İKİ SAFHALI ÖRNEKLEME YÖNTEMİNDE MEDYAN TAHMİN EDİCİLERİ

Sibel AL[1]
Hülya ÇINGI[2]

Özet: Kitleye ilişkin çeşitli parametrelerin tahmin edilmesinde yardımcı değişken bilgisinin kullanımına sık rastlanılmaktadır. Bazı çalışmalarda yardımcı değişkene ilişkin kitle bilgisine ulaşılamadığından iki safhalı örnekleme yöntemi kullanılmaktadır. Literatürde kitle ortalaması, toplamı ve varyansının tahmini için çeşitli örnekleme yönteminde kullanılan tahmin edicilere sık rastlanmaktadır. Ancak gelir, üretim gibi değişkenlerin yer aldığı çalışmalarda değişkenler oldukça çarpık dağıldığından medyan değeri ortalamaya göre daha çok tercih edilen bir konum ölçüsü olmaktadır. Bu çalışmada amaç, iki safhalı örnekleme yönteminde literatürde günümüze kadar kitle medyanı tahmini için önerilen çeşitli oransal, çarpımsal ve regresyon tahmin edicilerini tanıtmak, bu tahmin edicilerin etkinliklerini hem teorik olarak hem de uygulamalı olarak incelenmektir.

Anahtar sözcükler: İki safhalı örnekleme, Medyan tahmin edicisi, Yardımcı değişken, Yan, Hata kareler ortalaması (HKO), Etkinlik.

Abstract: Use of auxiliary variables is very common in estimating various population parameters. The information of auxiliary variable has not been encountered in some of studies so two-phase sampling is employed. In literature estimation of population mean, sum and variance is very common in various types of sampling methods. However, variables have a highly skewed distribution, such as income, production is studied median is often regarded as more appropriate measure of location than mean. In this study, we purposed to introduce the ratio, product and regression median estimators in two-phase sampling in literature and examine the efficiency of these estimators both theoretical and practical.

Keywords: Two-Phase sampling, Median estimator, Auxiliary information, Bias, Mean square error (MSE), Efficiency.

I. Giriş

Örneklem üzerinde çalışmak, araştırmacıya zaman, para ve insan gücü bakımından tasarruf sağlar. Bazı kitlelere büyüklüğü nedeniyle örneklemenin uygulanması zorunludur. Örnekleme kuramının konusu, kitleden, kitlenin yapısına en uygun örnekleme yöntemiyle örneklem seçim süreci ve örneklemden kitlenin özelliklerinin tahmin edilmesi sürecidir. Bu iki süreç birbiriyle ilgilidir. Hangi seçim süreci kullanılmış ise, o sürece göre tahminler yapılır (Çıngı, 1994: 2).

Kitleye ilişkin çeşitli parametrelerin tahmin edilmesinde yardımcı değişken bilgisinin kullanımına sık rastlanılır. Ancak bazı çalışmalarda yardımcı değişkene ilişkin kitle bilgisine ulaşılamadığından iki safhalı

[1] Arş. Gör., Hacettepe Üniversitesi İstatistik Bölümü
[2] Prof. Dr., Hacettepe Üniversitesi İstatistik Bölümü

örnekleme yöntemi kullanılır. İki safhalı örnekleme yönteminde, ilk aşamada X yardımcı değişkenine ait bilgilerin tahmini için ön örneklem seçilir. İkinci aşamada ise ön örneklemden Y değişkeninin tahmin edilmesi için alt örneklem seçilir.

Literatürde kitle ortalaması, toplamı ve varyansının tahmini için çeşitli örnekleme yönteminde kullanılan tahmin edicilere oldukça sık rastlanmaktadır (Kadılar ve Çıngı, 2006: 1047-1059). Ancak gelir, gider, üretim gibi değişkenlerin yer aldığı çalışmalarda değişkenler oldukça çarpık dağıldığından medyan değeri ortalamaya göre daha çok tercih edilen bir konum ölçüsü olmaktadır.

Bu çalışmada kitle medyanı tahmini için önerilen çeşitli oransal, çarpımsal ve regresyon tahmin edicileri iki safhalı örnekleme yönteminde ele alınmıştır. Bu tahmin edicilere ilişkin yan ve hata kareler ortalamalarının elde edilmesi ayrıntılı olarak incelenmiştir. Tahmin ediciler birbirleriyle karşılaştırılarak hangi koşullar altında hangi tahmin edicilerin etkin olduğu araştırılmıştır. Ayrıca sayısal örnek ile tahmin edicilerin etkinlikleri incelenmiştir.

II. İki Safhalı Örnekleme Yönteminde Çeşitli Medyan Tahmin Edicileri

İki safhalı örnekleme yönteminde, N birimli sonlu bir kitleden yerine koymaksızın x ve z yardımcı değişkenlerine ilişkin kitle medyanı M_X ve M_Z'in tahmin edildiği n' birimlik bir ön örneklem seçilsin. Ön örneklemden y değişkenine ilişkin kitle medyanını tahmin etmek için n birimlik bir alt örneklem seçilsin. M_Y, M_X, M_Z kitle medyanını $\hat{M}_Y$, $\hat{M}_X$ ve $\hat{M}_Z$ alt örneklemdeki örneklem medyanlarını ve $\hat{M}_X^{'}$ ile $\hat{M}_Z^{'}$ ise x ve z yardımcı değişkenleri için ön örneklem medyanlarını ifade etmektedir. Tahmin edicilere ilişkin yan ve hata kareler ortalama değerlerini bulmak için fark yönteminden yararlanılmaktadır (Kadılar vd., 2009: 301-309). Fark yönteminde kullanılan terimler Tablo 1 ve Tablo 2'de verilmektedir.

Tablo 1: *Fark Yönteminde Kullanılan Terimler ve Beklenen Değerleri*

Dönüşüm	$E(e_i)$	$E(e_i^2)$
$e_0 = \dfrac{\hat{M}_Y - M_Y}{M_Y}$	0	$\dfrac{1}{4} f_1 \{M_Y f_Y (M_Y)\}$
$e_1 = \dfrac{\hat{M}_X - M_X}{M_X}$	0	$\dfrac{1}{4} f_1 \{M_X f_X (M_X)\}$
$e_2 = \dfrac{\hat{M}'_X - M_X}{M_X}$	0	$\dfrac{1}{4} f_2 \{M_X f_X (M_X)\}$
$e_3 = \dfrac{\hat{M}_Z - M_Z}{M_Z}$	0	$\dfrac{1}{4} f_1 \{M_Z f_Z (M_Z)\}^{-}$
$e_4 = \dfrac{\hat{M}'_Z - M_Z}{M_Z}$	0	$\dfrac{1}{4} f_2 \{M_Z f_Z (M_Z)\}^{-}$

Tablo 2: *Fark Yönteminde Kullanılan Kovaryans Terimleri*

erim	$E(e_i e_j)$	erim	$E(e_i e_j)$
	$\frac{1}{4} f_1 \{M_X M_Y f_X (M_X) f_Y (M$		$\frac{1}{4} f_1 \{M_X M_Z f_X (M_X) f_Z (M$
	$\frac{1}{4} f_2 \{M_X M_Y f_X (M_X) f_Y (M$		$\frac{1}{4} f_2 \{M_X M_Z f_X (M_X) f_Z (M$
	$\frac{1}{4} f_1 \{M_Y M_Z f_Y (M_Y) f_Z (M$		$\frac{1}{4} f_2 \{M_X M_Z f_X (M_X) f_Z (M$
	$\frac{1}{4} f_2 \{M_Y M_Z f_Y (M_Y) f_Z (M$		$\frac{1}{4} f_2 \{M_X M_Z f_X (M_X) f_Z (M$
	$\frac{1}{4} f_2 \{M_X f_X (M_X)\}^{-2}$		$\frac{1}{4} f_2 \{M_Z f_Z (M_Z)\}^{-2}$

Tablo 1 ve Tablo 2' de $f_1 = \left(\dfrac{1}{n} - \dfrac{1}{N}\right)$, $f_2 = \left(\dfrac{1}{n'} - \dfrac{1}{N}\right)$ ve $\rho_{XY} = 4P_{11}(x;y) - 1$, ($P_{11}$, $X \le M_X$ ve $Y \le M_Y$ olan birimlerin oranı) olarak tanımlanmıştır.

A.Klasik Medyan Tahmin Edicisi

Örnekleme çalışmalarında ilk olarak Gross (1980: 181-184), kitle medyanının tahminini $\hat{M}_Y$ olarak tanımlamış ve tahminin asimptotik varyansını elde etmiştir. y'nin birikimli dağılım fonksiyonu F_Y, olasılık yoğunluk fonksiyonu ise f_Y ile gösterilsin. Taylor serisi açılımından yararlanarak

$$F_Y(\hat{M}_Y) = F_Y[M_Y + (\hat{M}_Y - M_Y)]$$
$$= F_Y(M_Y) + f_Y(M_Y)(\hat{M}_Y - M_Y) + o_p(n^{-1/2}) \qquad (1)$$

yazılabilir. Burada $(\hat{M}_Y - M_Y)$ ifadesi yalnız bırakılırsa,

$$\hat{M}_Y - M_Y = \{f_Y(M_Y)\}^{-1}[F_Y(\hat{M}_Y) - F_Y(M_Y)] + o_p(n^{-1/2}) \qquad (2)$$

eşitliği elde edilir. $\hat{F}_Y$, F_Y'nin tahmini olmak üzere,

$$F_Y(\hat{M}_Y) - F_Y(M_Y) = \hat{F}_Y(\hat{M}_Y) - \hat{F}_Y(M_Y) + o_p(n^{-1/2}) \qquad (3)$$

biçiminde yazılabilir. $\hat{F}_Y(\hat{M}_Y) = 0{,}5$ ve $\hat{F}_Y(M_Y) = p_Y$ olmak üzere, (2) numaralı eşitlikte (3) numaralı eşitlik yerine konulursa,

$$\hat{M}_Y - M_Y = \{f_Y(M_Y)\}^{-1}[\hat{F}_Y(\hat{M}_Y) - \hat{F}_Y(M_Y)] + o_p(n^{-1/2})$$
$$= \{f_Y(M_Y)\}^{-1}[0{,}5 - p_y] + o_p(n^{-1/2}) \qquad (4)$$

eşitliği elde edilmektedir. $P = Q = 1/2$ olmak üzere $V(p_Y) = f_1 PQ$ olarak elde edilir. (4) numaralı eşitlikte $E(\hat{M}_Y - M_Y)^2$ işlemi yapılırsa örneklem medyanının asimptotik varyansı,

$$V(\hat{M}_Y) = \{f_Y(M_Y)\}^{-2} V(p_Y)$$
$$= f_1(4)^{-1}\{f_Y(M_Y)\}^{-2} \qquad (5)$$

olarak elde edilir.

B.Oransal Medyan Tahmin Edici

Singh, Joarder ve Tracy (2001: 33-46), iki safhalı örnekleme yönteminde oransal medyan tahmin edicisini,

$$\hat{M}_{SJT1} = \frac{\hat{M}_Y}{\hat{M}_X}\hat{M}_X^{'} \qquad (6)$$

eşitlikte görüldüğü biçimde tanımlamışlardır. Tablo 1'de verilen dönüşümler uygulanırsa tahmin edici

$$\hat{M}_{SJT1} = M_Y(1 + e_0)(1 + e_1)^{-1}(1 + e_2) \qquad (7)$$

eşitliğinde görüldüğü biçimde elde edilir. $(1 + e_1)^{-1}$ terimi binom açılımına göre açılıp, ikinci dereceden sonraki terimler ihmal edilirse,

$$\hat{M}_{SJT1} \cong M_Y(1+e_0-e_1+e_2-e_0e_1+e_0e_2-e_1e_2+e_1^2) \quad (8)$$

$f_3=\left(\frac{1}{n}-\frac{1}{n'}\right)$ olmak üzere, Tablo 1 ve Tablo 2'de ki beklenen değerlerden yararlanılarak $E(\hat{M}_{SJT1}-M_Y)$ işlemi yapılırsa, tahmin ediciye ilişkin yan,

$$Yan(\hat{M}_{SJT1}) \cong \frac{M_Y f_3}{4\{M_X f_X(M_X)\}^2}\left[1-\frac{M_X f_X(M_X)}{M_Y f_Y(M_Y)}\rho_{XY}\right] \quad (9)$$

biçiminde elde edilir. Tahmin ediciye ilişkin hata kareler ortalaması ise ikinci dereceden sonraki terimler ihmal edildiğinde,

$$E(\hat{M}_{SJT1}-M_Y)^2 \cong M_Y^2 E(e_0-e_1+e_2-e_0e_1+e_0e_2-e_1e_2+e_1^2)^2 \quad (10)$$

$$E(\hat{M}_{SJT1}-M_Y)^2 \cong M_Y^2 E(e_0^2+e_1^2+e_2^2-2e_0e_1+2e_0e_2-2e_1e_2) \quad (11)$$

$$HKO(\hat{M}_{SJT1}) \cong \frac{1}{4\{f_Y(M_Y)\}^2}\times$$

$$\left[f_1+f_3\left(\frac{M_Y f_Y(M_Y)}{M_X f_X(M_X)}\right)\left(\frac{M_Y f_Y(M_Y)}{M_X f_X(M_X)}-2\rho_{XY}\right)\right] \quad (12)$$

(12) numaralı eşitlikte görüldüğü biçimde elde edilir.

C.Singh, Joarder ve Tracy Genelleştirilmiş Tahmin Edici Ailesi

Singh, Joarder ve Tracy (2001: 33-46), Srivastava (1971: 404-407) ve Srivastava ve Jhajj (1981: 341-343)'nin önermiş oldukları ortalama için genelleştirilmiş tahmin edici sınıflarına benzer şekilde medyan tahmini için eşitlikte görüldüğü gibi bir genelleştirilmiş tahmin edici sınıfı önermişlerdir.

$$H=\left\{\hat{M}_Y^{(H)}:\hat{M}_Y^{(H)}=H\left(\hat{M}_Y,\frac{\hat{M}_X}{\hat{M}_X^{'}}\right)\right\} \quad (13)$$

$u=\frac{\hat{M}_X}{\hat{M}_X^{'}}$, $H(M_Y,1)=M_Y$ ve $\frac{\partial H}{\partial \hat{M}_Y}\Big|_{\hat{M}_Y=M_Y,u=1}=1$ olmak üzere $M_Y^{(H)}$'in $(M_Y,1)$ noktası etrafında birinci dereceden Taylor serine göre açılımı,

$$\hat{M}_Y^{(H)}=H(M_Y+(\hat{M}_Y-M_Y),1+(u-1))$$

$$=H(M_Y,1)+(\hat{M}_Y-M_Y)\frac{\partial H}{\partial \hat{M}_Y}\Big|_{\hat{M}_Y=M_Y,u=1}+(u-1)\frac{\partial H}{\partial u}\Big|_{\hat{M}_Y=M_Y,u=1} \quad (14)$$

$H_1=\frac{\partial H}{\partial u}\Big|_{\hat{M}_Y=M_Y,u=1}$ ve $u=e_2+1$ olmak üzere,

$$\hat{M}_Y^{(H)}=M_Y+(\hat{M}_Y-M_Y)+(u-1)H_1$$

$$\hat{M}_Y^{(H)}=M_Y(1+e_0)+e_2H_1 \quad (15)$$

biçiminde elde edilir. Tahmin edicinin hata kareler ortalaması,

$$E(\hat{M}_Y^{(H)} - M_Y)^2 = E(M_Y e_0 + e_2 H_1)^2$$

$$\cong E(M_Y^2 e_0^2 + e_2^2 H_1^2 + 2M_Y H_1 e_0 e_2)$$

$$HKO(\hat{M}_Y^{(H)}) = \frac{M_Y^2}{4} f_1 \{M_Y f_Y (M_Y)\}^{-2} + H_1^2 \frac{1}{4} f_2 \{M_X f_X (M_X)\}^{-2}$$

$$+\frac{2}{4} M_Y H_1 f_2 \{M_X M_Y f_X (M_X) f_Y (M_Y)\}^{-1} \rho_{XY} \qquad (16)$$

olarak elde edilir. (16) numaralı eşitlikte H_1 'e göre birinci dereceden türev alıp, sıfıra eşitlenirse optimal H_1 değeri $H_1 = -\frac{M_X f_X (M_X)}{f_Y (M_Y)} \rho_{XY}$ olarak elde edilir. (16) numaralı eşitlikte tanımlanan hata kareler ortalamasında optimal H_1 değeri yerine konulursa genelleştirilmiş tahmin edici ailesine ilişkin minimum hata kareler ortalama değeri,

$$HKO_{Min}(\hat{M}_Y^{(H)}) = \frac{1}{4\{f_Y (M_Y)\}^{-2}} \left[\left(\frac{1}{n} - \frac{1}{N} \right) - \left(\frac{1}{n} - \frac{1}{n'} \right) \rho_{XY}^2 \right] \qquad (17)$$

biçiminde elde edilir.

$$\hat{M}_{SJT1} = \frac{\hat{M}_Y}{\hat{M}_X} \hat{M}_X^{'}, \quad \hat{M}_{SJT2} = \hat{M}_Y \left(\frac{\hat{M}_X}{\hat{M}_X^{'}} \right)^{\alpha} \quad \text{ve} \quad \hat{M}_{SJT3} = \hat{M}_Y \left(\frac{\hat{M}_X^{'}}{c\hat{M}_X + (1-c)\hat{M}_X^{'}} \right)$$

tahmin edicileri genelleştirilmiş tahmin edici ailesine dahil tahmin ediciler olarak örnek verilebilir.

D.Çarpımsal Medyan Tahmin Edici

Singh ve Joarder (2001) çarpımsal medyan tahmin edicisini iki safhalı örnekleme yönteminde incelemişlerdir. Önerilen tahmin edici,

$$\hat{M}_{SJ} = \hat{M}_Y \left(\frac{\hat{M}_X}{\hat{M}_X^{'}} \right) \qquad (18)$$

olarak tanımlanmıştır. Tahmin edicide Tablo 1'de verilen dönüşümler yapılırsa,

$$\hat{M}_{SJ} = M_Y (1 + e_0)(1 + e_1)(1 + e_2)^{-1} \qquad (19)$$

$$\hat{M}_{SJ} \cong M_Y (1 + e_0 + e_1 - e_2 + e_0 e_1 - e_0 e_2 - e_1 e_2 + e_2^2) \qquad (20)$$

olarak elde edilir. Tahmin edicinin yan ve hata kareler ortalama değerleri aşağıdaki gibi elde edilir.

$$Yan(\hat{M}_{SJ}) = \frac{1}{4} f_3 \{M_X f_X (M_X) f_Y (M_Y)\}^{-1} \rho_{XY} \qquad (21)$$

$$HKO(\hat{M}_{SJ}) \cong \frac{1}{4\{f_Y(M_Y)\}^2} \times$$
$$\left[f_1 + f_3 \left(\frac{M_Y f_Y(M_Y)}{M_X f_X(M_X)} \right) \left(\frac{M_Y f_Y(M_Y)}{M_X f_X(M_X)} + 2\rho_{XY} \right) \right] \quad (22)$$

E.Regresyon Medyan Tahmin Edici

Singh, Joarder ve Tracy (2001: 33-46) iki safhalı örnekleme yönteminde regresyon medyan tahmin edicisini,

$$\hat{M}_{SJT4} = \hat{M}_Y + b(\hat{M}'_X - \hat{M}_X) \quad (23)$$

biçiminde tanımlamışlardır. Tablo 1' de tanımlanan dönüşümler yapılırsa,

$$\hat{M}_{SJT4} = M_Y(1+e_0) + bM_X(e_2 - e_1) \quad (24)$$

(24) numaralı eşitlikte tahmin ediciden kitle değeri çıkartılarak beklenen değeri alınırsa regresyon medyan tahmin edicinin yansız bir tahmin edici olduğu görülmektedir. Tahmin edicinin hata kareler ortalama değerini bulmak için $E(\hat{M}_{SJT4} - M_Y)^2$ işlemi yapılırsa,

$$E(\hat{M}_{SJT4} - M_Y)^2 = E(M_Y e_0 + bM_X e_2 - bM_X e_1)^2$$
$$= E(M_Y^2 e_0^2 + b^2 M_X^2 e_1^2 + b^2 M_X^2 e_2^2 - 2bM_X M_Y e_0 e_1$$
$$+ 2bM_X M_Y e_0 e_2 - 2b^2 M_X^2 e_1 e_2)$$

$$V(\hat{M}_{SJT4}) = \frac{1}{4}[f_1 \{f_Y(M_Y)\}^{-2}$$
$$+ b^2 f_3 f_X(M_X)\}^{-2} - 2bf_3 \{f_X(M_X) f_Y(M_Y)\}^{-1} \rho_{XY}] \quad (25)$$

biçiminde elde edilir. Eşitlik (25)'te b'ye göre birinci dereceden türev alınıp sıfıra eşitlenirse optimal b değeri $b = \frac{f_X(M_X)}{f_Y(M_Y)} \rho_{XY}$ biçiminde elde edilir. Elde edilen optimal b değeri eşitlik (25)'te yerine konulursa regresyon medyan tahmin edicisine ilişkin minimum varyans (26) numaralı eşitlikte görüldüğü gibi elde edilir.

$$V_{Min}(\hat{M}_{SJT4}) = \frac{1}{4\{f_Y(M_Y)\}^2} \left[f_1 - f_3 \rho_{XY}^2 \right] \quad (26)$$

F.Singh, Singh ve Puertas Tahmin Edicileri

Singh, Singh ve Puertas (2003: 369-382), basit rasgele örneklemede yardımcı değişkene ilişkin sabit bir değerin bilindiğini varsayarak yeni birer oransal ve çarpımsal medyan tahmin edicileri önermişlerdir. Önerdikleri bu tahmin edicileri iki safhalı örnekleme yöntemine uyarlayarak bu tahmin edicilerin yan ve hata kareler ortalama değerlerini elde etmişlerdir.

$\theta = \dfrac{M_X}{A - M_X}$ ve $\eta = \dfrac{M_X}{A + M_X}$ biçiminde tanımlandığında tahmin ediciler ve onlara ilişkin yan ve hata kareler ortalama değerleri Tablo 3'te özetlenmiştir. Burada A yardımcı değişkene ilişkin dağılım genişliği, tepe değeri gibi bir değer olabilir.

Tablo 3: *Singh, Singh ve Puertas Tahmin Edicileri*

	SSP1	**SSP2**
ahmi n Edici	$\hat{M}^{(t_1)}_{SSP1} = \hat{M}_Y \left(\dfrac{A - \hat{M}_X}{A - \hat{M}'_X} \right)$	$\hat{M}^{(t_2)}_{SSP2} = \hat{M}_Y \left(\dfrac{\hat{M}'_X + A}{\hat{M}_X + A} \right)$
an	$Yan(\hat{M}^{(t_1)}_{SSP1}) \cong -\dfrac{f_3\, \theta\rho}{4M_X f_X(M_X}$	$Yan(\hat{M}^{(t_2)}_{SSP2}) \cong \dfrac{f_3 M_Y \eta}{4\{M_X f_X(M_X)\}^2} \Big[\eta$
KO	$HKO(\hat{M}^{(t_1)}_{SSP1}) \cong \dfrac{1}{4\{f_Y(M_Y)\}^2}$ $+\theta f_3 \left(\dfrac{M_Y f_Y(M_Y)}{M_X f_X(M_X)} \right)^2 \Big\{ \theta - 2\rho$	$HKO(\hat{M}^{(t_2)}_{SSP2}) \cong \dfrac{1}{4\{f_Y(M_Y)\}^2} \Big[f_1$ $+\eta f_3 \left(\dfrac{M_Y f_Y(M_Y)}{M_X f_X(M_X)} \right)^2 \Big\{ \eta - 2\rho_{XY} \dfrac{M_X}{M_Y}$

G.Singh, Singh ve Upadhyaya Tahmin Edicileri

Singh, Singh ve Upadhyaya (2006: 23-46) zincirleme oransal ve regresyon medyan tahmin edicilerini iki safhalı örnekleme yöntemi altında incelemişlerdir. Bu tahmin edicilere ilişkin yan ve hata kareler ortalama değerlerini elde etmişlerdir. Önerilen tahmin ediciler, minimum yan ve minimum hata kareler ortalama değerleri Tablo 4'te, tahmin edicilere ilişkin minimum yan ve minimum hata kareler ortalama değerlerinin bulunmasında kullanılan optimal değerler Tablo 5'te verilmiştir.

Tablo 4: *Singh, Singh ve Upadhyaya Tahmin Edicileri*

Tahmin Edici	Yan	HKO
$\hat{M}_{SSU1} = \hat{M}_Y \left(\frac{\hat{M}'_X}{\hat{M}_X}\right)\left(\frac{M_Z}{M'_Z}\right)$	$Yan(\hat{M}_{SSU1}) \cong \frac{M_Y}{4}\left[f_3\{M_X f_X(M_X)\}^{-2}\left(1 - \frac{M_X f_X(M_X)}{M_Y f_Y(M_Y)}\rho\right.\right.$ $\left.\left. + f_2\{M_Z f_Z(M_Z)\}^{-2}\left(1 - \frac{M_Z f_Z(M_Z)}{M_Y f_Y(M_Y)}\rho_{YZ}\right)\right]$	$HKO(\hat{M}_{SSU1}) \cong \frac{1}{4\{f_Y(M_Y)\}^2}\left[f_1 + f_3 \frac{M_Y f_Y(M_Y)}{M_X f_X(M_X)}\left(\frac{M_Y f_Y(M_Y)}{M_X f_X(M_X)} - \right.\right.$ $\left.\left. + f_2 \frac{M_Y f_Y(M_Y)}{M_Z f_Z(M_Z)}\left(\frac{M_Y f_Y(M_Y)}{M_Z f_Z(M_Z)} - 2\rho_{YZ}\right)\right]$
$\hat{M}_{SSU2} = \hat{M}_Y \left(\frac{\hat{M}'_X}{\hat{M}_X}\right)^{\alpha_1}\left(\frac{M_Z}{M'_Z}\right)^{\alpha_2}$	$Yan_{Min}(\hat{M}_{SSU2}) = \frac{1}{8f_Y(M_Y)}\left[f_3\{M_X f_X(M_X)\}^{-1}\rho_{XY}\{1 - \rho_{XY} \frac{M}{M}\right.$ $\left. + f_2\{M_Z f_Z(M_Z)\}^{-1}\rho_{YZ}\{1 - \rho_{YZ} \frac{M_Z f_Z(M_Z)}{M_Y f_Y(M_Y)}\}\right]$	$HKO_{Min}(\hat{M}_{SSU2}) = \frac{1}{4\{f_Y(M_Y)\}^2}\left[f_1 - f_3\rho_{XY}^2 - f_2\rho_{YZ}^2 \right]$

$\hat{M}_{SSU3} = \hat{M}_Y \left(\frac{\hat{M}'_X}{\hat{M}_X}\right)^{\alpha_1} \left(\frac{M_Z}{\hat{M}'_Z}\right)^{\alpha_2} \left(\frac{M_Z}{\hat{M}_Z}\right)^{\alpha}$	$Yan_{Min}(\hat{M}_{SSU3}) = \frac{M_Y}{4\{M_Y f_Y(M_Y)\}^2 (1-\rho_{XZ}^2)^2} [f_3\{(\rho_{XY} - \rho_{XZ}\rho_{YZ})^2$ $-2\rho_{XY}(\rho_{XY} - \rho_{XZ}\rho_{YZ})(1-\rho_{XZ}^2) + \frac{M_Y f_Y(M_Y)}{M_X f_X(M_X)}(\rho_{XY} - \rho_{XZ}\rho_{YZ})(1-\rho_{XZ}^2)$ $+f_2\{\rho_{XZ}^2(\rho_{XY} - \rho_{XZ}\rho_{YZ})^2 - 2\rho_{XZ}\rho_{YZ}(\rho_{XY} - \rho_{XZ}\rho_{YZ})(1-\rho_{XZ}^2)$ $+\frac{M_Y f_Y(M_Y)}{M_Z f_Z(M_Z)}\rho_{XZ}(\rho_{XY} - \rho_{XZ}\rho_{YZ})(1-\rho_{XZ}^2)\}$ $+f_1\{-(\rho_{YZ} - \rho_{XY}\rho_{XZ})^2 + \frac{M_Y f_Y(M_Y)}{M_Z f_Z(M_Z)}(\rho_{YZ} - \rho_{XY}\rho_{XZ})(1-\rho_{XZ}^2)\}]$	$HKO_{Min}(\hat{M}_{SSU3}) = \frac{1}{4\{f_Y(M_Y)\}^2}\left[f_1 - f_2\rho_{YZ}^2 - f_3 R_{Y.Xz}^2\right]$

Tahmin Edici	Yan	HKO
$\hat{M}_{SSU4} = \hat{M}_Y + b_1(\hat{M}'_X - \hat{M}_X)$ $+ b_2(M_Z - \hat{M}'_Z)$	0	$HKO_{Min}(\hat{M}_{SSU4}) = \frac{1}{4\{f_Y(M_Y)\}^2}\left[f_1 - f_3\rho_{XY}^2 - f_2\rho_{YZ}^2\right]$
$\hat{M}_{SSU5} = \hat{M}_Y + b_1(\hat{M}'_X - \hat{M}_X)$ $+ b_2(M_Z - \hat{M}'_Z) + b_3(M_Z - \hat{M}$	0	$HKO_{Min}(\hat{M}_{SSU5}) = \frac{1}{4\{f_Y(M_Y)\}^2}\left[f_1 - f_2\rho_{YZ}^2 - f_3 R_{Y.Xz}^2\right]$

$$R_{Y.XZ}^2 = \frac{\rho_{XY}^2 + \rho_{YZ}^2 - 2\rho_{XY}\rho_{XZ}\rho_{YZ}}{1-\rho_{XZ}^2}$$

Tablo 5: *Tahmin Edicilere İlişkin Optimal Değerler*

Tahmin Edici	**Optimal Değerler**
$\hat{M}_{SSU2} = \hat{M}_Y \left(\frac{\hat{M}'_X}{\hat{M}_X}\right)^{\alpha_1} \left(\frac{M_Z}{M'_Z}\right)^{\alpha_2}$	$\alpha_1 = \frac{M_X f_X(M_X)}{M_Y f_Y(M_Y)} \rho_{XY}$ $\alpha_2 = \frac{M_Z f_Z(M_Z)}{M_Y f_Y(M_Y)} \rho_{YZ}$
$\hat{M}_{SSU3} = \hat{M}_Y \left(\frac{\hat{M}'_X}{\hat{M}_X}\right)^{\alpha_1} \left(\frac{M_Z}{\hat{M}'_Z}\right)^{\alpha_2} \left(\frac{M_Z}{\hat{M}_Z}\right)$	$\alpha_1^* = \frac{M_X f_X(M_X)}{M_Y f_Y(M_Y)} \left(\frac{\rho_{XZ}\rho_{YZ} - \rho_{XY}}{\rho_{XZ}^2 - 1}\right)$ $\alpha_2^* = \frac{M_Z f_Z(M_Z)\rho_{XZ}}{M_Y f_Y(M_Y)} \left(\frac{\rho_{YZ}\rho_{XZ} - \rho_{XY}}{\rho_{XZ}^2 - 1}\right)$ $\alpha_3^* = \frac{M_Z f_Z(M_Z)}{M_Y f_Y(M_Y)} \left(\frac{\rho_{XY}\rho_{XZ} - \rho_{YZ}}{\rho_{XZ}^2 - 1}\right)$
Tahmin Edici	**Optimal Değerler**
$\hat{M}_{SSU4} = \hat{M}_Y + b_1(\hat{M}'_X - \hat{M}_X) + b_2(M_Z - \hat{M}'_Z)$	$b_1 = \frac{f_X(M_X)}{f_Y(M_Y)} \rho_{XY}$ $b_2 = \frac{f_Z(M_Z)}{f_Y(M_Y)} \rho_{YZ}$
$\hat{M}_{SSU5} = \hat{M}_Y + b_1(\hat{M}'_X - \hat{M}_X) + b_2(M_Z - \hat{M}'_Z)$	$b_1^* = \frac{f_X(M_X)}{f_Y(M_Y)} \left(\frac{\rho_{XY} - \rho_{YZ}\rho_{XZ}}{1 - \rho_{XZ}^2}\right)$ $b_2^* = \frac{f_Z(M_Z)}{f_Y(M_Y)} \left(\frac{\rho_{XY} - \rho_{YZ}\rho_{XZ}}{1 - \rho_{XZ}^2}\right)\rho_{XZ}$ $b_3^* = \frac{f_Z(M_Z)}{f_Y(M_Y)} \left(\frac{\rho_{YZ} - \rho_{XY}\rho_{XZ}}{1 - \rho_{XZ}^2}\right)$

H. Gupta, Shabbir ve Ahmad Tahmin Edicisi

Gupta, Shabbir ve Ahmad (2008: 1815-1822) iki yardımcı değişken bilgisinden yararlanarak ve Z yardımcı değişkenine ilişkin dağılım genişliğinin bilindiği varsayımı altında yeni bir tahmin edici önermişlerdir. Önerilen tahmin edici,

$$\hat{M}_{GSA} = \hat{M}_Y \left(\frac{\hat{M}'_X}{\hat{M}_X}\right)^{\gamma_1} \left(\frac{M_Z + R_Z}{\hat{M}'_Z + R_Z}\right)^{\gamma_2} \left(\frac{M_Z + R_Z}{\hat{M}_Z + R_Z}\right)^{\gamma_3} \tag{27}$$

biçiminde tanımlanmıştır. $\omega = \dfrac{M_Z}{M_Z + R_Z}$ olmak üzere Tablo 1'de tanımlanan dönüşümler uygulanırsa tahmin edici,

$$\hat{M}_{GSA} = M_Y(1+e_0)(1+e_1)^{-\gamma_1}(1+e_2)^{\gamma_1}(1+\omega e_3)^{-\gamma_3}(1+\omega e_4)^{-\gamma_2} \tag{28}$$

şeklinde elde edilir. $(1+e_1)^{-\gamma_1}$, $(1+e_2)^{\gamma_1}$, $(1+\omega e_3)^{-\gamma_3}$ ve $(1+\omega e_4)^{-\gamma_2}$ terimleri binom açılımına göre açılıp, çarpım işlemi yapıldığında ikinci dereceden sonraki terimler ihmal edilirse tahmin edici,

$$\begin{aligned}\hat{M}_{GSA} \cong M_Y(&1+e_0-\gamma_1 e_1+\gamma_1 e_2-\gamma_3\omega e_3-\gamma_2\omega e_4-\gamma_1 e_0 e_1+\gamma_1 e_0 e_2-\gamma_3\omega e_0 e_3\\ &-\gamma_2\omega e_0 e_4-\gamma_1^2 e_1 e_2+\gamma_1\gamma_3\omega e_1 e_3+\gamma_1\gamma_2\omega e_1 e_4-\gamma_1\gamma_3\omega e_2 e_3-\gamma_1\gamma_2\omega e_2 e_4\\ &+\gamma_2\gamma_3\omega^2 e_3 e_4+\frac{\gamma_1(\gamma_1+1)}{2}e_1^2+\frac{\gamma_1(\gamma_1-1)}{2}e_2^2+\frac{\gamma_3(\gamma_3+1)}{2}\omega^2 e_3^2+\frac{\gamma_2(\gamma_2+1)}{2}\omega^2 e_4^2)\end{aligned} \tag{29}$$

biçiminde elde edilir. Tablo 3'te verilen $\hat{M}_{SSU3}$ tahmin edicisinin optimal α_i^* değerlerinden yararlanarak optimal γ_i değerleri, $\gamma_1 = \alpha_1^*$, $\gamma_2 = \dfrac{\alpha_2^*}{\omega}$, $\gamma_3 = \dfrac{\alpha_3^*}{\omega}$ biçiminde elde edilir. Optimal değerlerden yararlanarak tahmin ediciye ilişkin minimum yan ve minimum hata kareler ortalama değerleri aşağıdaki eşitliklerdeki gibi elde edilir.

$$\begin{aligned}Yan_{Min}(\hat{M}_{GSA}) = \frac{M_Y}{4\{M_Y f_Y(M_Y)\}^2(1-\rho_{XZ}^2)^2}\Big[&\big\{f_3\{(\rho_{XY}-\rho_{XZ}\rho_{YZ})^2\\ &-2\rho_{XY}(\rho_{XY}-\rho_{XZ}\rho_{YZ})(1-\rho_{XZ}^2)\\ &+\frac{M_Y f_Y(M_Y)}{M_X f_X(M_X)}(\rho_{XY}-\rho_{XZ}\rho_{YZ})(1-\rho_{XZ}^2)\}\\ &+f_2\{\rho_{XZ}^2(\rho_{XY}-\rho_{XZ}\rho_{YZ})^2-2\rho_{XZ}\rho_{YZ}(\rho_{XY}-\rho_{XZ}\rho_{YZ})(1-\rho_{XZ}^2)\\ &+\frac{M_Y f_Y(M_Y)}{M_Z f_Z(M_Z)}\omega\rho_{XZ}(\rho_{XY}-\rho_{XZ}\rho_{YZ})(1-\rho_{XZ}^2)\}\\ &+f_1\left\{-(\rho_{YZ}-\rho_{XY}\rho_{XZ})^2+\frac{M_Y f_Y(M_Y)}{M_Z f_Z(M_Z)}\omega(\rho_{YZ}-\rho_{XY}\rho_{XZ})(1-\rho_{XZ}^2)\right\}\Big]\end{aligned} \tag{30}$$

$$HKO_{Min}(\hat{M}_{GSA}) = \frac{1}{4\{f_Y(M_Y)\}^2}\left[f_1 - f_2\rho_{YZ}^2 - f_3 R_{Y.Xz}^2\right] \tag{31}$$

III. Çeşitli Tahmin Edicilerin Karşılaştırılması

Çeşitli tahmin edicilerin etkinlikleri hata kareler ortalamaları bakımından teorik olarak karşılaştırıldığında hangi tahmin edicinin hangi koşul altında daha etkin olduğu Tablo 6'da özetlenmiştir.

Tablo 6: *Çeşitli Tahmin Edicilerin Etkinliklerinin Karşılaştırılması*

Karşılaştırma	**Koşul**	**Karşılaştırma**	**Koşul**
$HKO(\hat{M}_{SJT1}) < V(\hat{M}_Y)$	$\rho_{XY} > \frac{M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$	$HKO(\hat{M}_{SJ}) < V(\hat{M}_Y)$	$\rho_{XY} < -\frac{M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$
$HKO(\hat{M}_Y^{(H)}) < V(\hat{M}_Y)$	$\rho_{XY}^2 > 0$, Daima	$HKO(\hat{M}_{SJT4}) < V(\hat{M}_Y)$	$\rho_{XY}^2 > 0$, Daima
$HKO(\hat{M}_{SSP1}^{(t_1)}) < V(\hat{M}_Y)$	$\theta > 0$ ise $\rho_{XY} > \frac{\theta M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$ $\theta < 0$ ise $\rho_{XY} < \frac{\theta M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$	$HKO(\hat{M}_{SSP2}^{(t_2)}) < V(\hat{M}_Y)$	$\eta > 0$ ise $\rho_{XY} > \frac{\eta M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$ $\eta < 0$ ise $\rho_{XY} < \frac{\eta M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$
$HKO(\hat{M}_{SSP1}^{(t_1)}) < HKO(\hat{M}_{SJT1})$	$\theta > 1$ ise $\rho_{XY} > \frac{(1+\theta) M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$ $\theta < 1$ ise $\rho_{XY} < \frac{(1+\theta) M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$	$HKO(\hat{M}_{SSP2}^{(t_2)}) < HKO(\hat{M}_{YO})$	$\eta > 1$ ise $\rho_{XY} > \frac{(1+\eta) M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$ $\eta < 1$ ise $\rho_{XY} < \frac{(1+\eta) M_Y f_Y(M_Y)}{2M_X f_X(M_X)}$
$HKO(\hat{M}_{SSU1}) < HKO(\hat{M}_{SJT1})$	$\rho_{YZ} > \frac{M_Y f_Y(M_Y)}{2M_Z f_Z(M_Z)}$		

IV. Uygulama

Uygulamada 'Türkiye genelinde ilk ve ortaöğretim olanaklarının incelenmesi ve belirlenen aksaklıklara çözüm önerilerinin getirilmesi' konulu projede kullanılan 2006–2007 öğretim yılında Milli Eğitim Bakanlığı'nın okullardan derlediği eğitim verilerinden yararlanılmıştır. Projede eğitim olanaklarına göre gelişmişlik düzeyleri belirlenmiştir. Bu çalışmada iki safhalı örnekleme yönteminde medyan tahmin edicilerinin etkinliklerini incelemek amacıyla orta gelişmişlik seviyesine sahip ilçeler kitle olarak ele alınmıştır. İlgilenilen değişken Y, 2006 yılında orta düzeydeki ilçelerde ÖSS'ye göre yerleşen öğrenci sayısı; ilk yardımcı değişken olan X, orta düzeydeki ilçelerde orta öğretimdeki toplam derslik sayısı ve ikinci yardımcı değişken olan Z, orta düzeydeki ilçelerde ÖSS'ye hazırlık dershane sayısı olarak alınmıştır. 340 ilçeden hoş görülebilecek hata miktarı yaklaşık olarak 50 alındığında, 150 ilçe yerine koymadan basit rasgele örnekleme yöntemi ile ön örneklem olarak seçilmiştir. Ön örneklemden ise hoş görülebilecek hata miktarı 73 alındığında basit rasgele örnekleme yöntemi ile yerine koymaksızın 50 ilçe alt örneklem olarak seçilmiştir.

Verilere ilişkin istatistikler

$$N = 340, \qquad n' = 150, \qquad n = 50,$$

$$\rho_{XY} = 0{,}88, \qquad \rho_{YZ} = 0{,}96, \qquad \rho_{XZ} = 0{,}96,$$

$$M_Y = 193, \qquad M_X = 50, \qquad M_Z = 1,$$

$$f_Y(M_Y) = 0{,}001212, \qquad f_X(M_X) = 0{,}006262$$

$$f_Z(M_Z) = 0{,}295174 \qquad R_X = 1388 \qquad R_Z = 74$$

biçiminde özetlenmiştir. Burada R_X, x yardımcı değişkenine ilişkin dağılım genişliğini, R_Z ise z yardımcı değişkenine ilişkin dağılım genişliğini ifade etmektedir. İlgilenilen değişken ile x yardımcı değişkeninin log-normal dağılıma, z yardımcı değişkeninin ise log-lojistik dağılıma Kolmogrov-Smirnov testine göre %95 güven düzeyinde uyduğu görülmüştür. Olasılık yoğunluk fonksiyonlarında medyan değerlerinin yerine konması ile $f_Y(M_Y)$, $f_X(M_X)$ ve $f_Z(M_Z)$ değerleri elde edilmiştir. Tahmin edicilerin etkinlikleri

$$\text{Etkinlik} = \frac{V(\hat{M}_Y)}{HKO(\hat{M}_j)} \times 100$$

$$(j = STJ1, STJ4, SJ, \text{regresyon}, SSP1, SSP2, SSU1, SSU2, SSU3, SSU4, SSU5, GSA)$$

eşitliği ile hesaplanmaktadır. Tahmin edicilere ilişkin HKO değerleri ve tahmin edicilerin etkinlikleri Tablo 7'de görülmektedir.

Tablo 7: *Tahmin Edicilerin Etkinlikleri*

Tahmin Edici	**HKO**	**Etkinlik**
SSU3 ($\hat{M}_{SSU3}$)	177,53	1635,40
SSU5 ($\hat{M}_{SSU5}$)	177,53	1635,40
GSA ($\hat{M}_{GSA}$)	177,53	1635,40
SSU2 ($\hat{M}_{SSU2}$)	561,64	516,92
SSU4 ($\hat{M}_{SSU4}$)	561,64	516,92
SSU1 ($\hat{M}_{SSU1}$)	619,52	468,63
Regresyon ($\hat{M}_{SJT4}$)	1145,97	253,34
SJT Genelleştirilmiş Tahmin Edici ($\hat{M}_{Y}^{(H)}$)	1145,97	253,34
Oransal ($\hat{M}_{SJT1}$)	1186,06	244,78
SSP2 ($\hat{M}_{SSP2}^{(t_2)}$)	2789,33	104,08
SSP1 ($\hat{M}_{SSP1}^{(t_1)}$)	2797,41	103,78
Klasik ($\hat{M}_{Y}$)	2903,25	100,00
Çarpımsal ($\hat{M}_{SJ}$)	7153,56	40,58

Tablo 7 incelendiğinde SSU3, SSU5 ve GSA tahmin edicilerinin en etkin tahmin ediciler olduğu görülmektedir. En düşük etkinliğe sahip tahmin edici ise Singh ve Joarder (2001)'ın önerdikleri çarpımsal medyan tahmin edicisidir. Yardımcı değişken ile ilgilenilen değişken arasındaki korelasyon yüksek olduğundan oransal tahmin edicilerin kullanılması daha uygundur. Uygulama sonucunda çarpımsal medyan tahmin edicinin ($\hat{M}_{SJ}$), en düşük etkinliğe sahip olması bu durumun bir sonucudur.

Kaynakça

GROSS, T. S., (1980), "Median Estimation in Sample Surveys", Proc. Surv. Res. Meth. Sect. Amer. Statist. Ass., ss.181 184.

GUPTA S., SHABBIR J., AHMAD S., (2008), "Estimation of Median in Two-Phase Sampling Using Two Auxiliary Variables", Commun. Statist. Theory Methods 37, ss.1815 1822.

ÇINGI, H., (1994), Örnekleme Kuramı, H.Ü. Fen Fakültesi Basımevi, Beytepe.

ÇINGI, H., KADILAR C., KOÇBERBER G., (2007), Türkiye Genelinde İlk ve Orta Öğretim Olanaklarının İncelenmesi ve Belirlenen Aksaklıklara Çözüm Önerilerinin Getirilmesi. TÜBİTAK, SOBAG, 106K077.

KADILAR, C., ÇINGI H., (2006), "Ratio Estimators for the Population Variance in Simple and Stratified Random Sampling", Applied Mathematics and Computation, 173(2), ss.1047 1059.

KADILAR, C., ÜNYAZICI, Y., ÇINGI, H., (2009), "Ratio Estimator for the Population Mean Using Ranked Set Sampling", Statistical Papers, 50, ss.301 309.

SINGH, S., JOARDER, A., TRACY, D. S., (2001), "Median Estimation Using Double Sampling", Austral. & New Zealand J. Statistics, 43(1), ss.33 46.

SINGH, S., JOARDER, A., (2001), "Estimation of Distribution Function and Median in Two Phase Sampling", Techinal Report Series TR270.

SINGH, H. P., SINGH, S., Puertas, S., (2003), "Ratio Type Estimators for the Median of Finite Populations", Allgemeines Statistisches Archiv, 87, ss. 369 382.

SINGH ,S., SINGH, H. P. , UPADHYAYA, L. N., (2006), "Chain Ratio and Regression Type Estimators for Median Estimation in Survey Sampling", Statistical Papers, 48, ss.23 46.

SRIVASTAVA, S. K., (1971), "A Generalized Estimator for the Mean of a Finite Population Using Multi-Auxiliary Information", Journal of the American Statistical Association, 66, ss.404 407.

SRIVASTAVA, S. K., JHAJJ, H. S., (1981), "A Class of Estimators of the Population Mean in Survey Sampling Using Auxiliary Information", Biometrika, 68 (1), ss.341 343.