taking Form X after taking Form B does not differ from the effect of taking Form Y after taking Form A. So, there was no order effect and different application order of the test forms were not taken into account in analyzes. The computer program Rage-rgequate_wg(v3.15) was used to conduct the smoothing and

equating. Beta4 presmoothing method yield accurate fit with the sample distributions for both Form X and Form Y relative frequency distributions. According to graphical analyzes polynomial loglinear presmoothing method for C parameter 2 to 7 yield accurate fit with the sample distributions for both Form X and Form Y relative frequency distributions.

Moment preservations of the beta4 and loglineer presmoothed Form X and Form Y distributions were determined. In beta4 presmoothing for Form Y first three moments were very close or almost the same with sample moments. Kurtosis, the fourth moment was differentiated only in a very non-serious way. These results shows that equipercetile equating with beta4 presmoothing was quite suitable both for Form Y. In beta4 presmoothing for Form X all of the four moments were very close or almost the same with sample moments. Therefore, in beta4 presmoothing for Form X score distribution, moment preservition has been provided perfectly for all of the four sample moments.

As a result of chi square goodness of fit statistics, beta4 presmoothing method was found fit for both Form X and Form Y score distribution. Chi square goodness of fit statistics showed that polynomial loglinear presmoothing methods with $C \geq 2$ were found fit for both Form X and Form Y. All of these models show no significant differences considered as accurate models. In this case, polynomial model that has the lowest degree is selected. Therefore, in polynomial loglinear presmoothing the most accurate model has been found for $C=2$ for both Form X and Form Y. In other words, the quadratic polynomial model was found to be the most appropriate.

In equipercetile equating with cubic spline postsmoothing, determining what S level to use in accuracy of postsmoothing it is important to examine the graphics and moments. According to graphical analysis, cubic spline postsmoothings with $S=.30$ were found the most accurate model.

The all three smoothing methods, used in this study, were reduced bootstrap standard errors of equating and beta4 and polynomial loglinear presmoothing methods relatively produced less bootstrap standart errors of equipercetile equating. As a conclusion, It is found that all of the three smoothing methods, used in the study, have been found to reduce equating errors as compared with unsmoothed equepercetile equating. In other words, all of the three smoothing methods, used in the study, were effective. In term of relative efficiency, beta4 presmoothing method was more effective than both polynomial loglineer presmoothing and cubic postsmoothing methods for small sample data, used in the study.