T.C.
SELÇUK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İSTATİSTİK ANABİLİM DALI

İKİ-YÖNLÜ TABAKALI ÖRNEKLEMEDE TAMSAYI ÖRNEK HACİMLERİNİN BELİRLENMESİ

YÜKSEK LİSANS TEZİ

Danışman

Yrd. Doç. Dr. Mustafa SEMİZ

Hazırlayan

Erol GENCER
068215001007

KONYA, 2008

İÇİNDEKİLER

Sayfa

İÇİNDEKİLER ..... **i**

TABLOLAR ..... **iii**

ÖZET ..... **iv**

ABSTRACT ..... **v**

1. GİRİŞ **1**

KAYNAK TARAMASI **2**

2. TEK-YÖNLÜ TABAKALI ÖRNEKLEME **3**

2.1 Örnek Hacminin Tabakalara Dağıtılması 4

2.1.1 En Uygun Paylaştırma 5

2.1.2 Neyman paylaştırma 6

2.1.3 Orantılı Paylaştırma 6

2.1.4 Eşit paylaştırma 6

2.2 Örnek Hacminin Bulunması 7

2.2.1 Belirli bir duyarlılıkta örnek hacminin belirlenmesi 7

2.2.1.1 Eşit paylaştırma 7

2.2.1.2 Orantılı paylaştırma 8

2.2.1.3 Neyman paylaştırma 8

2.2.1.4 En uygun paylaştırma 8

2.2.2 Belirlenen maliyetle örnek hacminin belirlenmesi 9

3. İKİ-YÖNLÜ TABAKALI ÖRNEKLEME 12

3.1 Marjinal Olasılıklara Göre Cochran Orantılı Tabakalı Paylaştırma Yöntemi ... 12

3.2 Marjinal Olasılıklara Göre Matematiksel Programlama Yaklaşımı 16

3.3 İki-Yönlü Tabakalı Örneklemede Birim Seçimi İçin Alternatif Yöntem 17

4. UYGULAMA 18

4.1 Marjinal Olasılıklara Göre Cochran Orantılı Tabakalı Paylaştırma Yöntemi ile Yığın Hacminin Tahmini 21

4.2 Marjinal Olasılıklara Göre Matematiksel Programlama Yaklaşımı ile Yığın Hacminin Tahmini 23

4.3 Ortak Olasılıklara Göre Önerilen Orantılı Tabakalı Paylaştırma Yöntemi İle Yığın Hacminin Tahmini 27

5. SİMÜLASYON 29

5.1 Cochran Metodu için SPSS Programı 30

5.2 Alternatif Metod için SPSS Programı 36

6. SONUÇ 40

KAYNAKLAR 41

TABLOLAR

Tablo2.1 Üç Tabakaya Ayrılmış 60 kişilik yığın için veriler 10

Tablo3.1 Tabaka örnek hacimlerinin belirlenmesi..... 12

Tablo3.2 $n = 10$ için Örnek Çekimi 13

Tablo3.3 20 Hücreye Örneklemin Paylaşımı..... 14

Tablo4.1 Normal dağılımdan üretilecek sayıların ortalama ve varyansları 18

Tablo4.2 Normal dağılımdan üretilen sayılar 19

Tablo4.3 Üretilen sayıların ortalama ve varyansları..... 20

Tablo4.4 $n = 15$ için Örnek Çekimi 22

Tablo4.5 Cochran yöntemiyle çekilen örneklerin tabakalara paylaştırılması..... 22

Tablo4.6 Cochran yöntemi ile hücrelerden seçilen gözlemler..... 23

Tablo4.7 Matematiksel programlama yöntemi ile örnek hacminin tabakalara paylaştırılması 25

Tablo4.8 Matematiksel programlama yöntemi ile hücrelerden seçilen gözlemler 26

Tablo4.9 Alternatif yöntemde hücre içi olasılıklar ve birikimli olasılıkların hesaplanması 27

Tablo4.10 Tesadüfi sayılar ile birikimli olasılıkların karşılaştırılması..... 28

Tablo4.11 Alternatif yöntem ile örnek hacminin tabakalara paylaştırılması..... 28

Tablo4.12 Alternatif yöntem ile hücrelerden seçilen gözlemler..... 28

Tablo5.1 Tabaka hacimleri..... 29

Tablo5.2 Cochran metodu ile alternatif metodun karşılaştırılması..... 30

ÖZET

Yüksek Lisans Tezi

İKİ-YÖNLÜ TABAKALI ÖRNEKLEMEDE TAMSAYI ÖRNEK HACİMLERİNİN BELİRLENMESİ

Erol GENCER

Selçuk Üniversitesi Fen Bilimleri Enstitüsü

İstatistik Anabilim Dalı

Danışman : Yrd.Doç.Dr.Mustafa SEMİZ

2008, 43 Sayfa

Bu çalışmada, iki-yönlü tabakalı örneklemede, tabaka örnek hacimlerinin orantılı olarak tabakalardan seçilmesinde alternatif bir metot sunulmuştur. Bu amaçla 1960 yılında Bryant, Hartley ve Jessen tarafından geliştirilen ve Cochran'ın Sampling Tecniques, 2^{nd} Edition kitabında önerdiği yöntem temel olarak kabul edilmiştir. Bu yöntemin eksiklikleri ve önerilen alternatif yöntemin üstünlüğü simülasyon ile gösterilmiştir. Monte Carlo simülasyonu için SPSS v15 ve Minitab v15 programları kullanılmıştır.

Simülasyon çalışması sonuçlarına göre önerilen alternatif metodun standart hatası temel yönteme göre düşürülmüştür. Alternatif metodun etkinliği yüksek örnek hacimlerinde daha net görülmektedir.

Anahtar Kelimeler: İki-yönlü tabakalı örnekleme, tabakalı örnekleme, tamsayı tabaka hacmi, orantılı paylaştırma.

ABSTRACT

Master Thesis

DETERMINATION OF INTEGER SAMLE SIZES IN TWO-WAY STRATIFIED SAMPLING

Erol GENCER

Selcuk University

Graduate School of Natural and Applied Sciences

Department of Statistic

Supervisor : Assist. Prof. Dr.Mustafa SEMİZ

2008, 43 Pages

In this study, an alternative method was provided for determining the stratified sample sizes via proportional allocation in two-way stratified sampling. For this reason, the method which was devoloped by Bryant, Hartley and Jessen in 1960 suggested by Cochran's Sampling Techniques 2^{nd} Edition textbook was accepted as a basic method. The lack of this basic method and the superiority of the suggested alternative method were shown by simulation. For the Monte Carlo simulation, SPSS v15 and Minitab v15 packege programmes were used.

According to the result of simulation, the standart error of the suggested alternative method was to be decreased with respect to the basic method. The efficiency of the alternative method was seemed much more clear with higher sample sizes.

Keywords: Two-way stratified sampling, stratified sampling, integer stratifed sizes, proportional allocation.

1. GİRİŞ

Bir istatistiksel araştırmanın ilk aşamasında, ilk önce ilgilenilecek yığın tanımlanır. Sonraki aşamalarda, tanımlanan yığının istenilen parametreleri hakkında bilgi elde edilmeye çalışılır. Tanımlanan yığın hakkında bilgi edinmede başvurulabilecek ilk yöntem, bu yığın tanımı içinde yer alan bütün birimler üzerinden, değişken ya da değişkenlerle ilgili veri elde etmek yani tamsayım yapmaktır. Ancak tamsayım, çeşitli nedenlerle her zaman mümkün olmaz. Bu durumda, istenilen bilginin üretilebilmesi, ancak, tanımlanan yığından, onu temsil edebilecek sınırlı sayıda birimin, yani bir örneklemin seçilmesi ve bu örneklem birimleri üzerinden, gereken verilerin derlenmesiyle mümkün olur.

Örnekleme birimlerinin herhangi bir ölçüsüne ilişkin birimden birime değişim büyükse, bu durumda yığın değişkenliği daha küçük alt gruplara ayrılabilir. Birbirine yakın değerlere sahip yığın birimlerinin aynı grup içinde yer alması farklı gruplardaki birimlerde farklılık göstermesi demektir. Bu gruplara **tabaka** denir. Bu yolla yığın varyansı büyükken alt grup varyansları daha küçük olacaktır. Yani tabaka içi varyans küçük, tabakalar arasında varyans büyük olacaktır (Semiz, 2007).

Tabakalı örnekleme, çerçevesi belirlenmiş bir yığında, alt tabakalar veya alt birim gruplarının var olduğu durumlarda kullanılır. Burada asıl amaçlanan, yığın içindeki alt tabakaların varlığından yola çıkarak yığın üzerinde sonuç çıkarımı yapmaktır.

Bu tezin 2. bölümünde Tek-yönlü tabakalı örnekleme teknikleri ayrıntılı bir şekilde ele alınmıştır. 3. bölümde ise İki-yönlü tabakalı örneklemeye ilişkin bir yöntem anlatılmış, bu yöntem Matematiksel programlama yaklaşımı ile tekrar ele alınmıştır. Alternatif yöntem önerilerek sonraki bölümde bir uygulama yapılmıştır. Yapılan bu uygulama neticesinde ulaşılan sonuçlar Bulgular bölümünde ele alınmıştır.

KAYNAK TARAMASI

Snedecor (1939), sosyal bilim denemeleri için örnekleme çalışmaları yapmıştır.

Yates (1946), istatistiksel çalışmalarında ilk kez örnekleme yöntemlerini kullanmıştır.

Finkner (1950), örnekleme metodları ile Kuzey Karolina' daki şeftali ürünlerinin sayısını tahmin etmek için kullanmıştır.

Deming (1960), iş hayatı araştırmalarında kullanılmak üzere örnekleme yöntemlerini geliştirmiştir. Aynı yılda yayınladığı *Some Theory of Samplig* örnekleme alanındaki klasik kitaplardandır.

Yamane (1967), örneklem ile ilgili araştırmalar yapmış, daha önce yapılan araştırmaları derleyerek *Sampling Theory* adlı kitabını piyasaya sürmüştür.

Cochran (1977), örnekleme teknikleri adındaki klasikleşen kitabını çıkarmıştır. Bu kitapta yapılmış olan bütün teknikleri detaylı bir şekilde açıklamıştır.

Patricia ve Philip Ramsey (1981), Cochran' ın Q-Testi (1950) için en düşük örnek hacmi konusunda bir araştırma yapmışlardır.

Çıngı (1994), Örnekleme kuramı adlı kitabı yayınlanmıştır.

Semiz (2000), tabaka örnek hacmini belirlemek için tamsayılı programlama yöntemini kullanmıştır.

Semiz (2007), örnekleme yöntemleri adlı kitabın ilk baskısı çıkmıştır.

2. TEK-YÖNLÜ TABAKALI ÖRNEKLEME

Öğrencilerin fiziksel gelişimleri üzerinde çalışmak isteyen bir araştırmacının, ilköğretim 1. ve 5. sınıflar arasında okuyan öğrencileri benzer bir grup olarak düşünüp, örneklemini oluşturmak istesin. Yansız olarak yapılacak bu araştırma için araştırmacının elde edeceği örneklemde, herhangi bir sınıf düzeyindeki öğrenci sayısı diğer sınıf düzeylerine göre tesadüfen daha fazla veya daha az olabilir. Bunu engellemek için araştırmacı, 1'den 5'e kadar olan her sınıf düzeyini yığının alt tabakaları olarak düşünerek, her tabakadan belli sayıda öğrenci çekerek örneklemini oluşturabilir. Bu şekilde seçilen örneklem içinde her sınıf eşit düzeyde veya yığındaki oranı ölçüsünde temsil edilir. Böylelikle, elde edilecek bulguların yığını temsil etme gücü de o ölçüde artar (Semiz, 2007).

Kullanılacak olan bazı gösterimler şöyledir,

T : tabaka sayısı

t : tabaka indisi $(t = 1,\ldots,T)$

B : Örnekleme için ayrılan toplam bütçe

m_t : t nci tabakadan alınacak bir örneğin maliyeti

N : yığın hacmi

N_t : t nci tabakanın hacmi

n_t : t nci tabakadan alınacak örneğin hacmi

S_t : t nci tabakadaki standart sapma

A_t : t nci tabakanın ana kütledeki ağırlığı (N_t / N)

2.1. Örnek Hacminin (*n*) Tabakalara Dağıtılması

Ortalamanın veya toplam değerin tahminine ait varyansın ya da bu varyansın tahmininin hesaplanmasında tabakalardan alınacak örnek hacimlerinin (n_t) biliniyor olması gereklidir. Eğer örnek hacmi (n_t) ise ortalamaya $(\bar{y}_{tb})$ ve toplam değerin $(\hat{Y}_{tb})$ bu değerlerine ilişkin varyanslar sırasıyla şu şekildedir (Çıngı, 1994).

$$V(\bar{y}_{tb}) = \sum \left(\frac{N_t}{N}\right)^2 \frac{N_t - n_t}{N_t} \frac{S_t^2}{n_t} \quad \text{ve}$$

$$V(\hat{Y}_{tb}) = \sum_{t=1}^{T} N_t^2 \frac{N_t - n_t}{N_t} \frac{S_t^2}{n_t}$$

Tabakalardan alınacak örneklerin sayısı (n_t)'lerin toplamı, yığından alınacak örneğin hacmini (n) oluşturur. Başka bir ifadeyle;

$n = n_1 + \cdots + n_T = \sum_{t=1}^{T} n_t$ şeklinde gösterilir (Çıngı, 1994).

Tabakalı örneklemede yöntem, her tabakadan o tabakayı temsil edecek örneklerin alınmasıdır. Sonuç olarak alınan bu örneklerden toplanan bilgiler tabakaların ağırlıkları ile orantılı olarak kullanılır. Yığın ortalaması için de bu sistematik aynı şekilde çalışır. Her tabakadan alınan örnekten hesaplanan ortalama o tabakanın ortalamasının bir tahminidir ve o tabakanın toplam değerinin tahmini için kullanılır. Bu durumda, *t.* tabakanın toplam değerinin tahmini $\hat{Y}_t = N_t \bar{y}_t$ olur. Tabakalı örneklemede yığının toplam değerinin tahmini ise

$\hat{Y}_{tb} = N_1 \bar{y}_1 + \cdots + N_T \bar{y}_T = \sum_{t=1}^{T} N_t \bar{y}_t$ olur (Yamane, 1967).

Yığın ortalamasının yansız tahmini ise

$\bar{y}_{tb} = \frac{\hat{Y}_{tb}}{N} = \frac{N_1 \bar{y}_1 + \cdots + N_T \bar{y}_T}{N_1 + \cdots + N_T}$ olacaktır.

Ortalamanın nokta tahmini olan $\bar{y}_{tb}$'nin aralık tahminini bulabilmemiz için bu istatistiğin standart hatası; bunun için kendi varyansı, bütün tabakalardan alınan ortalamaların varyanslarının toplamı şeklinde yazabiliriz (Semiz, 2007).

$$V(\bar{y}_{tb}) = V\left(\frac{N_1\bar{y}_1 + \cdots + N_T\bar{y}_T}{N}\right)$$

seçim iadesiz olarak gerçekleştirilirse ve gerekli işlemler yapılırsa;

$V(\bar{y}_{tb}) = \frac{1}{N^2}\sum_{t=1}^{T} N_t(N_t - n_t)\frac{S_t^2}{n_t}$ şeklinde bulunur.

Yeniden bir düzenleme yapılırsa;

İadesiz seçimde $: V(\bar{y}_{tb}) = \frac{1}{N^2}\sum\frac{(N_tS_t)^2}{n_t} - \frac{1}{N^2}\sum N_tS_t^2$,

İadeli seçimde $: V(\bar{y}_{tb}) = \frac{1}{N^2}\sum\frac{(N_tS_t)^2}{n_t}$

Eğer, (n) örnek hacmi biliniyorsa, bu durumda hangi tabakadan ne kadar birim seçileceğine dört farklı şekilde karar verilebilir. Çünkü tabakaların yapıları (hacimleri, varyansları veya maliyetleri) birbirinden farklı olabilir. Bu sebepler ve yöntemler şu şekildedir:

2.1.1. ***En uygun paylaştırma***: Eğer, tabakaların hacimleri (N_t), varyansları ve tabakalardan alınacak örneklerin maliyetleri (m_t) tabakadan tabakaya değişiyorsa en uygun paylaştırma yöntemi kullanılır. Örnek hacmi (n) tabakalara paylaştırılır ve hangi tabakadan ne kadar birim çekileceğine karar verilir.

Örnek hacminin tabakalara hacimleri, varyansları ve maliyetleri ile orantılı en uygun paylaştırılması:

$$n_t = \frac{N_tS_t/\sqrt{m_t}}{\sum_{t=1}^{T} N_tS_t/\sqrt{m_t}}n; \quad V(\bar{y}_{eu}) = \frac{1}{N^2}\frac{1}{n}\left(\sum_{t=1}^{T} N_tS_t\sqrt{m_t}\right)\left(\sum_{t=1}^{T}\frac{N_tS_t}{\sqrt{m_t}}\right) - \frac{1}{N^2}\sum_{t=1}^{T} N_tS_t^2$$

2.1.2. ***Neyman paylaştırma*****:** Tabakadan tabakaya örnek birimlerinin maliyeti aynı fakat tabaka hacimleri ve varyansları farklılık gösteriyorsa bu kez Neyman paylaştırma yöntemi ile (n) tabakalara paylaştırılır.

Örnek hacminin tabakalara hacimleri ve varyansları ile orantılı olarak Neyman paylaştırması:

$$n_t = \frac{N_t S_t}{\sum_{t=1}^{T} N_t S_t} n \text{ yazılırsa } V(\bar{y}_{Ney}) = \frac{1}{N^2} \frac{\left(\sum_{t=!}^{T} N_t S_t\right)^2}{n} - \frac{1}{N^2} \sum_{t=1}^{T} N_t S_t^2$$

2.1.3. ***Orantılı paylaştırma*****:** Tabakadan tabakaya maliyetler ve varyanslar büyük değişiklik göstermiyorsa sadece tabakaların hacimleri birbirinden önemli derecede farklı ise örnek hacmi (n) tabaka hacimleri ile orantılı olarak tabaklara paylaştırılır.

Örnek hacminin tabakalara hacimleri ile orantılı paylaştırılması:

$$n_t = \frac{N_t}{N} n \text{ yazılırsa } V(\bar{y}_{or}) = \frac{1}{N} \sum_{t=1}^{T} \frac{N_t S_t^2}{n} - \frac{1}{N^2} \sum_{t=1}^{T} N_t S_t^2 \text{ elde edilir.}$$

2.1.4. ***Eşit paylaştırma*****:** Tabakalar arasında maliyet, varyans ve hacim farklılığı yoksa her tabakadan eşit oranda ya da eşit sayıda örnek birimi almak uygun olacaktır. Bu durumda, örnek hacmi (n) tabakalara eşit olarak paylaştırılacaktır (Semiz, 2007).

Örnek hacminin tabakalara eşit paylaştırılması:

$$n_t = \frac{n}{T} \text{ yazılırsa } V(\bar{y}_{eş}) = \frac{T}{N^2} \sum_{t=1}^{T} \frac{N_t^2 S_t^2}{n} - \frac{1}{N^2} \sum_{t=1}^{T} N_t S_t^2$$

2.2. Örnek Hacminin Bulunması

Örnek hacminin belirlenmesinde iki temel nokta vardır:

i) Belirli bir duyarlılıkla araştırmanın yapılması,

ii) Belirli bir maliyetle yapılmasıdır.

2.2.1. Belirli bir duyarlılıkta örnek hacminin belirlenmesi:

Bu aşamada hangi paylaştırma yöntemini kullanmamız gerektiğine karar vermemiz gerekmektedir. Daha önce gösterildiği gibi tabaka hacimlerini, varyanslarını ve her bir tabakadan seçilecek 1 birim örneklemin maliyetlerini bilmek gerekir. Her durum için örnek hacmini belirleyecek formüller şu şekildedir;

Duyarlılıkla ilgili olarak $d^2 = z^2 V(\bar{y}_{tb})$ olmak üzere güvenilirlik sabit tutulmak üzere duyarlılık, tahmin edicinin arzu edilen varyansına göre belirlenebilir. Bu durumda arzu edilen varyans $V(\bar{y}_{tb}) = \frac{d_0^2}{z_0^2} = D^2$ olmak üzere her paylaştırma yöntemi için örnek hacimleri şu şekilde belirlenir (Yamane, 1967).

2.2.1.1. *Eşit paylaştırma uygulanacak ise arzu edilen varyans ve örnek hacmi formülü:*

İadesiz örneklemede

$D^2 = \frac{T}{N^2} \sum_{t=1}^{T} \frac{N_t^2 S_t^2}{n} - \frac{1}{N^2} \sum_{t=1}^{T} N_t S_t^2$ ve sonuç olarak örnek hacmi

$n = \frac{T \sum_{t=1}^{T} N_t^2 S_t^2}{N^2 D^2 + \sum_{t=1}^{T} N_t S_t^2}$ olur.

İadeli örnekleme için

$D^2 = \frac{T}{N^2}\sum_{t=1}^{T}\frac{N_t^2 S_t^2}{n}$ ve sonuç olarak örnek hacmi

$n = \frac{T\sum_{t=1}^{T} N_t^2 S_t^2}{N^2 D^2}$ olur (Yamane, 1967).

2.2.1.2. ***Orantılı paylaştırma uygulanacaksa;***

İadesiz örneklemede : $n = \frac{N\sum_{t=1}^{T} N_t S_t^2}{N^2 D^2 + \sum_{t=1}^{T} N_t S_t^2}$

İadeli örneklemede : $n = \frac{N\sum_{t=1}^{T} N_t S_t^2}{N^2 D^2}$ (Yamane, 1967).

2.2.1.3. ***Neyman paylaştırma yönteminde;***

İadesiz örneklemede : $n = \frac{\left(\sum_{t=1}^{T} N_t S_t\right)^2}{N^2 D^2 + \sum_{t=1}^{T} N_t S_t^2}$

İadeli örneklemede : $n = \frac{\left(\sum_{t=1}^{T} N_t S_t\right)^2}{N^2 D^2}$ (Yamane, 1967).

2.2.1.4. ***En uygun paylaştırma yönteminde;***

İadesiz örneklemede : $n = \frac{\left(\sum_{t=1}^{T} N_t S_t \sqrt{m_t}\right)\left(\sum_{t=1}^{T} N_t S_t / \sqrt{m_t}\right)}{N^2 D^2 + \sum_{t=1}^{T} N_t S_t^2}$

İadeli örneklemede : $n = \dfrac{\left(\sum_{t=1}^{T} N_t S_t \sqrt{m_t}\right)\left(\sum_{t=1}^{T} N_t S_t / \sqrt{m_t}\right)}{N^2 D^2}$

Tabakaların hacimleri çerçeveyi oluşturma zorunluluğu sebebiyle bilinmesi gerekmektedir. Eğer örneklemede bir bütçe kısıtlaması varsa bu maliyetler de bilinmelidir.

Diğer taraftan hesaplamada kullanılacak tabaka varyanslarının bilinmesine imkan yoktur. Bu noktada, bu değerlerin daha önce yapılmış olan çalışmalardan tahminlerini veya pilot çalışma yapılarak kestirim değerlerini kullanmak zorunluluğu doğmaktadır (Semiz, 2007).

2.2.2. Belirlenen maliyetle örnek hacminin belirlenmesi:

Örnekleme için ayrılan toplam bütçe B ve sabit masraflar için ayrılan kısmı M_s toplam bütçeden çıkarıldıktan sonra örnekleme birimleri için kalan bütçe $B - M_s$ olacaktır. Örnekleme için kalan bu bütçeye göre alınacak en uygun örnek hacmi; tabaka hacmi, varyans ve tabakadan alınacak bir birimin maliyetine göre şu eşitlikle hesaplanır (Semiz, 2007).

$$n = \frac{(B - M_s)\sum_{t=1}^{T} N_t S_t / \sqrt{m_t}}{\sum_{t=1}^{T} N_t S_t \sqrt{m_t}}$$

Buradan hesaplanan örnek hacmi tabakalara aşağıdaki formülle paylaştırılır;

$$n_t = \frac{N_t S_t / \sqrt{m_t}}{\sum_{t=1}^{T} N_t S_t / \sqrt{m_t}} n$$

Şu anda Örnekleme derslerinde okutulan kitapların birçoğunda maliyet kısıtını içine alan Lagrange tekniğine dayalı çözümler önerilmektedir. Bu yöntemi klasik yöntem olarak isimlendirelim. Bu yöntemin çok önemli 2 eksiği vardır:

Bu eksikliklerin ilki; modele sadece maliyet kısıtını eklediğinden çözümünde örnek hacimleri ara sıra tabaka hacimlerinden daha yüksek çıkabilmektedir. Yani $n_t \le N_t$ kısıtı ihmal edilebilmektedir (Semiz ve Esin, 2000).

İkinci eksiklik ise; klasik yöntemde bulunan sonuçlar tamsayı olmadığından yığından alınacak örneğin hacmi (n)'in belirlenmesi sırasında ve (n)'in dağıtılarak tabakalardan alınacak örnek hacimleri (n_t)'nin belirlenmesi sırasında yuvarlamalar yapılmaktadır.

Tabaka varyansları, hacimleri ve tabakalardan alınacak her bir örneğin maliyetinin tabakadan tabakaya önemli ölçüde karmaşıklık gösterdiği durumlarda yuvarlama, araştırmacıyı zor durumda bırakabilir. Problem bir optimizasyon problemi şeklinde ele alınıp doğrusal olmayan tamsayılı programlama biçiminde ifade edilip çözülürse bu tür bir çözüm daha doğru, daha gerçekçi ve objektif olacaktır (Semiz ve Esin, 2000).

A_t : t nci tabakanın yığın hacmindeki oranı, (N_t / N) olmak üzere;

Örnek 2.1

Üç tabakaya ayrılmış 60 kişilik yığına ait bilgiler aşağıdadır. Örnekleme için kalan bütçe 100 birim ise klasik en uygun paylaştırma şu şekildedir,

TABLO 2.1

Üç tabakaya ayrılmış 60 kişilik yığın için veriler

t	N_t	A_t	S_t	m_t	$(A_tS_t)^2$	$N_tS_t/\sqrt{m_t}$	$N_tS_t\sqrt{m_t}$
1	18	0.30	1.50	3	0.2025	15.588	46.765
2	27	0.45	2.00	4	0.81	27.000	108.000
3	15	0.25	35.00	5	76.5625	234.787	1173.935

$$n = \frac{(B - M_s)\sum_{t=1}^{T} N_t S_t / \sqrt{m_t}}{\sum_{t=1}^{T} N_t S_t \sqrt{m_t}} = \frac{100\,(15.588 + 27.000 + 234.787)}{(46.765 + 108.000 + 1173.935)} = 20.875 \to 20$$

$n_t = \dfrac{N_t S_t / \sqrt{m_t}}{\sum_{t=1}^{T} N_t S_t / \sqrt{m_t}} n$ formülü ile hesaplanırsa $(n_t) = \begin{pmatrix} 1.123 \\ 1.946 \\ 16.929 \end{pmatrix} \rightarrow \begin{pmatrix} 1 \\ 2 \\ 17 \end{pmatrix}$

Görüldüğü gibi 3. tabakadan 17 birim örnek alınması gerekirken tabaka hacmi sadece 15 birimdir yani çözüm uygun olmayacaktır (Semiz ve Esin; 2000).

Ancak çözüm Semiz' in (2000) önerdiği metodla yani doğrusal olmayan tamsayılı programlama ile çözülmek istenirse;

Amaç fonksiyonu,

Enk $\sum_{t=1}^{T} \dfrac{A_t^2 S_t^2}{n_t}$

şekline olmalı ve kısıtlar da

$1 \le n_t \le N_t$ ve $\sum_{t=1}^{T} m_t n_t \le B$

olarak belirlenir. Bu şekilde çözüm;

$(n_t) = \begin{pmatrix} 3 \\ 4 \\ 15 \end{pmatrix}$

şeklinde bulunur. Bu çözüm bütün kısıtları sağladığından bu örnek için en uygun çözümdür (Semiz ve Esin, 2000).

3. İKİ-YÖNLÜ TABAKALI ÖRNEKLEME

3.1.Marjinal Olasılıklara Göre Cochran Orantılı Tabakalı Paylaştırma Yöntemi

S satır ve K kolondan oluştuğu bilinen SK hücrede, tabakalama değişkenin iki farklı özelliği için planlansın. Eğer $n \geq \text{SK}$ ise her hücre, denemede temsil edilir, diğer bir deyişle eğer n değeri planlanan S satır ve K kolon sayılarının çarpımlarından büyükse her hücreden en az bir eleman örneğe alınır. $n < \text{SK}$ olduğu zaman bir problem ortaya çıkar. Ancak denemede her bir değişkenin orantılı olarak temsil edilmesi istenir. Bu konuda basit bir yöntem Bryant, Hartley ve Jessen tarafından 1960 yılında geliştirilmiştir. Bu yöntem bir örnekle açıklanırsa, *165* okulu bir yığın farz edilsin, şehir büyüklüğü yönünden beş sınıf ve her bir öğrenci başına düşen ortalama masraf yönünden dört sınıf olarak tabakalar düzenlenmiştir. Okul numaraları m_{ij} ve bunların toplam okul sayısına oranları $P_{ij} = m_{ij} / 165$ olarak 20 hücrenin her birindeki değerleri Tablo 3.1'de gösterilmiştir (Cochran, 1977).

TABLO 3.1

Tabaka örnek hacimlerinin belirlenmesi

Şehrin büyüklüğü		Öğrenci başına ortalama masraf					Toplam	$n_{i.}$
		A	*B*	*C*	*D*			
I	m_{1j}	*15*	*21*	*17*	*9*	$m_{1.}$	*62*	
	P_{1j}	*.091*	*.127*	*.103*	*.055*	$P_{1.}$	*.376*	***4***
II	m_{2j}	*10*	*8*	*13*	*7*	$m_{2.}$	*38*	
	P_{2j}	*.061*	*.049*	*.079*	*.042*	$P_{2.}$	*.231*	***2***
III	m_{3j}	*6*	*9*	*5*	*8*	$m_{3.}$	*28*	
	P_{3j}	*.036*	*.055*	*.030*	*.049*	$P_{3.}$	*.170*	***2***
IV	m_{4j}	*4*	*3*	*6*	*6*	$m_{4.}$	*19*	
	P_{4j}	*.024*	*.018*	*.036*	*.036*	$P_{4.}$	*.114*	***1***
V	m_{5j}	*3*	*2*	*5*	*8*	$m_{5.}$	*18*	
	P_{5j}	*.018*	*.012*	*.030*	*.049*	$P_{5.}$	*.109*	***1***
Toplam	$m_{.j}$	*38*	*43*	*46*	*38*		*165*	
	$P_{.j}$	*.230*	*.261*	*.278*	*.231*		*1.000*	
	$n_{.j}$	***2***	***3***	***3***	***2***			

Amaç her bir okula yaklaşık olarak eşit seçim şansı vererek her bir marjinal sınıfı orantılı olarak temsil etmektir. Bu örneği $n = 10$ için yapacak olursak

$n_{i.} = nP_{i.}$ ve $n_{.j} = nP_{.j}$ sayıları hesaplanır ve en yakın tamsayıya yuvarlanır. Bu değerler Tablo3.1'de gösterilmiştir (Cochran, 1977).

Sonraki adım, $n = 10$ için hücrelerden $n_{i.}\, n_{.j} / n^2$ olasılıkla ij. hücrenin çekilmesidir. Burada $n \times n$'lik kare matris şeklinde tasarlanır. Tasarlanan matriste satır veya kolon üzerinden işlemler yapılabilmektedir. Biz satır işlemlerini seçelim. Satır 1, bir kolonda rasgele çekilir. Satır 2 artakalan kolonlardan birinden rasgele çekilir ve süreç bu şekilde devam eder. Sonunda her bir satır ve kolon bir adet birimi içerir. Bu şekildeki bir sonuç Tablo3.2 de gösterilmiştir (Cochran, 1977).

TABLO 3.2

n = 10 için Örnek Çekimi

		1	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
			A		***B***			***C***		***D***	
1		×									
2	***I***				×						
3			×								
4								×			
5	***II***						×				
6									×		
7	***III***			×							
8										×	
9	***IV***					×					
10	***V***										×

Dikkat edilirse kolon 1 ve 2 , $n_{i.} = 2$ olduğundan marjinal tabaka A' ya atanır. Aynı şekilde $n_{.j} = 4$ olduğundan Satır 1' den 4' e kadarı da marjinal tabaka I' ya atanır, süreç böylece devam eder. 20 hücrede örneklerin paylaşımı tamamlanır.

Paylaşım Tablo3.3' te daha anlaşılır bir şekilde gösterilmektedir. IA hücresine 15 okuldan 2' si rasgele olarak seçilmiştir. i. satır ve j. kolondaki okulun seçilme olasılığı $n_{i.}\, n_{.j} / P_{ij}$ şeklindedir. Bu sebeple her satır ve kolonun seçilme olasılıkları eşit olmayacaktır (Cochran, 1977).

TABLO 3.3

20 Hücreye Örneklemin Paylaşımı

	A	***B***	***C***	***D***	**Toplam**
I	*2*	*1*	*1*	*0*	***4***
II	*0*	*0*	*2*	*0*	***2***
III	*0*	*1*	*0*	*1*	***2***
IV	*0*	*1*	*0*	*0*	***1***
V	*0*	*0*	*0*	*1*	***1***
Toplam	***2***	***3***	***3***	***2***	***10***

Her bir hücrenin ortalamasının yansız tahmini aşağıdaki formülle hesaplanır;

$$\bar{y}_U = \frac{1}{n}\sum \frac{n^2 P_{ij}}{n_{i.}\, n_{.j}} y_{ij}$$

y_{ij}, $i.$ sütun ve $j.$ kolondaki toplam örneklemdir. Bununla beraber yansızlık önemli ise $P_{ij} = n_{i.}\, n_{.j} / n^2$, $\bar{y}$ için daha iyi bir tahmin edicidir. Varyans için de yanlı ve yansız tahmin edicilere ulaşılabilir, örnek hacmi n'in iyi temsil edilebilmesi her satır ve kolondan en az iki birimin örneğe alınması ile sağlanır, bu da örnek hacminin en az satır sayısı S ve kolon K sayılarının çarpımının iki katı olması ile mümkün olmaktadır (Cochran, 1977).

Eğer bazı hücrelerde P_{ij} önemli derecede $n_{i.}\, n_{.j} / n^2$'den ayrılırsa, ekstra adım olarak, okulların neredeyse daha sabit seçilme olasılıklarını korur. $n_{i.}$ ve $n_{.j}$ hesapladıktan sonra nicelikleri $D_{ij} = nP_{ij} - n_{i.}\, n_{.j} / n$ ile incelenir ve bunlar tamsayıya yuvarlanır. Eğer herhangi bir hücrede D_{ij} pozitif tamsayıysa bu hücrelere D_{ij} otomatik olarak atanır. Örnek hacmi n , $n_{i.}$ ve $n_{.j}$ paylaştırıldıktan sonraki miktarına bağlı olarak azaltılabilir (Cochran, 1977).

Yığın ortalamasının tahmini;

$\bar{y}_{tb} = \frac{N_1\bar{y}_1+\cdots+N_T\bar{y}_T}{N_1+\cdots+N_T}$

şeklinde hesaplanır ve bu ortalamaya ilişkin varyans;

$$\hat{V}\ (y_{tb}) = \frac{\sum_{i=1}^{ds}(\bar{y}_i - \bar{y}_{tb})}{ds-1}$$

yine bu ortalamaya ilişkin standart hata;

$$s_{\bar{x}}^2 = \sqrt{\hat{V}\ (y_{tb})}$$

formülleri ile hesaplanır.

Cochran'ın önerdiği metod temel olarak üç adımdan oluşmaktadır;

Adım I: Cochran' ın önerdiği metotla $n \times n$' lik çapraz tablo oluşturulur.

Adım II: Kolonlar ve satırlar tabakalardan alınacak örnek hacimlerine göre tabaka sayıları kadar bölünürler. Bunu belirlemek için her tabakanın hacmi ile yığın hacmi oranlanır bulunan sonuç en yakın tamsayıya yuvarlanır. $P_{ij} = \frac{m_{ij}}{N}$, $n_{ij} = n.P_{ij}$

Adım III: İşleme satır veya kolon işlemlerinden bir tanesi tercih edilerek başlanır. Bu örnekte satır işlemleri tercih edilmiştir. Bu durumda 1. satırdan başlamak üzere 1 ile n arasında tesadüfi bir kolon belirlenir, bu kolon x. kolon olsun $(x < n)$. 1. satır ve x. kolon işaretlenir. Daha sonra 2. satıra geçilir ve x. kolon dışında tesadüfi bir kolon belirlenir. Bu işlem her satırda tesadüfi bir kolon olmak üzere tamamlanır.

3.2. Orantılı Tabakalı Paylaştırma Yönteminde Matematiksel Programlama Yaklaşımı

Cochran' ın önerdiği metod için kısıtlar ve amaç faklı bir şekilde düşünülür ve bir tamsayılı programlama şeklinde ele alınsın. Bunu yapmaktaki amaç, Cochran metodu ile en iyi sonucu elde ederek yığın ortalamasının daha iyi tahminini gerçekleştirmektir. Bunun için;

Amaç fonksiyonu;

$max = n_{11} + \ldots + n_{T1T2}$;

$n_{t1.}$birinci tabakadan alınması gereken örnek hacmi olmak üzere kısıtlar;

$\sum_{t1=1}^{T1} n_{t1.} = {}^{P_{t1.}}/_{n}$;

$n_{t2.}$ ikinci tabakadan alınması gereken örnek hacmi olmak üzere kısıtlar;

$\sum_{t2=1}^{T2} n_{t2.} = {}^{P_{t2.}}/_{n}$;

örnek hacmi için kısıt;

$\sum_{t1=1}^{T1} \sum_{t2=1}^{T2} n_{t1t2} = n$;

ve n_{t1t2} alt tabakalardan alınacak örnek hacmi olmak üzere bu örnekler ondalıklı sayı olamayacağından tamsayı olma kısıtları;

$@gin(n11);$

⋮

$@gin(nt1t2);$

şeklinde LINGO programı yazılarak problem çözülür.

3.3. İki-Yönlü Tabakalı Örneklemede Birim Seçimi İçin Alternatif Yöntem

Marjinal olasılıkları kullanmak; alt tabakaların, örnek hacimlerinden bağımsız olarak hesaplanarak örneğe alınacak birimleri seçmektir. Ancak yığın hakkında daha iyi bilgi edinmek için örnek hacmi fazla olan tabakalardan daha çok birim örneğe alınmalıdır. Bunun için alternatif bir yöntem belirlenmiştir. Bu programın amacı tabaka hacmi daha çok olan alt tabakalardan örneğe daha çok birim almaktır. Böylece daha az örnek seçerek yığın hakkında daha iyi tahmin yapmak amaçlanmaktadır.

Temel olarak $n \geq T_1 \cdot T_2$ olduğu durumlarda kullanılacaktır. Kutulardaki yığın hacmi, toplam yığın hacmine oranlanarak yığın hacim oranları hesaplanır ve bir kolona yazılır. Bir kolon daha hazırlanarak hacim oranları yığılımlı olarak hesaplanır ve yazılır. Daha sonra Düzgün (0-1) aralığından rastgele sayı çekilir, çekilen bu sayı yığılımlı oranlar arasında taranır ve düştüğü aralığa 1 birim örnek atanır. Bu yöntem için hazırlanan algoritma aşağıdaki gibidir;

Adım I: Her alt tabakaya 1 birim başlangıç olarak atanır.

Adım II: Her alt tabaka için ilgili tabakalara ait yığın hacim oranları (π_{ij}) alt alta yazılır ve yanına bir kolon yazılarak bu kolona birikimli π_{ij} hesaplanarak yazılır.

Adım III: İlk $T_1 \cdot T_2$ birim, hücrelere atandığından düzgün $0-1$ aralığından tesadüfi olarak $(n - T_1 \cdot T_2)$ kadar sayı belirlenir.

Adım IV: Seçilen bu sayılardan 1. si birikimli ortak olasılıkların içinde taranır ve hangi satır bu sayıyı kapsıyorsa o satıra 1 eklenir. Daha sonra rasgele sayılardan 2. birikimli ortak olasılıkların içinde taranır, yine hangi satır bu sayıyı kapsıyorsa o satıra 1 eklenir. Bu işlem bu şekilde $n - T_1 \cdot T_2$ tane seçilen rasgele sayı kadar yapılır.

Adım V: Daha sonra hangi satırda kaç birim seçilmişse bunlar alt tabakalara dağıtılır.

4. UYGULAMA

Üniversite öğrencilerinin geldikleri coğrafi bölge ve yaşadıkları yerleşim yerlerine göre tabakalanarak aylık ortalama masrafları ve varyansları aşağıdaki tabloda gösterilen gibi olduğu düşünülsün. Amacımız 3. Bölümde verilen üç yöntem ile bu tabakalardan 15 birimlik bir örnek seçmek ve hangi yöntem için yığın ortalamasının tahmininin daha iyi olduğunu araştırmaktır.

Bu uygulama için Minitab programının 15. sürümü yardımıyla Tablo4.1' de verilen değerlere göre Normal dağılımdan sayılar üretilmiş ve üretilen sayılar düzenlenerek Tablo4.2' de gösterilmiştir.

Tablo4.1

Normal dağılımdan üretilecek sayıların ortalama ve varyansları

	İl	**İlçe**	**Köy**
Karadeniz	$\mu_{11} = 500$ $\pi_{11} = 500$ $N_{11} = 20$	$\mu_{12} = 700$ $\pi_{12} = 900$ $N_{12} = 20$	$\mu_{13} = 400$ $\pi_{13} = 200$ $N_{13} = 50$
Marmara	$\mu_{21} = 1000$ $\pi_{21} = 900$ $N_{21} = 30$	$\mu_{22} = 800$ $\pi_{22} = 900$ $N_{22} = 10$	$\mu_{23} = 500$ $\pi_{23} = 500$ $N_{23} = 20$
Akdeniz	$\mu_{31} = 700$ $\pi_{31} = 700$ $N_{31} = 30$	$\mu_{32} = 800$ $\pi_{32} = 800$ $N_{32} = 20$	$\mu_{33} = 900$ $\pi_{33} = 900$ $N_{33} = 30$
Ege	$\mu_{41} = 800$ $\pi_{41} = 900$ $N_{41} = 20$	$\mu_{42} = 800$ $\pi_{42} = 900$ $N_{42} = 30$	$\mu_{43} = 600$ $\pi_{43} = 600$ $N_{43} = 20$

Tablo4.2

Normal dağılımdan üretilen sayılar

	İl			İlçe			Köy				
Karadeniz	500	497		703	690		405	387	379	406	407
	517	488		688	635		416	375	399	407	398
	504	476		636	714		368	386	415	374	420
	519	503		736	670		415	411	394	422	412
	490	498		677	663		408	414	394	394	399
	496	455		722	694		395	399	387	398	409
	472	525		675	676		429	401	393	395	405
	499	554		709	711		401	415	371	394	416
	477	501		714	716		399	399	398	397	407
	461	497		701	700		407	388	390	390	404
Marmara	984	967	1032	787			498	526			
	1029	1020	1002	837			505	475			
	970	981	984	789			495	500			
	1061	1014	1046	788			487	491			
	949	1009	984	809			546	521			
	1004	1020	1029	771			502	500			
	962	962	1037	794			487	478			
	1021	948	973	800			515	536			
	998	1027	932	819			529	506			
	973	947	1020	766			511	498			
Akdeniz	711	712	656	790	753		883	858	929		
	719	698	724	754	784		875	881	903		
	691	710	702	824	811		871	894	953		
	716	697	687	819	788		899	908	900		
	711	693	685	790	791		878	935	921		
	646	690	694	783	802		910	858	895		
	704	722	754	816	801		926	900	915		
	670	699	662	810	852		863	869	902		
	663	696	691	802	791		829	888	922		
	683	724	698	782	821		894	900	917		
Ege	777	817		787	807	808	595	601			
	787	789		781	827	871	602	635			
	773	801		804	801	835	620	591			
	763	786		846	816	750	636	582			
	853	793		771	796	771	597	617			
	803	802		798	869	812	624	618			
	774	767		764	792	813	616	603			
	738	773		770	800	769	567	600			
	828	783		802	790	778	590	613			
	777	811		783	801	740	574	645			

Üretilen bu verilere ilişkin bazı temel istatistikler şöyle hesaplanmıştır;

Tablo4.3

Üretilen sayıların ortalama ve varyansları

	İl	**İlçe**	**Köy**	
Karadeniz	$\mu_{11} = 496.60$ $\pi_{11} = 510.25$ $N_{11} = 20$ $P_{11} = 0.066$	$\mu_{12} = 691.50$ $\pi_{12} = 726.05$ $N_{12} = 20$ $P_{12} = 0.066$	$\mu_{13} = 399.84$ $\pi_{13} = 174.91$ $N_{13} = 50$ $P_{13} = 0.166$	$\boldsymbol{\mu_{1.} = 486.16}$ $\boldsymbol{\pi_{1.} = 14045.93}$ $\boldsymbol{N_{1.} = 90}$ $\boldsymbol{P_{1.} = 0.300}$ $\boldsymbol{n_{1.} \cong 5}$
Marmara	$\mu_{21} = 996.17$ $\pi_{21} = 1115.31$ $N_{21} = 30$ $P_{21} = 0.100$	$\mu_{22} = 796.00$ $\pi_{22} = 457.55$ $N_{22} = 10$ $P_{22} = 0.033$	$\mu_{23} = 505.30$ $\pi_{23} = 356.85$ $N_{23} = 20$ $P_{23} = 0.066$	$\boldsymbol{\mu_{2.} = 799.18}$ $\boldsymbol{\pi_{2.} = 49741.77}$ $\boldsymbol{N_{2.} = 60}$ $\boldsymbol{P_{2.} = 0.200}$ $\boldsymbol{n_{2.} \cong 3}$
Akdeniz	$\mu_{31} = 697{,}44$ $\pi_{31} = 516{,}44$ $N_{31} = 30$ $P_{31} = 0.100$	$\mu_{32} = 798.20$ $\pi_{32} = 537.01$ $N_{32} = 20$ $P_{32} = 0.066$	$\mu_{33} = 895.87$ $\pi_{33} = 703.29$ $N_{33} = 30$ $P_{33} = 0.100$	$\boldsymbol{\mu_{3.} = 797{,}04}$ $\boldsymbol{\pi_{3.} = 8052{,}91}$ $\boldsymbol{N_{3.} = 80}$ $\boldsymbol{P_{3.} = 0.267}$ $\boldsymbol{n_{3.} \cong 4}$
Ege	$\mu_{41} = 789.75$ $\pi_{41} = 643.88$ $N_{41} = 20$ $P_{41} = 0.066$	$\mu_{42} = 798.40$ $\pi_{42} = 928.45$ $N_{42} = 30$ $P_{42} = 0.100$	$\mu_{43} = 606.30$ $\pi_{43} = 425.48$ $N_{43} = 20$ $P_{43} = 0.066$	$\boldsymbol{\mu_{4.} = 741.04}$ $\boldsymbol{\pi_{4.} = 8065.20}$ $\boldsymbol{N_{4.} = 70}$ $\boldsymbol{P_{4.} = 0.233}$ $\boldsymbol{n_{4.} \cong 3}$
	$\boldsymbol{\mu_{.1} = 765{,}35}$ $\boldsymbol{\pi_{.1} = 32952{,}64}$ $\boldsymbol{N_{.1} = 100}$ $\boldsymbol{P_{.1} = 0.333}$ $\boldsymbol{n_{.1} \cong 5}$	$\boldsymbol{\mu_{.2} = 771.33}$ $\boldsymbol{\pi_{.2} = 2848.19}$ $\boldsymbol{N_{.2} = 80}$ $\boldsymbol{P_{.2} = 0.266}$ $\boldsymbol{n_{.2} \cong 4}$	$\boldsymbol{\mu_{.3} = 575.83}$ $\boldsymbol{\pi_{.3} = 40195.09}$ $\boldsymbol{N_{.3} = 120}$ $\boldsymbol{P_{.3} = 0.400}$ $\boldsymbol{n_{.3} \cong 6}$	$\boldsymbol{\mu_{..} = 691{,}13}$ $\boldsymbol{\pi_{..} = 36559{,}01}$ $\boldsymbol{N_{..} = 300}$ $\boldsymbol{P_{..} = 1.00}$ $\boldsymbol{n_{..} = 15}$

4.1. Marjinal Olasılıklara Göre Cochran Orantılı Tabakalı Paylaştırma Yöntemi İle Yığın Hacminin Tahmini

Adım I: Cochran' ın önerdiği metotla $n = 15$ alınıp 15×15' lik çapraz tablo çizilsin.

Adım II: Kolonlar ve satırlar tabakalardan alınacak örnek hacimlerine göre tabaka sayıları kadar bölünürler. Bu durumda kolonlar 3 tabakaya ayrılacak (il, ilçe, köy) ve sırasıyla İlde $n_{ij} = n.P_{ij} = 15 * 0.3333 \cong 5$ kolon, İlçede 4 kolon ve Köyde 6 kolon olmak üzere toplam 15 birimlik örnek alınmış olacaktır.

Diğer değişkene göre tabaklama yapılırsa satırlar 4 tabakaya ayrılacak (Karadeniz, Marmara, Akdeniz, Ege) ve sırasıyla Karadeniz'de 5 sütun, Marmara'da 3 sütun, Akdeniz'de 3 sütun ve Ege'de 4 sütun olmak üzere toplam 15 birimlik örnek alınmış olacaktır.

Adım III: Bu işlem satır veya kolon işlemlerinden bir tanesi tercih edilerek yapılır. Bu örnekte satır işlemleri tercih edilmiştir. Bu durumda 1. satırdan başlamak üzere 1 ile 15 arasında tesadüfi bir kolon belirlenir, bu kolon 1. kolon olsun. 1. satır ve 1. kolon işaretlenir. Daha sonra 2. satıra geçilir ve 1. kolon dışında tesadüfi bir kolon belirlenir. bu kolon 10. kolon olsun. 2. satır ve 10. kolon işaretlenir ve bu işlem her satırda tesadüfi bir kolon olmak üzere tamamlanır.

Tablo4.4

n = 15 için Örnek Çekimi

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		İl					İlçe				Köy					
1	Karadeniz	×														
2											×					
3					×											
4													×			
5														×		
6	Marmara		×													
7												×				
8								×								
9	Akdeniz									×						
10							×									
11				×												
12	Ege														×	
13									×							
14																×
15						×										

Tablo4.5

Cochran yöntemiyle çekilen örneklerin tabakalara paylaştırılması

	İl	İlçe	Köy	
Karadeniz	2	0	3	**5**
Marmara	1	1	1	**3**
Akdeniz	1	2	0	**3**
Ege	1	1	2	**4**
	5	**4**	**6**	**15**

Karadeniz il merkezinden 2 kişi ve köylerinden 3 kişi; Marmara bölgesi yerleşim yerlerinden 1' er; Akdeniz il merkezlerinden 1, ilçelerinden 2; Ege bölgesi illerinden ve ilçelerinden 1, köylerinden 2 kişi örneğe alındığında bölgesel ya da yerleşim yerine göre marjinal paylaştırma yöntemi sağlanmış olur.

Bu hücrelerden belirlenen sayıdaki birimleri tesadüfi olarak seçilsin. Bu şekildeki hazırlamış olan Tablo4.6 aşağıda verilmiştir.

Tablo.4.6
Cochran yöntemi ile hücrelerden seçilen gözlemler

	İl	İlçe	Köy
Karadeniz	472 525		401 399 404
Marmara	981	819	500
Akdeniz	687	802 791	
Ege	828	869	597 597

Bu gözlemler ile yığın hacmini tahmin etmek için aşağıdaki formül kullanılacaktır;

$$\bar{y}_{tb} = \frac{N_1\bar{y}_1 + \cdots + N_T\bar{y}_T}{N_1 + \cdots + N_T}$$

Buna göre yığın ortalamasının tahmini;

$$\bar{y}_{tb} = (20 * \frac{472+525}{2} + 50 * \frac{(401+399+404)}{3} + 30 * 981 + 10 * 819 + 20 * 500$$
$$+30 * 687 + 20 * \frac{802+791}{2} + 20 * 828 + 30 * 869 + 20 * \frac{597+597}{2})/250$$
$$= 756.18$$ YTL olarak hesaplanır.

4.2. Marjinal Olasılıklara Göre Tabakalı Paylaştırma Yönteminde Matematiksel Programlama Yaklaşımı İçin Yığın Hacminin Tahmini

Bölüm 3.2. deki açıklamalara göre Cochran' ın önerdiği metot için en uygun çözüm araştırıldığında amaç fonksiyonu aşağıdaki gibi olacaktır.

max = n11 + n12 + n13 + n21 + n22 + n23 + n31 + n32 + n33 + n41 + n42 + n43;

bölgelere göre alınması gereken örnek hacimleri için kısıtlar;

n11 + n12 + n13 = 5; *(Karadeniz bölgesinden alınacak örneğin hacmi=5)*

n21 + n22 + n23 = 3; *(Marmara bölgesinden alınacak örneğin hacmi=3)*

n31 + n32 + n33 = 3; *(Akdeniz bölgesinden alınacak örneğin hacmi=3)*

n41 + n42 + n43 = 4; *(Ege bölgesinden alınacak örneğin hacmi=4)*

Yerleşim yerlerine göre alınması gereken örnek hacimleri için kısıtlar;

n11 + n21 + n31 + n41 = 5; *(İllerden alınacak örneğin hacmi=5)*

n12 + n22 + n32 + n42 = 4; *(İlçelerden alınacak örneğin hacmi=4)*

n13 + n23 + n33 + n43 = 6; *(Köylerden alınacak örneğin hacmi=6)*

Örnek hacmi = $n = \sum_{t1=1}^{4} \sum_{t2=1}^{3} n_{t1t2} = 15$ kısıtı;

n11 + n12 + n13 + n21 + n22 + n23 + n31 + n32 + n33 + n41 + n42 + n43 =15;

Tamsayı olma kısıtları;

@gin(n11);

@gin(n12);

@gin(n13);

@gin(n21);

@gin(n22);

@gin(n23);

@gin(n31);

@gin(n32);

@gin(n33);

@gin(n41);

@gin(n42);

@gin(n43);

Olmak üzere matematiksel programlama yaklaşımının çözümü;

Rows=9 Vars=12 No. integer vars=12 (all are linear)

Nonzeros=56 Constraint nonz=36(36 are +- 1) Density=0.479

Smallest and largest elements in absolute value= 1.00000 15.0000

No. < : 0 No. =: 8 No. > : 0, Obj=MAX, GUBs <= 4

Single cols= 0

Optimal solution found at step: 3

Objective value: 15.00000

Variable	*Value*	*Reduced Cost*
N11	*0.0000000E+00*	*0.0000000E+00*
N12	*4.000000*	*0.0000000E+00*
N13	*1.000000*	*0.0000000E+00*
N21	*0.0000000E+00*	*0.0000000E+00*
N22	*0.0000000E+00*	*0.0000000E+00*
N23	*3.000000*	*0.0000000E+00*
N31	*1.000000*	*0.0000000E+00*
N32	*0.0000000E+00*	*0.0000000E+00*
N33	*2.000000*	*0.0000000E+00*
N41	*4.000000*	*0.0000000E+00*
N42	*0.0000000E+00*	*0.0000000E+00*
N43	*0.0000000E+00*	*0.0000000E+00*
Row	***Slack or Surplus***	***Dual Price***
1	*15.00000*	*1.000000*
2	*0.0000000E+00*	*1.000000*
3	*0.0000000E+00*	*1.000000*
4	*0.0000000E+00*	*1.000000*
5	*0.0000000E+00*	*1.000000*
6	*0.0000000E+00*	*0.0000000E+00*
7	*0.0000000E+00*	*0.0000000E+00*
8	*0.0000000E+00*	*0.0000000E+00*
9	*0.0000000E+00*	*0.0000000E+00*

Bu sonuca göre tabakalardan alınacak örnek hacimleri şu şekilde belirlenmiştir.

Tablo4.7

Matematiksel programlama yöntemi ile örnek hacminin tabakalara paylaştırılması

	İl	**İlçe**	**Köy**	
Karadeniz	0	4	1	**5**
Marmara	0	0	3	**3**
Akdeniz	1	0	2	**3**
Ege	4	0	0	**4**
	5	**4**	**6**	**15**

Tablo4.8' deki veriler Tablo4.2. deki veri tablosundan tesadüfi olarak seçilmiştir. Seçilen verilerle yığın ortalaması tahmin edilmiştir.

Tablo4.8

Matematiksel programlama yöntemi ile hücrelerden seçilen gözlemler

	İl	İlçe	Köy
Karadeniz		703 677 711 676	394
Marmara			498 546 491
Akdeniz	722		883 922
Ege	777 789 817 783		

$$\bar{y}_{tb} = \frac{\bar{Y}_{tb}}{N} = \frac{N_1\bar{y}_1 + \cdots + N_T\bar{y}_T}{N_1 + \cdots + N_T}$$

Formülüyle,

$$\bar{y}_{tb} = (20 * \frac{(703+677+711+676)}{4} + 50 * 394 + 20 * \frac{(498+546+491)}{3} + 30 * 722 + 30 * \frac{(883+922)}{2}$$

$$+20 * \frac{(777+789+817+783)}{4})/170$$

= 637.25 YTL şeklinde hesaplanır.

4.3. Ortak Olasılıklara Göre Önerilen Orantılı Tabakalı Paylaştırma Yöntemi İçin Yığın Hacminin Tahmin

Bölüm 3.3. te bahsedilen ve tabaka hacmi büyük olan alt tabakalardan daha çok birimi örneğe almak için geliştirilen yönteme göre örneğe alınacak 15 birimi örneğe alma işlemi şu şekilde yapılacaktır;

Adım I: Her çapraz tabakaya 1 birim başlangıç olarak atanır.

Adım II: Her çapraz tabaka için ilgili tabakalara ait yığın hacim oranları (π_{ij}) alt alta yazılır ve yanına bir kolon yazılarak bu kolona birikimli π_{ij} hesaplanarak yazılır.

Adım III: İlk $T_1 \cdot T_2$ birim hücrelere atandığından düzgün $0-1$ aralığından tesadüfi olarak $n - T_1 \cdot T_2$ kadar sayı belirlenir. Bu örnek için $n = 15$ olduğundan 3 tane tesadüfi sayı çekilir.

Adım IV: Seçilen bu sayılardan 1. si 0.33073, birikimli ortak olasılıkların içinde taranır ve hangi satır bu sayıyı kapsıyorsa o satıra 1 eklenir, bu örnek için 4. satır bu sayıyı kapsar. 2. sayı 0.89277 olduğundan 11. satıra, ve 3. sayı da 0.28592 olduğundan 3. satıra eklenir.

Adım V: Daha sonra hangi satırda kaç birim seçilmişse bunlar tabakalara dağıtılır. Örneğin; $Karadeniz \times İl$ çapraz tabakasından önceki adımda birim örneğe alınmadığından bu çapraz tabakanın hacmi 1 olarak kalacaktır. Ancak $Karadeniz \times Köy$ çapraz tabakasından 2 birim örneğe alınacaktır.

Tablo4.9

Alternatif yöntemde hücre içi olasılıklar ve birikimli olasılıkların hesaplanması

	Bölge	Yerleşim Birimi	Yığın Hacim Oranları π_{ij}	Birikimli Ortak Olasılıklar
1	Karadeniz	İl	20/300=0.0666	0.0666
2	Karadeniz	İlçe	20/300=0.0666	0.1332
3	Karadeniz	Köy	50/300=0.1666	0.2998
4	Marmara	İl	30/300=0.1000	0.3998
5	Marmara	İlçe	10/300=0.0333	0.4331
6	Marmara	Köy	20/300=0.0666	0.4997
7	Akdeniz	İl	30/300=0.1000	0.5997
8	Akdeniz	İlçe	20/300=0.0666	0.6663
9	Akdeniz	Köy	30/300=0.1000	0.7663
10	Ege	İl	20/300=0.0666	0.8329
11	Ege	İlçe	30/300=0.1000	0.9329
12	Ege	Köy	20/300=0.0666	1.0000

Tablo4.10

Tesadüfi sayılar ile birikimli olasılıkların karşılaştırılması

	Düzgün 0-1 aralığından tesadüfi sayılar	Birikimli olasılık değeri (karşılaştırılan değer)	Bölge	Yerleşim Birimi
1	0.33073	0.3998	Marmara	İl
2	0.89277	0.9329	Ege	İlçe
3	0.28592	0.2998	Karadeniz	Köy

Tablo4.11

Alternatif yöntem ile örnek hacminin tabakalara paylaştırılması

	İl	**İlçe**	**Köy**	
Karadeniz	1	1	2	**4**
Marmara	2	1	1	**4**
Akdeniz	1	1	1	**3**
Ege	1	2	1	**4**
	5	**5**	**5**	**15**

Verilen yöntemle çapraz hücrelerden kaç birim örneğe alınacağı belirlenmiş ve Tablo4.2. deki veri setinden rasgele çekilen veriler yardımıyla yığın ortalamasının tahmini hesaplanmıştır.

Tablo4.12

Alternatif yöntem ile hücrelerden seçilen gözlemler

	İl	**İlçe**	**Köy**
Karadeniz	554	675	399 388
Marmara	962 1004	794	498
Akdeniz	664	802	878
Ege	817	807 827	617

Yığın ortalamasının tahmini için kullanılacak formül;

$$\bar{y}_{tb} = \frac{\bar{Y}_{tb}}{N} = \frac{N_1\bar{y}_1 + \cdots + N_T\bar{y}_T}{N_1 + \cdots + N_T}$$

olacaktır.

Buna göre yığın ortalaması

$$\bar{y}_{tb} = (20 * 554 + 20 * 675 + 50 * \frac{(399 + 388)}{2} + 30 * \frac{(1004 + 962)}{2} + 10 * 794 + 20 * 498$$

$$+30 * 664 + 20 * 802 + 30 * 878 + 20 * 817 + 30 * \frac{807+827}{2} + 20 * 617)/300$$

$= 690.45$ YTL olarak tahmin edilmiştir.

5. SİMÜLASYON

Cochran' ın önerdiği metod ile alternatif olarak önerilen metod yığın ortalamasının tahmini açısından karşılaştırılmıştır. Daha sonra da bu tahminlere ilişkin standart hata değerleri kullanılarak bu iki yöntem arasında karşılaştırma yapılmıştır. Bununla amaçlanan değişik tabaka hacimleri için Cochran metodu ile alternatif olarak önerilen metodun yığın ortalamasının tahmin değerlerinin incelenmesidir ve sonuçlar Tablo5.2'de gösterilmiştir.

Tablo5.1

Tabaka hacimleri

	İl	**İlçe**	**Köy**	
Karadeniz	20	20	50	90
Marmara	30	10	20	60
Akdeniz	30	20	30	80
Ege	20	30	20	70
	100	80	120	300

Tablo5.2

Cochran metodu ile alternatif metodun karşılaştırılması

Deneme=1000	**Cochran Metodu**		**Alternatif Metod**	
n	Ortalama Tahmini	Standart Hata	Ortalama Tahmini	Standart Hata
15	673.18	34.65	690.87	6.65
30	687.03	13.59	691.09	4.73
45	690.25	5.95	690.97	4.02
60	690.76	4.79	691.07	3.37
75	691.11	3.21	691.24	2.87
90	691.08	2.85	691.18	2.68
100	691.10	2.85	691.13	2.56

Burada dikkat edilmesi gereken, örneğe alınacak birimlerin sayısı arttıkça Cochran ve Alternatif metodlar için standart hatanın birbirine yaklaşmasıdır. Ancak örneğe alınacak birim sayısı az olduğunda alternatif yöntem kullanılırsa yığın ortalamasını tahmin etmek için daha iyi bir seçim yapılmış olacaktır.

5.1. Cochran Metodu için SPSS Programı

```
set mxloop=10000.
matrix.
compute densay=1000.
compute otah=make(densay,1,0).
compute KiT=make(densay,1,0).
compute n=100.
compute A=ident(n).
compute tbkhacmi={20,20,50;30,10,20;30,20,30;20,30,20}.          /* veriler giriliyor */
compute altsnr={1,21,41,91,121,131,151,181,201,231,251,281}.
get X/file*/variables=gelir.                                     /* veriler giriliyor */
compute th=reshape(tbkhacmi,1,12).
compute yort=msum(x)/300.
compute strsay=ncol(tbkhacmi).
compute stnsay=nrow(tbkhacmi).
compute strstn=strsay*stnsay.
compute WT=make(stnsay,strsay,0).
compute w=make(stnsay,strsay,1).
compute B=tbkhacmi.
compute tort=0.
compute tsapma=0.
compute shata=0.
loop trt=1 to densay.
/* tabakalardan seçilecek orneklem hacmi hesaplanıyor */
compute ornhacmi=tbkhacmi/msum(tbkhacmi).
compute str=make(1,strsay,1).
compute stn=make(stnsay,1,1).
/* sutunlardan seçilecek orneklem hacmi */
loop sutun=1 to strsay.
 compute ysut=rnd(n*(msum(ornhacmi(:,sutun)))).
 compute g=str(:,sutun).
```

```
 compute str(:,sutun)=ysut.
 compute ysut=g.
/* satırlardan seçilecek orneklem hacmi */
loop satir=1 to stnsay.
 compute ysat=rnd(n*(msum(ornhacmi(satir,:)))).
 compute f=stn(sutun,:).
 compute stn(satir,:)=ysat.
 compute ysat=f.
end loop.
end loop.
compute sum1=msum(stn).
compute sum2=msum(stn).
/* tabakardan alınan örneklem yuvarlama sonucunda istenene eşit değilse */
 compute yty=0.
 compute stnmax=cmax(stn).
loop i=1 to strsay.
do if (stn(i)=stnmax).
 compute  yty=i.
end if.
end loop.
compute kln=0.
compute strmax=rmax(str).
loop i=1 to strsay.
 do if (str(i)=strmax).
compute  kln=i.
end if.
end loop.
do if (csum(stn)<n).          /* satırdaki fazlaların veya eksiklerin kontrolü */
compute stn(kln)=stn(kln)+1.
else if (csum(stn)>n).
compute stn(kln)=stn(kln)-1.
```

```
end if.
/* sütundaki fazlaların veya eksiklerin kontrolü */
do if (rsum(str)<n).
compute str(yty)=str(yty)+1.
else if (rsum(str)>n).
compute str(yty)=str(yty)-1.
end if.
/* kolondaki yığılmalı orneklem hacmi */
compute kolon=make(stnsay,1,stn(1)).
loop i=1 to stnsay.
loop j=i+1 to stnsay.
compute kolon(j)=kolon(i)+stn(j).
end loop.
end loop.
/* satırdaki yığılmalı orneklem hacmi */
compute yatay=make(1,strsay,str(1)).
loop i=1 to strsay.
loop j=i+1 to strsay.
compute yatay(j)=yatay(i)+str(j).
end loop.
end loop.
/* A matrisinin sutunları karıştırıldı */
loop sut=1 to n.
compute s1=trunc(n*uniform(1,1)+1).
compute s2=trunc(n*uniform(1,1)+1).
compute dvek=A(:,s1).
compute A(:,s1)=A(:,s2).
compute A(:,s2)=dvek.
compute d1={s1,s2}.
end loop.
/* A matrisinin satırları değiştirildi */
```

```
loop sat=1 to n.
compute s1=trunc(n*uniform(1,1)+1).
compute s2=trunc(n*uniform(1,1)+1).
compute avek=A(s1,:).
compute A(s1,:)=A(s2,:).
compute A(s2,:)=avek.
compute d2={s1,s2}.
end loop.
/* toplayacağımız satır ve sütun sayılarının alt sınırı */
compute q=make(1,strsay,1).
compute r=make(stnsay,1,1).
loop i=1 to (stnsay-1).
compute r(i+1)=r(i)+stn(i).
end loop.
loop j=1 to (strsay-1).
compute q(j+1)=q(j)+str(j).
end loop.
/* A matrisinin kolon(i) ile yatay(j) arasında kalan kısmının toplamı */
compute W= make(stnsay,strsay,0).
loop i=1 to stnsay.
loop j=1 to strsay.
loop m=q(j) to yatay(j).
loop l=r(i) to kolon(i).
compute W(i,j)=W(i,j)+A(l,m).
end loop.
end loop.
end loop.
end loop.
compute Ki=make(stnsay,strsay,0).
compute KiT(trt,1)=msum((B-W)&**2/B).
Compute Wt=Wt+W.
```

```
compute indis=0.
compute Gtop=0.
/* Ortalama tahmini başlıyor */
compute tx=1.
compute tth=0.
loop i=1 to 4.
loop j=1 to 3.
compute nt=w(i,j).
do if (nt>0).
compute ht=0.
loop k=1 to nt.
compute XX=0.
compute indis=trunc(uniform(1,1)*th(tx))+altsnr(tx).
compute XX=X(indis).
compute ht=ht+XX.
end loop.
compute hkat=th(tx)*ht/nt.
compute gtop=gtop+hkat.
compute tth=tth+th(tx).
end if.
compute tx=tx+1.
end loop.
end loop.
/* Ortalama tahmini bitti */
compute Ortah=Gtop/tth.
compute otah(trt)=ortah.
compute tort=tort+ortah.
compute tsapma=tsapma+(ortah-691.13)**2.
end loop.
compute dort=tort/densay.
compute shata=(tsapma/(densay-1))**0.5.
```

```
print dort.
print shata.
save otah /outfile=*.
end matrix.
```
(Semiz, 2007)

5.2. Alternatif Metod İçin SPSS Programı

Yapılan bazı simülasyonlarla Cochran metodunun bazı eksiklikleri fark edilmiştir. Tabakaların hacimleri ile hücre hücre ilgilenmek yapılacak sonuç çıkarımlarında daha doğru bilgi sağlayacaktır. Program temel olarak; her hücreden en az 1 birimi örneğe alarak rasgele $0-1$ aralığından sayılar çekmek ve bu sayıları yığılımlı ortak olasılıklarla karşılaştırma işlemi yapmaktadır. Seçilen rasgele sayı hangi birikimli olasılığın içinde kalıyorsa ona karşılık gelen hücreden 1 birim daha örneğe alınacaktır. Böylece hangi hücrenin $\frac{n_{ij}}{N}$ oranı daha büyükse o hücreden daha çok örnek alınacaktır.

```
set mxloop=10000.
matrix.
compute densay=1.
compute n=15.
compute otah=make(densay,1,0).
compute tort=0.
compute tsapma=0.
compute shata=0.
compute tbkhacmi={20,20,50;30,10,20;30,20,30;20,30,20}.          /* veriler giriliyor */
compute altsnr ={1,21,41,91,121,131,151,181,201,231,251,281}.
get X/file*/variables=gelir.                                     /* veriler giriliyor */
compute th=reshape(tbkhacmi,1,12).
compute strsay=ncol(tbkhacmi).
compute stnsay=nrow(tbkhacmi).
```

```
compute strstn=strsay*stnsay.
compute A=make(stnsay,strsay,1).
compute AT=make(stnsay,strsay,0).
compute thoran=tbkhacmi/(msum(tbkhacmi)).                    /*  tabakalar oranlanıyor */
compute dikey=reshape(thoran,strstn,1).      /* tabakalar oranlara göre sıralandı */
compute KiT=make(densay,1,0).
loop r=1 to densay.
do if n>(strstn).
compute dky=dikey.
compute random=uniform(n-strstn,1).                          /* rasgele sayılar üretildi  */
loop i=2 to strstn.                          /* oranlar yığılımlı olarak toplandı hesaplandı  */
compute dky(i,1)=dikey(i,1)+dky(i-1,1).
end loop.
/* üretilen rasgele sayılara göre hangi tabakadan kaç birim çekileceği hesaplanıyor */
compute B=make(strstn,1,1).
loop i=1 to (n-strstn).
loop j=1 to strstn.
do if (random(i,1)<dky(j,1)).
compute B(j,1)=B(j,1)+1.
end if.
end loop.
end loop.
compute C=make(strstn,1,1).
compute C(1,1)=B(1,1).
loop i=2 to strstn.
compute C(i,1)=B(i,1)-B(i-1,1)+1.
end loop.
compute A=reshape(C,stnsay,strsay).
compute Ki=make(stnsay,strsay,0).
loop i=1 to stnsay.
loop j=1 to strsay.
```

```
compute Ki(i,j)=((A(i,j)-tbkhacmi(i,j))*(A(i,j)-tbkhacmi(i,j)))/(tbkhacmi(i,j)).
end loop.
end loop.
compute KiT(r,1)=msum(Ki).
compute topl=msum(a).
compute at=at+a.
else if (n=12).
print a.
end if.
compute indis=0.
compute Gtop=0.
/* Ortalama tahmini başlıyor */
compute tx=1.
compute tth=0.
loop i=1 to 4.
loop j=1 to 3.
compute nt=a(i,j).
do if (nt>0).
compute ht=0.
loop k=1 to nt.
compute XX=0.
compute indis=trunc(uniform(1,1)*th(tx))+altsnr(tx).
compute XX=X(indis).
compute ht=ht+XX.
end loop.
compute tth=tth+th(tx).
compute hkat=th(tx)*ht/nt.
compute gtop=gtop+hkat.
end if.
compute tx=tx+1.
end loop.
```

```
end loop.
/* Ortalama tahmini bitti */
compute Ortah=Gtop/tth.
compute otah(r)=ortah.
compute tort=tort+ortah.
compute tsapma=tsapma+(ortah-691.13)**2.
end loop.
compute dort=tort/densay.
compute shata=(tsapma/(densay-1))**0.5.
print dort.
print shata.
save otah /outfile=*.
end matrix.
```
(Semiz, 2007)

6. SONUÇ

Bu tezde iki yönüyle ilgilenilen tabakalı bir veri seti için daha az örnek alınarak yığın hakkında bilgi sağlanması amaçlanmıştır. Bunun için Cochran' ın önerdiği metot ile bu metottan esinlenilerek geliştirilen bir metot karşılaştırılmıştır.

Yapılan simülasyon çalışması ile daha küçük örneklemlerde alternatif olarak sunulan yöntemin yığın ortalamasına ilişkin daha doğru sonuç verdiği gözlenmiştir. Bir birimi örneğe almak için belirli bir maliyet gerektiğinden tabakalardan az sayıda örnek alarak yığın hakkında daha iyi tahmin yapılması örnekleme maliyeti açısından da önemlidir.

Verilen 3 metod ile örneğe alınacak 15 birim, alt tabakalara nasıl dağıtılacağı hesaplanmış ve alt tabakalardan rasgele gözlem alınmış, bu verilerle ortalama tahmini yapılmıştır. Genel ortalama 691.13 YTL olduğu Tablo4.3'te gösterilmiştir.

Cochran' ın önerdiği metodla alt tabakalardan seçilen bu 15 birim ile yığın ortalamasının tahmini 756.18 YTL şeklinde hesaplanmıştır.

Matematiksel programlama yaklaşımıyla seçilen 15 birim için yığın ortalamasının tahmini 637.25 YTL şeklinde hesaplanmıştır.

Alternatif yöntemle alt tabakalara dağıtılan ve seçilen 15 birim yardımıyla hesaplanan yığın ortalamasının tahmini 690.45 olarak hesaplanmıştır.

Buna göre tahminde alternatif yöntemin değerlendirilen yığın ortalamasını tahmin etmesi açısından daha iyi olduğu söylenebilir. Bunun en önemli nedeni her hücreden örneğe en az 1 birimin alınıyor olmasıdır. Bu şekilde yığın ortalaması hem daha yansız hem de daha az örnek hacmi ile daha doğru tahmin yapılmaktadır.

KAYNAKLAR

1. Cochran, W.G. 1977. Sampling Techniques, 3^{nd} Edition, John Wiley and Sons Inc., New York.

2. Çıngı, H. 1994. Örnekleme Kuramı 2, Hacettepe Üniversitesi Fen Fakültesi Yayınevi, Ankara.

3. Deming, William Edvards (1950), Some Theory of Sampling, Dower Publications Inc., New York.

4. Ersin, Alptekin ve diğerleri (2001), Temel Örnekleme Yöntemleri, Literatür Yayınları, İstanbul.

5. Semiz, M. , Esin, A. A. 2000. Tabaka Örnek Hacimlerinin Doğrusal Olmayan Tamsayılı Programlama İle Belirlenmesi. İstatistik Sempozyumu Syf:199 Ankara, 27-28 Nisan.

6. Semiz, M. 2007. Örnekleme Yöntemleri, Dizgi Ofset ve Matbaacılık, Konya

7. Semiz, M. 2007. İstatistik Laboratuarı ve SPSS-MP Dizgi Ofset ve Matbaacılık, Konya.

8. Cochran, W. G. , 1942. Sampling Theory When The Sampling Units are of Unequal Sizes, J. Amer Statistics Assocation 37, Page: 199-212.

9. Cochran, W. G. , 1946. Relative Accuracy of Systematic and Stratified Random Samples for a Certain Class of Populations, Ann. Math. Statist. 17, Page:164-177

10. Yamane T., (2001). Çevirenler: Alptekin E, Temel Örnekleme Teknikleri, Literatür Yayınları, İstanbul.

11. Sampford M. R., 1962 An Introduction to Sampling Theory, The Statistician, Vol. 13, No. 1, pp. 81-82