Süleyman Demirel Üniversitesi,
Fen Bilimleri Enstitüsü Dergisi,
15-2 (2011),127-134

Çok Aşamalı Sıralı Küme Örneklemesi Tasarımlarının Etkinlikleri Üzerine Bir Çalışma

Nilay AKINCI[1], Yaprak Arzu ÖZDEMİR[2*]
[1]*TRT Genel Müdürlüğü Reklam Tasarım Tanıtım Dairesi Başkanlığı / ANKARA*
[2]*Gazi Üniversitesi Fen Fakültesi İstatistik Bölümü / ANKARA*
Alınış Tarihi:12.05.2011, Kabul Tarihi:26.07.2011

Özet: Sıralı küme örneklemesinde küme çapı arttıkça yığın ortalamasına ilişkin tahmin edicinin etkinliği artmaktadır. Ancak etkinliği arttırmak amacıyla, örnek seçim işleminde, küme çapı aynı kalmak üzere aşama sayısı arttırılarak, çeşitli çok aşamalı sıralı küme örneklemesi tasarımları önerilmiştir. Bu çalışmada, son yıllarda önerilen çok aşamalı sıralı küme örneklemesi tasarımlarından çift sıralı küme örneklemesi, çok aşamalı sıralı küme örneklemesi ve yüzde çift sıralı küme örneklemesi tasarımları incelenmiştir. Bu tasarımların sıralı küme örneklemesine göre etkinlikleri simülasyon çalışması ile çeşitli dağılımlar altında elde edilerek, yığın ortalamasını tahmin etmek üzere, en etkin tasarım belirlenmeye çalışılmıştır.

Anahtar Kelimeler: Sıralı Küme Örneklemesi, Çift Sıralı Küme Örneklemesi, Çok Aşamalı Sıralı Küme Örneklemesi, Yüzde Çift Sıralı Küme Örneklemesi, Göreli Etkinlik.

A Study on The Efficiencies of Multistage Ranked Set Sampling Designs

Abstract: In Ranked Set Sampling, the relative efficiency of the population mean estimator increases when the set size increases. However, different multistage ranked set sampling designs which are based on the sample selection with more than one stage with fixed set size are proposed to increase the efficiency. In this study, double ranked set sampling, multistage ranked set sampling and percentile double ranked set sampling designs are investigated. The relative efficiencies of these designs according to classical ranked set sampling are obtained by a simulation study under various distributions and most efficient design to estimate the population mean is determined.

Keywords: Ranked Set Sampling, Double Ranked Set Sampling, Multistage Ranked Set Sampling, Percentile Double Ranked Set Sampling, Relative Efficiency.

Giriş

İstatistiksel araştırmalarda, yığından seçilen örnek çapı ne kadar büyük ise elde edilen bilgiler de gerçeğe o kadar yakın olur. Ancak bütçe, eleman veya zaman yetersizliği gibi etkenlerden dolayı büyük çaplı örnekler seçilmesi mümkün olmayabilir. Bu durumda en az örnek çapı ile yığını en iyi şekilde temsil edecek bir örnekleme yöntemine ihtiyaç duyulur. Bu amaca yönelik olarak Mclntyre, 1952, yılında Sıralı Küme Örneklemesi (SKÖ) tasarımını önermiş ve SKÖ' nün yığın ortalamasını tahmin etmede Basit Tesadüfî Örnekleme(BTÖ)'ye göre daha etkin bir tasarım olduğunu göstermiştir. SKÖ' de örnek seçim işlemi aşağıdaki adımlar izlenerek yapılır.

1) Öncelikle ilgili yığından seçilen n^2 çaplı tesadüfî bir örnek, her biri n çaplı n kümeye tesadüfî olarak paylaştırılır. Böylece birbirinden bağımsız n çaplı n sayıda tesadüfî örnek elde edilmiş olur. Bu örneklerin her biri küme olarak adlandırılır.

2) Kümelerdeki birimler ilgilenilen değişken bakımından kendi içinde hassas ölçüm yapılmadan ucuz ve kolay bir ölçümle küçükten büyüğe doğru sıralanır. Bu sıralama görsel yolla veya daha ucuz ölçüm yapılması mümkün olan bir yardımcı değişkenden yararlanılarak yapılabilir.
3)Birimleri kendi içinde sıralanan kümelerin birincisinden ilk sıradaki birim, ikinci kümeden ikinci sıradaki birim ve bu şekilde devam edilerek n. kümeden n. sıradaki birim seçilir.
4)Seçilen birimler ilgilenilen değişken bakımından istenilen hassaslıktaki bir ölçümle ölçülür ve n çaplı sıralı küme örneği elde edilir. Bu işlem ilgilenilen örnek çapını elde etmek üzere r kez tekrarlanabilir. SKÖ nün etkinliği r tekrar sayısından etkilenmediği için bu çalışmada tekrar sayısı r=1 alınmıştır. Çizelge 1'de n=5 ve r=1 için SKÖ ile örneğe seçilen birimler gösterilmektedir. SKÖ ile elde edilen örnek $X_{1[1,5]1}$, $X_{2[2,5]1}$,…, $X_{5[5,5]1}$ olmak üzere $X_{i[i,n]j}$; j. tekrarda i. kümedeki i. sıra istatistiğini ifade etmektedir (i=1,2,…,n; j=1,2,…,r).

*yaprak@gazi.edu.tr

Çizelge 1. *n=5 ve r=1 için SKÖ ile örneğe seçilen birimler*

$X_{1[1,5]1}$	$X_{1[2,5]1}$	$X_{1[3,5]1}$	$X_{1[4,5]1}$	$X_{1[5,5]1}$
$X_{2[1,5]1}$	$X_{2[2,5]1}$	$X_{2[3,5]1}$	$X_{2[4,5]1}$	$X_{2[5,5]1}$
$X_{3[1,5]1}$	$X_{3[2,5]1}$	$X_{3[3,5]1}$	$X_{3[4,5]1}$	$X_{3[5,5]1}$
$X_{4[1,5]1}$	$X_{4[2,5]1}$	$X_{4[3,5]1}$	$X_{4[4,5]1}$	$X_{4[5,5]1}$
$X_{5[1,5]1}$	$X_{5[2,5]1}$	$X_{5[3,5]1}$	$X_{5[4,5]1}$	$X_{5[5,5]1}$

SKÖ tasarımının matematiksel teorisi Takahasi ve Wakimoto(1968) tarafından oluşturulmuştur. Patil vd.(1999) SKÖ' yü detaylı olarak inceleyerek, SKÖ ile ilgili temel çalışmaları özetlemişlerdir. SKÖ çeşitli dağılımlar altında, dağılım parametrelerinin etkin tahminlerini elde etmek amacıyla da kullanılmıştır. Bu konuda Log- normal, genelleştirilmiş geometrik, normal ve üstel dağılım parametrelerinin tahmini için sırasıyla Shen (1994), Bhoj ve Absanullah (1996) ve Sinha vd. (1996)'nın yapmış olduğu çalışmalar örnek gösterilebilir. Ayrıca dağılımın şekline bağlı olarak yığın parametrelerinin daha etkin tahmin edicilerinin elde edilmesi amacıyla, SKÖ de örnek seçim işleminde çeşitli değişiklikler yapılarak farklı SKÖ tasarımları da önerilmiştir. Medyan SKÖ(MSKÖ), Uç SKÖ(USKÖ), tesadüfî seçime dayalı SKÖ, Yüzde SKÖ(YSKÖ) ve L tahmin edici fikrine dayalı L-Sıralı Küme Örneklemesi(LSKÖ) önerilen önemli SKÖ tasarımlarıdır (Muttlak, 1997; Samawi vd., 1996; Li vd., 1999; Muttlak, 2003; Al-Nasser, 2007). SKÖ ile ilgili bazı Türkçe çalışmalarda bulunmaktadır (Çıngı, 2009; Özdemir, 2005).

Son yıllarda SKÖ de örnek çapı n sabit kalmak üzere, sıralama için örneğe seçilen birim sayısı ve örnek seçim aşamaları arttırılarak yeni SKÖ tasarımları da önerilmiştir. Bu konuda ilk çalışma, Al-Saleh ve Al-Kadiri (2000) tarafından yapılmıştır. Al-Saleh ve Al-Kadiri (2000) SKÖ deki örnek seçim işlemini n^3 birimle başlatacak şekilde aşama sayısı 2 olmak üzere, Çift Sıralı Küme Örneklemesini(ÇSKÖ) önermişlerdir. ÇSKÖ de Tekdüze (0,1), Üstel (1) ve Normal (0,1) dağılım altında ÇSKÖ' nün BTÖ' ye göre etkinliğini incelemişlerdir. ÇSKÖ ile elde edilen yığın ortalamasına ilişkin tahmin edicinin Normal ve Tekdüze dağılımları altında, Üstel dağılıma göre daha etkin olduğu gösterilmiştir. Al-Saleh ve Al-Omari (2002) ÇSKÖ' yü geliştirerek Çok Aşamalı Sıralı Küme Örneklemesi (ÇASKÖ)' ni önermişlerdir. ÇASKÖ tasarımı, ÇSKÖ tasarımının genel bir k aşama sayısına göre genelleştirilmiş şeklidir. ÇASKÖ tasarımını örnek çapının sadece n=2 olduğu durum için ele alarak, farklı k aşama sayıları için çeşitli dağılımlar altında BTÖ' ye göre etkinliklerini incelemişlerdir. Özellikle k aşama sayısı arttıkça etkinliğin arttığı gözlenmiştir. Al-Omari ve Jaber (2008) ise araştırmacı tarafından belirlenecek belli bir p oranına bağlı olarak örnek seçiminin yapılacağı Yüzde Çift SKÖ (YÇSKÖ) tasarımını önermişlerdir. YÇSKÖ tasarımına göre elde edilen ortalama tahmin edicisinin simetrik dağılımlar altında yığın ortalaması için yansız bir tahmin edici olduğunu ve SKÖ, MSKÖ ve USKÖ tasarımlarına göre elde edilen tahmin edicilerden daha etkin bir tahmin edici olduğunu göstermişlerdir.

Bu çalışmada, ÇSKÖ, ÇASKÖ ve YÇSKÖ tasarımları karşılaştırmalı olarak incelenecektir. Al-Saleh ve Al-Kadiri (2000), Al-Saleh ve Al-Omari (2002) ve Al-Omari ve Jaber (2008) yaptıkları çalışmalarda, önerdikleri SKÖ tasarımlarının BTÖ ye göre etkinliklerini incelemişlerdir. Ancak ÇASKÖ tasarımları, SKÖ ye alternatif olarak önerildikleri için bu çalışmada önerilen tasarımların SKÖ ye göre etkinlikleri incelenecektir. Ayrıca çeşitli dağılımlar ve farklı örnek çapları altında etkinlik değerleri simülasyon çalışması ile elde edilerek, hangi dağılım altında hangi tasarımın daha etkin olduğu tespit edilmeye çalışılacaktır.

Çalışmanın ikinci ve üçüncü bölümünde çok aşamalı SKÖ tasarımlarından ÇSKÖ, ÇASKÖ ve YÇSKÖ tasarımları ele alınarak, örnek seçim işlemlerine açıklık getirilecektir. Dördüncü bölümde incelenen tasarımların çeşitli dağılımlar altında SKÖ ye göre göreli etkinlik(GE) değerleri simülasyon çalışması ile elde edilmiştir. Beşinci bölümde ise sonuç ve önerilere yer verilmiştir.

Çok Aşamalı ve Çift Sıralı Küme Örneklemesi

ÇSKÖ de örnek seçim işlemi aşama sayısı k=2 olmak üzere, $n^{k+1}=n^3$ birimle başlamaktadır. ÇASKÖ ise ÇSKÖ nün genel bir k aşama sayısına genelleştirilmiş şeklidir. Bu nedenle ÇSKÖ, ÇASKÖ de k=2 olduğu durum olarak ele alınacaktır. ÇASKÖ'de örnek seçimi aşağıdaki adımlar izlenerek yapılır.

1) İlgili yığından n^{k+1} birim tesadüfî olarak seçilir.
2) Seçilen birimler her biri n^2 birimden oluşan n^{k-1} gruba tesadüfi olarak ayrılıp, her bir gruptaki birimler kendi içinde n çaplı n kümeye tesadüfi olarak paylaştırılır.
3) Her grup için ayrı ayrı bilinen SKÖ tasarımının 2. ve 3. adımları uygulanarak n çaplı n^{k-1} sayıda örnek elde edilir. Elde edilen örnekler birer küme olmak üzere, n kümeden tekrar yeni bir grup oluşturularak, hassas ölçümleri yapılmamış n^2 çaplı n^{k-2} grup elde edilir.
4) Elde edilen gruplara, n çaplı n küme elde edilinceye kadar 3. adım tekrar uygulanır.
5) Elde edilen n çaplı n kümeye bilinen SKÖ tasarımının 2., 3. ve 4. adımları uygulanarak hassas ölçümü yapılmış n çaplı çok aşamalı sıralı küme örneği elde edilir .

Yukarıda verilen örnek seçim aşamaları k=2 iken ÇSKÖ ne karşılık gelmektedir.

ÇASKÖ ile örnek seçim işlemini daha iyi açıklamak üzere, n=2 ve k=3 olduğu durum ele alınırsa, özellikle, n^{k+1}=16 birim yığından tesadüfî olarak seçilir. Birinci aşamada, seçilen birimler n^2=4 birimden oluşan n^{k-1}=4 gruba ayrılarak, her bir grup içindeki birimler 2 çaplı 2 kümeye tesadüfi olarak paylaştırılır. Bu 4 gruba SKÖ de örnek seçim işleminin 2. ve 3. adımları uygulanır. İkinci aşamada ise, 1. ve 2. gruptan seçilen örnek birimleri ile 3.ve 4. gruptan seçilen örnek birimleri 2 çaplı 2 küme oluşturacak şekilde 2 grup elde edilir. Birinci aşamada 1. gruptan seçilen birimler ikinci aşamada 1. grubun 1. kümesini oluşturmakta, 2. gruptan seçilen birimler ikinci aşamada 1. grubun, 2. kümesini oluşturmakta, bu şekilde devam edilerek, 4. gruptan seçilen birimler ikinci aşamada 2. grubun, 2. kümesini oluşturmaktadır. İkinci aşamada da benzer şekilde her iki grup için de SKÖ'de örnek seçim işleminin 2. ve 3. adımları uygulanır. 1. gruptan seçilen birimler üçüncü aşamada 1. grubun 1. kümesini oluşturmakta, 2. gruptan seçilen birimler ise 1. grubun 2. kümesini oluşturmaktadır.
3. aşamada elde edilen kümelere tekrar SKÖ de örnek seçim işleminin 3. ve 4. adımları uygulanarak, 2 birimlik çok aşamalı sıralı küme örneği elde edilir. Çizelge 2' de n=2 ve k=3 iken örneğe seçilen birimler gösterilmektedir.

Çizelge 2. *n=2 ve k=3 için ÇASKÖ tasarımı ile örneğe seçilen birimler*

Birinci aşama

$\boxed{X_{1\,1,2}^{(1,1)}}$	$X_{1\,2,2}^{(1,1)}$	$\boxed{X_{1\,1,2}^{(2,1)}}$	$X_{1\,2,2}^{(2,1)}$	$\boxed{X_{1\,1,2}^{(3,1)}}$	$X_{1\,2,2}^{(3,1)}$	$\boxed{X_{1\,1,2}^{(4,1)}}$	$X_{1\,2,2}^{(4,1)}$
$X_{2\,1,2}^{(1,1)}$	$\boxed{X_{2\,2,2}^{(1,1)}}$	$X_{2\,1,2}^{(2,1)}$	$\boxed{X_{2\,2,2}^{(2,1)}}$	$X_{2\,1,2}^{(3,1)}$	$\boxed{X_{2\,2,2}^{(3,1)}}$	$X_{2\,1,2}^{(4,1)}$	$\boxed{X_{2\,2,2}^{(4,1)}}$

İkinci aşama

$\boxed{X_{1\,1,2}^{(1,2)}}$	$X_{2\,2,2}^{(1,2)}$	$\boxed{X_{1\,1,2}^{(3,2)}}$	$X_{2\,2,2}^{(3,2)}$
$X_{1\,1,2}^{(2,2)}$	$\boxed{X_{2\,2,2}^{(2,2)}}$	$X_{1\,1,2}^{(4,2)}$	$\boxed{X_{2\,2,2}^{(4,2)}}$

Üçüncü aşama

$\boxed{X_{1\,1,2}^{(1,3)}}$	$X_{2\,2,2}^{(2,3)}$
$X_{1\,1,2}^{(3,3)}$	$\boxed{X_{2\,2,2}^{(4,3)}}$

Birinci aşamada, birinci gruptan seçilen birimler $S_1 = X_{1[1,2]}^{(1,1)}, X_{2[2,2]}^{(1,1)}$, ikinci gruptan seçilen birimler $S_2 = X_{1[1,2]}^{(2,1)}, X_{2[2,2]}^{(2,1)}$, üçüncü gruptan seçilen birimler $S_3 = X_{1[1,2]}^{(3,1)}, X_{2[2,2]}^{(3,1)}$ ve dördüncü gruptan seçilen birimler $S_4 = X_{1[1,2]}^{(4,1)}, X_{2[2,2]}^{(4,1)}$ olarak tanımlansın. İkinci aşamada hassas ölçüm yapılmadan seçilen birimler $X_{1[1,2]}^{(2)} = min(S_1)$, $X_{2[2,2]}^{(2)} = maks(S_2)$, $X_{1[1,2]}^{(2)} = min(S_3)$ ve $X_{2[2,2]}^{(2)} = maks(S_4)$ olmak üzere, üçüncü aşamada hassas ölçümleri yapılarak seçilen birimler $X_{1[1,2]}^{(3)} = \min\ X_{1[1,2]}^{(2)}, X_{1[2,2]}^{(2)}$ ve $X_{2[2,2]}^{(3)} = maks\ X_{2[1,2]}^{(2)}, X_{2[2,2]}^{(2)}$ şeklinde tanımlanabilir. Burada $X_{i[i,n]}^{(k)}$; k. aşama sonunda elde edilen i. kümedeki i. sıra istatistiğini ifade etmektedir.

ÇASKÖ' de yığın ortalamasının yansız tahmin edicisi ve bu tahmin edicinin varyansı sırasıyla,

$$\bar{X}_{ÇASKÖ} = \left(\frac{1}{n}\right)\sum_{i=1}^{n} X_{i[i,n]}^{(k)} \qquad (1)$$

$$Var(\bar{X}_{ÇASKÖ}) = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^{2(k)} = \frac{1}{n}\left[\sigma^2 - \frac{1}{n}\sum_{i=1}^{n}\ \mu_i^{(k)} - \mu\ ^2\right] \qquad (2)$$

olarak elde edilir (Al-Saleh ve Al-Omari (2002)). Burada $\mu_i^{(k)}$, k. aşamada elde edilen i. sıra istatistiğinin beklenen değeri ($E(X_{i[i,n]}^{(k)}) = \mu_i^{(k)}$) ve $\sigma_i^{2(k)}$ ise k. aşamada elde edilen i. sıra istatistiğinin varyansıdır ($Var(X_{i[i,n]}^{(k)}) = \sigma_i^{2(k)}$). ÇASKÖ tasarımında birinci aşamada seçilen birimler ikinci aşamada tekrar sıralanıp bir grup oluşturulmakta ve bu olay k. aşamaya kadar devam etmektedir. Bu nedenle, i. sıra istatistiğinin olasılık yoğunluk fonksiyonu her aşamada değişmektedir. Burada, $X_{i[i,n]}^{(k)}$ ' lar bağımsız fakat aynı dağılımlı değillerdir. Dağılımları ise (k-1). aşamada elde edilen i. sıra istatistiği $X_{i[i:n]}^{(k-1)}$ ' in dağılımına bağlıdır. k. aşamada elde edilen i.

sıra istatistiğinin olasılık yoğunluk fonksiyonu $f_{i:n}^{(k)}(x)$ olmak üzere, X’ in olasılık yoğunluk fonksiyonu;

$$f(x)=\frac{1}{n}\sum_{i=1}^{n} f_{i:n}^{(k)}(x) \tag{3}$$

olur. Böylece, X’in beklenen değeri,

$$\mu=\frac{1}{n}\sum_{i=1}^{n}\mu_i^{(k)} \tag{4}$$

dır (Al-Saleh ve Al-Omari, 2002). Buradan Eş. (1) de tanımlanan $\overline{X}_{ÇASKÖ}$ tahmin edicisinin yığın ortalaması için yansız olduğu

$$E(\overline{X}_{ÇASKÖ})=\frac{1}{n}E\left[\sum_{i=1}^{n}X_{i[i,n]}^{(k)}\right]=\frac{1}{n}\sum_{i=1}^{n}\mu_i^{(k)}=\mu \tag{5}$$

şeklinde gösterilebilir.

Yüzde Çift Sıralı Küme Örneklemesi

Yüzde çift sıralı küme örneklemesi (YÇSKÖ), yığın ortalamasını tahmin etmek için 2008 yılında Al-Omari ve Jaber tarafından önerilmiştir. YÇSKÖ, birimlerin önceden araştırmacı tarafından belirlenen bir p değerine (0<p<1)göre seçimine dayanmaktadır.

YÇSKÖ’ de örnek seçim işlemi aşağıdaki adımlara göre yapılır;
1) İlgili yığından n^3 birim tesadüfi olarak seçilir.
2) Birimler her biri n^2 büyüklüğünde n gruba tesadüfi olarak ayrılıp, her bir gruptaki birimler kendi içinde n çaplı n kümeye tesadüfi olarak paylaştırılır.
3) Her bir gruptaki kümelere SKÖ de örnek seçim işleminin 2. ve 3. adımları uygulanır.
4) 3. adımda elde edilen örnekler birer küme olmak üzere, elde edilen yeni gruptaki n çaplı n kümede n çift ise; ilk n/2 sayıda kümeden p(n+1). sıradaki birimler ve kalan n/2 sayıda kümeden q(n+1). sıradaki birimler seçilir. n tek ise; ilk (n–1)/2 sayıda kümeden p(n+1). sıradaki birimler, ((n+1)/2). kümeden medyan değeri ve son (n–1)/2 sayıda kümeden q(n+1). sıradaki birimler örneğe seçilerek, istenilen hassaslıktaki bir ölçümle X değişkeni bakımından ölçülür (p(n+1) ve q(n+1) için en yakın tamsayı değerleri alınır). Böylece n çaplı yüzde çift sıralı küme örneği elde edilir. Burada q=1-p’dir.

YÇSKÖ’de oluşturulan gruplardan örnek seçim işlemi iki aşamada gerçekleştirildiğinden k=2’dir. Örnek seçim işlemini daha iyi açıklamak üzere n=4 ve p=0.20 alınırsa, öncelikle $n^3=4^3=64$ birim yığından tesadüfi olarak seçilir. Birinci aşamada birimler, $4^2=16$ birim içeren 4 gruba ayrılır. Her bir grup içindeki birimler 4 çaplı 4 kümeye tesadüfi olarak paylaştırılır. Bu kümelere SKÖ de örnek seçim işleminin 2.ve 3. adımları uygulanır. Seçilen örnek birimleri kullanılarak, 1. gruptan seçilen birimler 1. kümeyi, 2. gruptan seçilen birimler 2. kümeyi ve bu şekilde devam edilerek, 4. gruptan seçilen birimler 4. kümeyi oluşturacak şekilde, 4 çaplı 4 küme oluşturulur. İkinci aşamada, bu kümelerin ilk ikisinden 0.20(5)=1. sıradaki birimler, son ikisinden ise 0.80(5)=4. sıradaki birimler örneğe seçilerek ilgilenilen değişken bakımından ölçülür. Böylece, yüzde çift sıralı küme örneği elde edilir. Çizelge 3’ de n=4 p=0.20 iken YÇSKÖ tasarımı uygulanarak örneğe seçilen birimler gösterilmektedir.

Çizelge 3. *n=4, p=0.20 için YÇSKÖ tasarımı ile örneğe seçilen birimler*

Birinci aşama

$\boxed{X_{1[1,4]}^{(1,1)}}$	$X_{1[2,4]}^{(1,1)}$	$X_{1[3,4]}^{(1,1)}$	$X_{1[4,4]}^{(1,1)}$
$X_{2[1,4]}^{(1,1)}$	$\boxed{X_{2[2,4]}^{(1,1)}}$	$X_{2[3,4]}^{(1,1)}$	$X_{2[4,4]}^{(1,1)}$
$X_{3[1,4]}^{(1,1)}$	$X_{3[2,4]}^{(1,1)}$	$\boxed{X_{3[3,4]}^{(1,1)}}$	$X_{3[4,4]}^{(1,1)}$
$X_{4[1,4]}^{(1,1)}$	$X_{4[2,4]}^{(1,1)}$	$X_{4[3,4]}^{(1,1)}$	$\boxed{X_{4[4,4]}^{(1,1)}}$

$\boxed{X_{1[1,4]}^{(2,1)}}$	$X_{1[2,4]}^{(2,1)}$	$X_{1[3,4]}^{(2,1)}$	$X_{1[4,4]}^{(2,1)}$
$X_{2[1,4]}^{(2,1)}$	$\boxed{X_{2[2,4]}^{(2,1)}}$	$X_{2[3,4]}^{(2,1)}$	$X_{2[4,4]}^{(2,1)}$
$X_{3[1,4]}^{(2,1)}$	$X_{3[2,4]}^{(2,1)}$	$\boxed{X_{3[3,4]}^{(2,1)}}$	$X_{3[4,4]}^{(2,1)}$
$X_{4[1,4]}^{(2,1)}$	$X_{4[2,4]}^{(2,1)}$	$X_{4[3,4]}^{(2,1)}$	$\boxed{X_{4[4,4]}^{(2,1)}}$

$\boxed{X_{1[1,4]}^{(3,1)}}$	$X_{1[2,4]}^{(3,1)}$	$X_{1[3,4]}^{(3,1)}$	$X_{1[4,4]}^{(3,1)}$
$X_{2[1,4]}^{(3,1)}$	$\boxed{X_{2[2,4]}^{(3,1)}}$	$X_{2[3,4]}^{(3,1)}$	$X_{2[4,4]}^{(3,1)}$
$X_{3[1,4]}^{(3,1)}$	$X_{3[2,4]}^{(3,1)}$	$\boxed{X_{3[3,4]}^{(3,1)}}$	$X_{3[4,4]}^{(3,1)}$
$X_{4[1,4]}^{(3,1)}$	$X_{4[2,4]}^{(3,1)}$	$X_{4[3,4]}^{(3,1)}$	$\boxed{X_{4[4,4]}^{(3,1)}}$

$\boxed{X_{1[1,4]}^{(4,1)}}$	$X_{1[2,4]}^{(4,1)}$	$X_{1[3,4]}^{(4,1)}$	$X_{1[4,4]}^{(4,1)}$
$X_{2[1,4]}^{(4,1)}$	$\boxed{X_{2[2,4]}^{(4,1)}}$	$X_{2[3,4]}^{(4,1)}$	$X_{2[4,4]}^{(4,1)}$
$X_{3[1,4]}^{(4,1)}$	$X_{3[2,4]}^{(4,1)}$	$\boxed{X_{3[3,4]}^{(4,1)}}$	$X_{3[4,4]}^{(4,1)}$
$X_{4[1,4]}^{(4,1)}$	$X_{4[2,4]}^{(4,1)}$	$X_{4[3,4]}^{(4,1)}$	$\boxed{X_{4[4,4]}^{(4,1)}}$

İkinci aşama

$\boxed{X^{(1,2)}_{1[1,4]}}$	$X^{(1,2)}_{2[2,4]}$	$X^{(1,2)}_{3[3,4]}$	$X^{(1,2)}_{4[4,4]}$
$\boxed{X^{(2,2)}_{1[1,4]}}$	$X^{(2,2)}_{2[2,4]}$	$X^{(2,2)}_{3[3,4]}$	$X^{(2,2)}_{4[4,4]}$
$X^{(3,2)}_{1[1,4]}$	$X^{(3,2)}_{2[2,4]}$	$X^{(3,2)}_{3[3,4]}$	$\boxed{X^{(3,2)}_{4[4,4]}}$
$X^{(4,2)}_{1[1,4]}$	$X^{(4,2)}_{2[2,4]}$	$X^{(4,2)}_{3[3,4]}$	$\boxed{X^{(4,2)}_{4[4,4]}}$

Birinci aşamada birinci gruptan seçilen birimler $S_1 = X^{(1,1)}_{1[1,4]}, X^{(1,1)}_{2[2,4]}, X^{(1,1)}_{3[3,4]}, X^{(1,1)}_{4[4,4]}$, ikinci gruptan seçilen birimler $S_2 = X^{(2,1)}_{1[1,4]}, X^{(2,1)}_{2[2,4]}, X^{(2,1)}_{3[3,4]}, X^{(2,1)}_{4[4,4]}$, üçüncü gruptan seçilen birimler $S_3 = X^{(3,1)}_{1[1,4]}, X^{(3,1)}_{2[2,4]}, X^{(3,1)}_{3[3,4]}, X^{(3,1)}_{4[4,4]}$ ve dördüncü gruptan seçilen birimler $S_4 = X^{(4,1)}_{1[1,4]}, X^{(4,1)}_{2[2,4]}, X^{(4,1)}_{3[3,4]}, X^{(4,1)}_{4[4,4]}$ olarak tanımlanırsa, ikinci aşamada sıralandıktan sonra ölçülerek seçilen birimler $X^{(2)}_{1[1,4]} = min(S_1)$, $X^{(2)}_{2[1,4]} = \min(S_2)$ ve $X^{(2)}_{3[4,4]} = maks(S_3)$, $X^{(2)}_{4[4,4]} = maks(S_4)$ olarak tanımlanabilir. $X^{(k)}_{i[i,n]}$; k. aşama sonunda elde edilen i. kümedeki i. sıra istatistiğini ifade etmektedir. Yığın ortalamasına ilişkin tahmin edici, n'in tek ve çift olmasına göre iki farklı biçimde tanımlanır. n çift ise, yığın ortalamasının YÇSKÖ tahmin edicisi ve bu tahmin edicinin varyansı sırasıyla,

$$\bar{X}_{YÇSKÖ} = \frac{1}{n}\left(\sum_{i=1}^{\frac{n}{2}} X^{(2)}_{i[p(n+1),n]} + \sum_{i=\frac{n+2}{2}}^{n} X^{(2)}_{i[q(n+1),n]}\right) \quad (6)$$

$$Var(\bar{X}_{YÇSKÖ}) = \frac{1}{n^2}\left(\sum_{i=1}^{\frac{n}{2}} Var(X^{(2)}_{i[p(n+1),n]}) + \sum_{i=\frac{n+2}{2}}^{n} Var(X^{(2)}_{i[q(n+1),n)})\right) \quad (7)$$

biçiminde tanımlanır. n tek olduğunda ise, yığın ortalamasının YÇSKÖ tahmin edicisi ve bu tahmin edicinin varyansı,

$$\bar{X}_{YÇSKÖ} = \frac{1}{n}\left(\sum_{i=1}^{\frac{n-1}{2}} X^{(2)}_{i[p(n+1),n]} + X^{(2)}_{\frac{n+1}{2}\left(\frac{n+1}{2},n\right)} + \sum_{i=\frac{n+3}{2}}^{n} X^{(2)}_{i[q(n+1),n]}\right) \quad (8)$$

$$Var(\bar{X}_{YÇSKÖ}) = \frac{1}{n^2}\left(\sum_{i=1}^{\frac{n-1}{2}} Var(X^{(2)}_{i(p(n+1),n)}) + Var\left(X^{(2)}_{\frac{n+1}{2}(\frac{n+1}{2},n)}\right) + \sum_{i=\frac{n+3}{2}}^{n} Var(X^{(2)}_{i(q(n+1):n)})\right) \quad (9)$$

olacaktır. Simetrik dağılımlar altında $\bar{X}_{YÇSKÖ}$ yığın ortalaması için yansız bir tahmin edicidir (Al-Omari ve Jaber, 2008). p değerine bağlı olarak YÇSKÖ farklı sıra istatistiklerinin örneğe seçilmesine ve dolayısıyla her örnek çapı ve p değeri için farklı tasarımların oluşmasına neden olmaktadır.

Çeşitli Dağılımlar Altında En Uygun SKÖ Tasarımının Belirlenmesi İçin Monte Carlo Simülasyon Çalışması

Bu bölümde, ÇASKÖ ve YÇSKÖ tasarımlarının SKÖ'ye göre etkinlikleri simülasyon yoluyla incelenecektir. Matlab paket programı kullanılarak, her bir tasarım için n=3, 4 ve 5 çaplı örnekler seçilerek, yığın ortalamasına ilişkin tahmin edicinin beklenen değeri ve Hata Kare Ortalama (HKO)'sı 100000 tekrarlı simülasyon çalışması ile hesaplanmıştır. ÇASKÖ için aşama sayısı k=2 ve 3, YÇSKÖ için p'nin 0.20 ve 0.40 değerleri alınmıştır. SKÖ' ye göre GE,

$$GE = \frac{Var(\bar{X}_{SKÖ})}{HKO(\bar{X}^{*}_{SKÖ})} \quad (10)$$

şeklinde tanımlanmaktadır. Burada $HKO(\bar{X}^{*}_{SKÖ})$, diğer SKÖ tasarımlarından (ÇASKÖ, YÇSKÖ) elde edilen tahmin edicinin HKO değerini ifade etmektedir. ÇASKÖ ve simetrik dağılımlar için YÇSKÖ tasarımlarında, yığın ortalamasına ilişkin tahmin edici yansız olduğu için $HKO(\bar{X}^*_{SKÖ})$ ifadesi $Var(\bar{X}^*_{SKÖ})$'ye dönüşecektir.

Yığın dağılımı olarak simetrik dağılımlardan, Normal (0,1), Tekdüze (0,1), Laplace (0,0.5), U şekilli dağılım U(0,1), simetrik olmayan dağılımlardan ise, Beta (2,9), Beta (9,2), Üstel (1) ve Weibull (1,0.5) dağılımları altında inceleme yapılmıştır. SKÖ' nün BTÖ'ye göre etkinliği özellikle dağılımın şekline göre değişmektedir. Çok aşamalı tasarımlar altında da benzer durumun ortaya çıkabileceği düşüncesi ile, dağılımlar belirlenirken basıklık ve çarpıklık katsayıları dikkate alınmıştır. İncelenen dağılımların basıklık ve çarpıklık katsayıları Çizelge 4' de verilmektedir. Elde edilen GE değerleri ise Çizelge 5-7 arasında verilmektedir.

***Çizelge 4**. Dağılımların basıklık ve çarpıklık katsayıları*

	Basıklık katsayısı	Çarpıklık katsayısı
Normal (0,1)	0	0
Tekdüze (0,1)	-1,2	0
Laplace (0,0.5)	3	0
U (0,1)	-1,5	0
Üstel (1)	6	2
Beta (2,9)	0,6483	0,8793
Beta (9,2)	0,6483	-0,8793
Weibull (1,0.5)	84,7199	6,6188

***Çizelge 5.** n=3 iken SKÖ ne göre GE değerleri*

Dağılımlar	YÇSKÖ(p=0,20)	YÇSKÖ(p=0,40)	ÇASKÖ(k=2)	ÇASKÖ(k=3)
Normal (0,1)	1.37	1.89	1.37	1.64
Tekdüze (0,1)	1.53	1.21	1.51	1.97
Laplace (0,0.5)	1.24	3.45	1.24	1.37
U (0,1)	1.54	0.96	1.53	2.08
Üstel (1)	1.22	1.72	1.23	1.37
Beta (2,9)	1.37	1.58	1.37	1.61
Beta (9,2)	1.36	1.59	1.36	1.62
Weibull (1,0.5)	1.06	2.93	1.03	1.04

***Çizelge 6.** n=4 iken SKÖ ne göre GE değerleri*

Dağılımlar	YÇSKÖ(p=0,20)	YÇSKÖ(p=0,40)	ÇASKÖ(k=2)	ÇASKÖ(k=3)
Normal (0,1)	1.15	2.15	1.49	1.85
Tekdüze (0,1)	2.22	1.37	1.69	2.33
Laplace (0,0.5)	0.79	3.81	1.31	1.48
U (0,1)	3.72	1.15	1.77	2.56
Üstel (1)	0.64	1.57	1.35	1.52
Beta (2,9)	0.99	1.70	1.47	1.84
Beta (9,2)	0.98	1.70	1.48	1.82
Weibull (1,0.5)	0.47	2.39	1.08	1.12

***Çizelge 7.** n=5 iken SKÖ ne göre GE değerleri*

Dağılımlar	YÇSKÖ(p=0,20)	YÇSKÖ(p=0,40)	ÇASKÖ(k=2)	ÇASKÖ(k=3)
Normal (0,1)	1.26	2.33	1.63	2.06
Tekdüze (0,1)	2.33	1.59	1.89	2.79
Laplace (0,0.5)	0.83	3.81	1.38	1.58
U (0,1)	3.12	1.42	1.96	2.98
Üstel (1)	0.66	1.49	1.39	1.61
Beta (2,9)	1.03	1.74	1.60	2.01
Beta (9,2)	1.03	1.74	1.60	2.01
Weibull (1,0.5)	0.47	2.16	1.11	1.16

Çizelge 5,6 ve 7 den görüldüğü gibi, simetrik tek modlu dağılımlardan Normal ve Laplace dağılımı altında, seçilen örnek çapları için, en yüksek GE değeri YÇSKÖ (p=0.40) de elde edilmektedir. Örnek çapı 2'den 5'e doğru arttıkça GE değeri de artmaktadır. YÇSKÖ'de p=0.40 iken elde edilen örnek, medyan gibi merkeze yakın sıra istatistiklerinden oluşmaktadır. Bu nedenle, yığın ortalamasına yakın tahminler elde edildiği için GE yüksek değerler almaktadır.

Simetrik tek modlu olmayan Tekdüze ve U dağılımları altında ise, en yüksek GE değerlerinin YÇSKÖ(p=0.20) ve ÇASKÖ (k=3) de elde edildiği görülmektedir. Ancak, YÇSKÖ de örnek seçim işlemi 2 aşamada gerçekleştirilmektedir. Bu nedenle YÇSKÖ ile bulunan GE değerlerini ÇASKÖ de k=2 ile karşılaştırmak daha uygun olacaktır. ÇASKÖ (k=3) durumu ise aşama sayısı arttıkça GE değerlerinin arttığını ifade etmektedir. Bu anlamda Tekdüze ve U dağılımları için en uygun tasarımın YÇSKÖ (p=0.20) olduğu söylenebilir. Diğer taraftan YÇSKÖ (p=0.20)'den sonra ÇASKÖ nin etkin bir tasarım olduğu ve aşama sayısı ile örnek çapı arttıkça GE değerlerinin çok az da olsa arttığı söylenebilir. Tek modlu olmayan simetrik dağılımlarda, yığın ortalaması, dağılımın uç değerlerinden daha fazla etkilenmektedir. YÇSKÖ (p=0.20) ile örneğe çoğunlukla uç sıra istatistikleri seçildiği için, bu tür tek modlu olmayan simetrik dağılımlarda GE değeri yüksek olacaktır.

Simetrik olmayan dağılımlardan Üstel, Beta ve Weibull dağılımlarında ise, en yüksek GE değeri YÇSKÖ (p=0.40) de elde edilmektedir. Üstel ve Weibull (1,0.5) dağılımları altında örnek çapı arttıkça daha çok merkeze yakın birimler örneğe seçildiği için GE değeri azalmakta, ancak Beta (2,9) ve Beta (9,2) dağılımlarında örnek çapı arttıkça GE değeri artmaktadır.

Ayrıca simetrik dağılımlar altında en yüksek GE değeri Laplace dağılımına aittir. Dağılımın basıklık katsayısına bakıldığında, diğer simetrik dağılımlardan daha büyük olduğu yani daha sivri bir dağılım olduğu görülmektedir. Dolayısıyla bu dağılım altında, YÇSKÖ (p=0.40) ile merkezden alınan sıra istatistikleri HKO nin daha küçük çıkmasını ve GE değerinin artmasını sağlamaktadır.

Simetrik olmayan dağılımlar içinde Beta (9,2) ve Beta (2,9) dağılımları sırasıyla sağa ve sola çarpık dağılımlar olduğundan basıklık katsayıları aynı, çarpıklık katsayıları da mutlak değerce birbirinin aynısıdır. GE değerlerine bakıldığında ise birbirine çok yakın değerler verdiği görülmektedir. Buradan da dağılımın çarpıklığının GE değeri üzerinde etkili olmadığı söylenebilir.

Çalışmada incelenen simetrik olmayan dağılımlardan basıklığı en yüksek olan Weibull (1,0.5) dağılımı altında, YÇSKÖ (p=0.40) iken elde edilen GE değerleri diğer simetrik olmayan dağılımlara göre daha yüksektir.

Ayrıca bazı durumlarda GE'in 1 den küçük değerler aldığı görülmektedir. Örneğin n=3 iken, U şekilli dağılım U (0,1) için YÇSKÖ (p=0.40)'de GE=0.96 dır. YÇSKÖ (p=0.40) ile dağılımın daha çok merkezinden örnekler alınmaktadır. Ancak dağılımın U şeklinde olmasından dolayı veriler uçlarda yoğunlaşmaktadır. Bilinen SKÖ tasarımında uçlarda yer alan birimler de örneğe dahil edildiği için SKÖ ile elde edilen HKO değeri, YÇSKÖ (p=0.40) dan elde edilen HKO değerinden daha küçük olacak ve GE değeri azalacaktır. Bu gibi durumlarda bilinen SKÖ tasarımının önerilen diğer SKÖ tasarımlarından daha etkin bir tasarım olduğu söylenebilir.

Sonuç

Bu çalışmada, SKÖ tasarımına alternatif olarak önerilen YÇSKÖ ve ÇASKÖ tasarımları incelenmiştir. YÇSKÖ ve ÇASKÖ de örnek seçim işlemlerine açıklık getirilerek, çeşitli dağılımlar altında SKÖ'ne göre etkinlikler GE ölçüsüne dayalı olarak Monte Carlo simülasyon çalışması ile elde edilmiştir. Elde edilen GE değerlerinden yararlanılarak, bu çalışmada incelenen tek modlu simetrik dağılımlar için yığın ortalamasını tahmin etmek üzere, en uygun tasarımın n=3,4 ve 5 iken YÇSKÖ (p=0.40) ve tek modlu olmayan simetrik dağılımlar için ise n=3, 4 ve 5 iken YÇSKÖ (p=0.20) olduğu söylenebilir. İncelenen simetrik olmayan dağılımlarda ise, yığın ortalamasını tahmin etmek üzere en uygun tasarım n=3, 4 ve 5 iken YÇSKÖ (p=0.40) dır. Ancak Üstel ve Weibull (1,0.5) dağılımları için YÇSKÖ (p=0.40) de örnek çapı arttıkça GE değeri n=3,4 ve 5'de azalmaktadır. ÇASKÖ de ise örnek çapı ve aşama sayısı arttıkça GE değeri incelenen her dağılım için artmaktadır. Bu nedenle, ÇASKÖ de aşama sayısı 2 den daha büyük değerler alınarak daha yüksek GE değerleri elde edilebilir. Ancak burada dikkat edilecek nokta, aşama sayısı arttıkça örneğe sıralama için seçilecek birim sayısının da artmasıdır. Bu durumda, sıralamada hata yapma olasılığını düşürmek amacıyla, örnek çapının küçük olması tercih edilir. Örnek çapı arttıkça, aşama sayısına bağlı olarak çok fazla sayıda birim örneğe alınacak, bu ise hem maliyet hem de sıralama açısından uygulamada zorluklar ortaya çıkaracaktır.

Kaynaklar

Al-Nasser, A.D., 2007. L Ranked Set Sampling: A Generalization Procedure for Robust Visual Sampling. Communications in Statistics-Simulation and Computation**,** 36, 33-43.

Al-Saleh, M.F., Al Kadiri, M.A. 2000. Double Ranked Set Sampling. Statistics & Probability Letters, 48, 205-212.

Al-Saleh, M.F., Al-Omari, A.I., 2002. Multistage Ranked Set Sampling. Journal of Statistical Planning And Inference, 102, 273-286.

Al-Omari, A.I., Jaber, K.H., 2008. Percentile Double Ranked Set Sampling. Journal of Mathematics and Statistics, 4(1), 60-64.

Bhoj, D.S., Ahsanullah, M., 1996. Estimation of Parameters of the Generalized Geometric Distribution using Ranked Set Sampling. Biometrics, 52, 685-694.

Çıngı, H., 2009. Örnekleme Kuramı. Üçüncü Baskı , Bizim Büro Basımevi Yayın Dağıtım Ticaret Limited Şirketi, Ankara.

Li, D., Sinha, B.K., Perron, F., 1999. Random Selection in Ranked Set Sampling and its Applications. Journal of Statistical Planning and Inference, 76,185-201.

Mclntyre, G.A., 1952. A Method of Unbiased Selective Sampling using Ranked Sets. Australian Journal of Agricultural Reseach**,** 3, 385-390.

Muttlak, H.A., 1997. Median Ranked Set Sampling. Applied Statistical Science**,** 6(4), 245-255.

Muttlak, H.A., 2003. Modified Ranked Set Sampling Methods. Pakistan Journal Statistic, 19(3), 315-323.

Patil, G.P., Sinha, A.K., Taillie, C., 1999. Ranked Set Sampling: a bibliography. Environmental and Ecological Statistics**,** 6, 91-98.

Samawi, H.M., Ahmed, M.S., Abu-Dayyeh, W., 1996. Estimating the Population Mean Using Extreme Ranked Set Sampling. Biometrical Journal**,** 38(5), 577-586.

Shen, W.H., 1994. On Estimation of a Log-Normal Mean Using a Ranked Set Sample. Sankhya, 54, 323-333.

Sinha, Bimal. K., Sinha, Bikas K., Purkayastha, S., 1996. On Some Aspects of Ranked Set Sampling for Estimation of Normal and Exponential Parameters. Statistical Decisions**,** 14, 223-240.

Takahasi, K., Wakimoto , K., 1968. On Unbiased Estimates of The Population Mean Based on The Sample Stratified By Means of Ordering. Annals of The Institude of Statistical Mathematics, 21, 249-255.

Özdemir, Y.A., 2005. Sıralı Küme Örneklemesiyle Doğrusal Regresyon Modelinde Parametre Tahminlerinin İncelenmesi. Doktora Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.