

**ÖRNEKLEME KURAMINDA AĞIRLIKLANDIRMA**

**Aylin ALKAYA**

**DOKTORA TEZİ  
İSTATİSTİK**

**GAZİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**HAZİRAN 2009  
ANKARA**

Aylin ALKAYA tarafından hazırlanan ÖRNEKLEME KURAMINDA AĞIRLIKLANDIRMA adlı bu tezin Doktora tezi olarak uygun olduğunu onaylarım.

Prof. Dr. Alptekin ESİN .....  
Tez Danışmanı, İstatistik Anabilim Dalı

Bu çalışma, jürimiz tarafından oy birliği ile İstatistik Anabilim Dalında Doktora tezi olarak kabul edilmiştir.

Prof. Dr. Alptekin ESİN .....  
İstatistik, Gazi Üniversitesi  
Prof. Dr. Ömer L. GEBİZLİOĞLU .....  
İstatistik, Ankara Üniversitesi  
Prof. Dr. Müslim EKNİ .....  
İstatistik, Gazi Üniversitesi  
Prof. Dr. Öztaş AYHAN .....  
İstatistik, ODTÜ  
Yrd. Doç. Dr. Celal AYDIN .....  
İstatistik, Gazi Üniversitesi

Tarih: 26/06/2009

Bu tez ile G.Ü. Fen Bilimleri Enstitüsü Yönetim Kurulu Doktora derecesini onamıştır.

Prof. Dr. Nail ÜNSAL .....  
Fen Bilimleri Enstitüsü Müdürü

## **TEZ BİLDİRİMİ**

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

Aylin ALKAYA

# ÖRNEKLEME KURAMINDA AĞIRLIKLANDIRMA

(Doktora Tezi)

Aylin ALKAYA

GAZİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

Haziran 2009

## ÖZET

Örnekleme araştırmalarında tahminlerin kalitesini arttırmak için veriye ağırlıklandırma uygulanır. Yığından örnek birimlerinin seçim olasılıklarındaki farklılıkları denkleştirmek, yanıtlanmama, kapsanmama sorununu gidermek ve örnek dağılımının yığın dağılımına uyumunu sağlamak amacıyla veride ağırlıklandırma çok aşamalı bir süreç olarak yapılır. Bu çalışmada, olasılı örnekleme yöntemlerinden basit tesadüfi örnekleme, tabakalı tesadüfi örnekleme ve küme örnekleme tasarımlarında parametrelerin Horvitz-Thompson [1952] ve Hájek [1971] tahmin edicileri incelenmiş ve tahmin edicilerin ağırlıklandırma aşamalarındaki yapıları verilmiştir. Ağırlıklandırma yöntemlerinden ağırlıklandırma sınıf düzeltmesi, sonradan tabakalama, tarama, ayarlama tahmin edicisi, genelleştirilmiş regresyon (GREG) tahmin edicisi yöntemleri tanıtılmıştır. Ağırlıklandırma yöntemlerinin Hájek tahmin edicileri tanımlamaları verilmiştir. Aşamalı ayarlama tahmin edicisi yöntemi önerilmiştir. 2003 Türkiye Nüfus ve Sağlık Araştırması verisinde örnekleme, yanıtlanmama ve ayarlama ağırlıklandırması, aşamalı ayarlama ve Ayhan (2003)'ün bileşik oran tahmin edicisiyle elde edilen parametre tahminlerinin duyarlılıkları araştırılmıştır.

**Bilim Kodu : 205.1.066**

**Anahtar Kelimeler : Ağırlıklandırma, örnekleme ağırlıklandırması, yanıtlanmama ağırlıklandırması, ayarlama ağırlıklandırması, Horvitz-Thompson tahmin edicisi, Hájek tahmin edicisi, ağırlıklandırma sınıf düzeltmesi, sonradan tabakalama, tarama, ayarlama tahmin edicisi, genelleştirilmiş regresyon (GREG) tahmin edicisi, aşamalı ayarlama, bileşik oran tahmin edicisi**

**Sayfa Adedi : 186**

**Tez Yöneticisi : Prof. Dr. Alptekin ESİN**

**WEIGHTING IN SAMPLING THEORY****(Ph.D. Thesis)****Aylin ALKAYA****GAZİ UNIVERSITY****INSTITUTE OF SCIENCE AND TECHNOLOGY****Haziran 2009****ABSTRACT**

**In sampling surveys weighting is applied to data to increase the quality of estimates. Weighting in data is realized as a multi stage process to reduce the differences of the selection probabilities of sampling units, to avoid nonresponse, noncoverage bias and so as to make the sampling distribution match with population distribution. In this study the Horvitz-Thompson [1952] and Hájek [1971] estimators of the population parameters are examined in the probability sampling methods of simple random sampling, stratified random sampling and cluster sampling and the estimators structures given at weighting stages. The weighting methods of weighting class adjustment, poststratification, raking, calibration estimator and generalized regression (GREG) estimators are illustrated. Hájek [1971] estimator definitions of the weighting methods are presented. The staged calibration estimator is suggested. Demographic and Health Survey/Turkey 2003 data are taken for application purposes. For the survey data the precision of the parameter estimates are investigated with respect to the sampling, nonresponse, calibration weighting, stage calibration and Ayhan (2003) combined ratio estimator.**

**Science Code : 205.1.066**

**Key Words : Weighting, sampling weighting, nonresponse weighting, calibration weighting, Horvitz-Thompson estimator, Hájek estimator, weighting class adjustment, poststratification, raking, calibration estimator, generalized regression (GREG), stage calibration, combined ratio estimator**

**Page Number: 186**

**Adviser : Prof. Dr. Alptekin ESİN**

## TEŞEKKÜR

Kıymetli danışman Hocam Prof. Dr. Alptekin ESİN'e çalışmalarım boyunca değerli yardım ve katkılarıyla beni yönlendirdiği, tezimin daha iyi olmasındaki çalışmalarımız için ve başarıma güvenip beni desteklediği için,

Değerli Hocam Prof. Dr. Öztaş AYHAN'a vaktini çok aldığım, tez danışmanım titizliğinde zaman ayırdığı, tezimin temellerinin atılmasında katkı sağlayarak bilgi birikimiyle tezimin her aşamasında beni yönlendirdiği için,

Sevgili Hocam Prof. Dr. Müslim EKNİ'ye çalışmalarına getirmiş olduğu bakış açısı ve yönlendirmelerle tezimin şekillenip titizlikle son haline getirilmesine katkı sağladığı için çok teşekkür ederim.

Değerli hocam Yrd. Doç. Dr. Celal Aydın'a tezime katkılarından dolayı teşekkür ederim.

Alanında uzman TUIK danışmanı Prof. Dr. Vijay Verma'ya sağladığı kaynaklar için, tezimin başlangıcında danıştığım TUIK Uzmanı Yılmaz Erşahin'e, tezimin uygulama aşmasında danıştığım ve yardımlarını gördüğüm Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü hocalarından Yrd.Doç.Dr. A.Sinan Türkyılmaz'a ve Arş.Gör. Mehmet Ali Eryurt'a teşekkür ederim. Gazi Üniversitesi Fen Bilimleri Enstitüsünden Munise Baş'a iyimserliği ve yardımları için teşekkür ederim.

Değerli akademisyen dostlarım Dr. Ahmet Tanç'a, Okt. Ayşegül Soylu, Yrd.Doç.Dr. Duygu Eren, Arş.Gör. Filiz Erdem Uçar, Yrd.Doç.Dr. Suzan Çoban ve Arş.Gör. Şükran Güngör Tanç'a hayatımda oldukları ve tezimde bana desteklerinden dolayı teşekkür ederim.

Tezimin bitmesini dört gözle bekleyen aileme sevgilerimle teşekkür ederim.

## İÇİNDEKİLER

|                                                                                      | Sayfa |
|--------------------------------------------------------------------------------------|-------|
| ÖZET .....                                                                           | iv    |
| ABSTRACT .....                                                                       | vi    |
| TEŞEKKÜR.....                                                                        | viii  |
| İÇİNDEKİLER .....                                                                    | ix    |
| ÇİZELGELERİN LİSTESİ.....                                                            | xiii  |
| SİMGELER VE KISALTMALAR.....                                                         | xv    |
| 1. GİRİŞ .....                                                                       | 1     |
| 2. ÖRNEKLEME AĞIRLIKLANDIRMASI.....                                                  | 11    |
| 2.1. Eşit Olasılı Örneklem .....                                                     | 15    |
| 2.2. Eşit Olmayan Olasılı Örneklem .....                                             | 15    |
| 2.3. Horvitz-Thompson Tahmin Edicisi .....                                           | 16    |
| 2.4. Hájek Tahmin Edicisi.....                                                       | 20    |
| 2.5. Basit Tesadüfi Örneklem.....                                                    | 24    |
| 2.5.1. Yığın ortalaması ve yığın toplamının Horvitz-Thompson<br>tahmin edicisi ..... | 25    |
| 2.5.2. Yığın ortalaması ve yığın toplamının Hájek tahmin edicisi .....               | 26    |
| 2.6. Tabakalı Tesadüfi Örneklem .....                                                | 27    |
| 2.6.1. Yığın ortalaması ve yığın toplamının Horvitz-Thompson<br>tahmin edicisi ..... | 29    |
| 2.6.2. Yığın ortalaması ve yığın toplamının Hájek tahmin edicisi .....               | 30    |
| 2.7. İki Aşamalı Küme Örneklemesi.....                                               | 31    |
| 2.7.1. Yığın ortalaması ve yığın toplamının Horvitz-Thompson<br>tahmin edicisi ..... | 33    |

**Sayfa**

|                                                                           |    |
|---------------------------------------------------------------------------|----|
| 2.7.2. Yığın ortalaması ve yığın toplamının Hájek tahmin edicisi .....    | 34 |
| 2.8. İki Değişkenin Birbirine Oranı R'nin Tahmini .....                   | 35 |
| 2.8.1. Basit tesadüfi örneklemede oran tahmini .....                      | 36 |
| 2.8.2. Tabakalı tesadüfi örneklemede oran tahmini .....                   | 37 |
| 2.8.3. İki aşamalı küme örneklemesinde oran tahmini.....                  | 38 |
| 2.8.4. Ayhan [2003]'ün önerdiği bileşik oran tahmin edicisi .....         | 39 |
| 2.9. Yığın Ortalaması Tahmin Edicileri için Genel Varyans Tahmini .....   | 41 |
| 2.10. Yığın Oranı Tahmin Edicileri için Genel Varyans Tahmini .....       | 44 |
| 3. YANITLANMAMA AĞIRLIKLANDIRMASI.....                                    | 47 |
| 3.1. Ağırlıklandırma Sınıflarının Belirlenmesi .....                      | 49 |
| 3.1.1. Yanıt eğilimi tabakalanması.....                                   | 52 |
| 3.1.2. Kestirim ortalama tabakalanması.....                               | 52 |
| 3.2. Ağırlıklandırma Sınıf Düzeltmesi .....                               | 53 |
| 3.3. Basit Tesadüfi Örneklemede Yanıtlanmama Ağırlıklandırması .....      | 55 |
| 3.3.1. Yanıtlayıcı ortalaması .....                                       | 56 |
| 3.3.2. Ağırlıklandırma sınıf düzeltmesi.....                              | 57 |
| 3.4. Tabakalı Tesadüfi Örneklemede Yanıtlanmama Ağırlıklandırması .....   | 59 |
| 3.4.1. Tabakalarda yanıtlanmama düzeltmesi .....                          | 59 |
| 3.4.2. Ağırlıklandırma sınıf düzeltmesi.....                              | 61 |
| 3.5. İki Aşamalı Küme Örneklemesinde Yanıtlanmama Ağırlıklandırması ..... | 62 |
| 4. AYARLAMA AĞIRLIKLANDIRMASI .....                                       | 64 |
| 4.1. Sonradan Tabakalama .....                                            | 69 |
| 4.1.1. Basit tesadüfi örneklemede sonradan tabakalama.....                | 73 |

**Sayfa**

|                                                                                  |     |
|----------------------------------------------------------------------------------|-----|
| 4.1.2. Tabakalı tesadüfi örneklemede sonradan tabakalama.....                    | 75  |
| 4.1.3. BTÖ de $\bar{y}_{st}$ tahmin edicisinin yanlılığı.....                    | 77  |
| 4.1.4. BTÖ de $\bar{y}_{st}$ sonradan tabakalama tahmin edicisinin varyansı..... | 79  |
| 4.2. Tarama .....                                                                | 82  |
| 4.2.1. Koşullu yan .....                                                         | 88  |
| 4.2.2. Koşullu varyans.....                                                      | 89  |
| 4.3. Ayarlama Tahmin Edicisi .....                                               | 90  |
| 4.3.1. Ayarlama ağırlıklarının belirlenmesi .....                                | 94  |
| 4.3.2. Yığın toplamının oran tahmin edicisinin ayarlamayla belirlenmesi .....    | 98  |
| 4.3.3. Genelleştirilmiş regresyon tahmin edicisi.....                            | 99  |
| 4.3.4. Regresyon tahmin edicisinin ayarlamayla belirlenmesi.....                 | 103 |
| 4.3.5. Yanlılık.....                                                             | 104 |
| 4.3.6. $\hat{Y}_C$ 'nin varyans ve varyans tahmini .....                         | 105 |
| 4.3.7. Neden ayarlanmış ağırlıklar? .....                                        | 107 |
| 4.5. Ayarlama Tahmin Edicisinde $\bar{Y}$ ve $Y$ 'nin Hájek Tahmin Edicisi.....  | 108 |
| 4.6. Aşamalı Ayarlama.....                                                       | 110 |
| 4.7. Ağırlıklandırma Yöntemlerini Kıyaslamada Varyans Şişme Faktörü .....        | 114 |
| 5. UYGULAMA .....                                                                | 117 |
| 5.1. Örnekleme Ağırlıklandırması.....                                            | 120 |
| 5.1.1. Yanıtlayıcı ortalaması .....                                              | 124 |
| 5.1.2. Yığın ortalamasının Horvitz-Thompson tahmin edicisi.....                  | 125 |
| 5.1.3. Yığın ortalamasının Hájek tahmin edicisi.....                             | 126 |

**Sayfa**

|                                                       |     |
|-------------------------------------------------------|-----|
| 5.2. Yanıtlanmama Ağırlıklandırması.....              | 128 |
| 5.2.1. Ağırlıklandırma sınıf düzeltmesi.....          | 130 |
| 5.3. Ayarlama Ağırlıklandırması.....                  | 133 |
| 5.3.1. Sonradan tabakalama .....                      | 135 |
| 5.3.2. Tarama .....                                   | 139 |
| 5.3.3. Ayarlama tahmin edicisi .....                  | 149 |
| 5.4. Aşamalı Ayarlama.....                            | 156 |
| 5.5. İki Değişkenin Birbirine Oranının Tahmini.....   | 164 |
| 5.5.1. Bileşik oran tahmin edicisi.....               | 165 |
| 5.5.2. Ayhan (2003) bileşik oran tahmin edicisi ..... | 167 |
| 6. SONUÇ .....                                        | 174 |
| KAYNAKLAR .....                                       | 179 |
| ÖZGEÇMİŞ .....                                        | 186 |

## ÇİZELGELERİN LİSTESİ

| Çizelge                                                                                                                                                                                    | Sayfa |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| Çizelge 2.1. Tanım kümeleri için ağırlıklandırma .....                                                                                                                                     | 40    |
| Çizelge 4.1. Ağırlıklandırma yöntemlerinin ayarlama ağırlıklandırması<br>sınıflaması .....                                                                                                 | 67    |
| Çizelge 4.2. $x_1$ ve $x_2$ değişkenlerine dayalı çapraz tablo göze yığın ve<br>örnek çaplarının gösterimi.....                                                                            | 84    |
| Çizelge 5.1. Bölge, NUTS 1 Bölgesi, Kır-kent, yerleşim yeri tipi ve illere<br>göre tabaka listesi, Türkiye 2003 .....                                                                      | 119   |
| Çizelge 5.2. Tabakalara göre kadınlar için örnekleme ağırlıkları, yanıt<br>oranları, yanıtlayıcı örnek çapları, düzeltilmiş örnekleme<br>ağırlıkları ve nihai ağırlıklar, TNSA- 2003 ..... | 121   |
| Çizelge 5.3. TNSA–2003 doğurgan yaşlardaki 15-49 yaşında evlenmiş<br>kadınlar verisi.....                                                                                                  | 124   |
| Çizelge 5.4. Örnek uygulaması: yerleşim yeri ve bölge .....                                                                                                                                | 129   |
| Çizelge 5.5. TNSA–2003 örneğindeki uygun kadınların, kentsel/kırsal<br>yerleşim yerlerine göre yanıt oranı ve yanıtlamama ağırlıkları .....                                                | 131   |
| Çizelge 5.6. $\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W_h' y_{chi}$ ve $\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W_h'$ toplam değerlerinin<br>hesaplamaları.....  | 131   |
| Çizelge 5.7. Yaş grupları ve eğitim durumu yanıt alınan örnek dağılımı<br>( $n_{klR}$ 'ler).....                                                                                           | 134   |
| Çizelge 5.8. Yaş grupları ve eğitim durumuna göre yığın dağılımı<br>( $N_{kl}$ 'ler) .....                                                                                                 | 134   |
| Çizelge 5.9. Sonradan tabaka ağırlıkları ( $W_{kl}^{**} = N_{kl} / n_{klR}$ ).....                                                                                                         | 136   |
| Çizelge 5.10. Sonradan tabaka ağırlıklandırması hesaplamaları .....                                                                                                                        | 138   |
| Çizelge 5.11. Çizelge 5.7 ve Çizelge 5.8'deki veri için tarama yinelemesi.....                                                                                                             | 142   |

| Çizelge                                                                                                                                                                | Sayfa |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|
| Çizelge 5.12. $\tilde{N}_{kl}$ , $k$ ıncı satır $l$ inci sütun için son adımdaki göze toplamı değeri .....                                                             | 145   |
| Çizelge 5.13. ${}_{tr}W_{kl}^{**} = \tilde{N}_{kl} / n_{klR}$ tarama ağırlıkları.....                                                                                  | 146   |
| Çizelge 5.14. Tarama ağırlıklandırması hesaplamaları.....                                                                                                              | 147   |
| Çizelge 5.15. $x_2$ ve $x_3$ yardımcı değişkenlerin gözlem değerleri, düzeltilmiş örnekleme ağırlıkları ve ayarlama ağırlıkları .....                                  | 152   |
| Çizelge 5.16. Ayarlama tahmin edicisi hesaplamaları .....                                                                                                              | 153   |
| Çizelge 5.17. Ağırlıklandırma aşamaları sonuçlarının ve ağırlıklandırma yöntemlerinin kıyaslanması .....                                                               | 155   |
| Çizelge 5.18. Birinci aşama ayarlama tahmin edicisi hesaplamaları .....                                                                                                | 158   |
| Çizelge 5.19. İkinci aşama ayarlama tahmin edicisi hesaplamaları.....                                                                                                  | 160   |
| Çizelge 5.20. Aşamalı ayarlama ağırlıklandırması sonuçlarının kıyaslanması.....                                                                                        | 164   |
| Çizelge 5.21. Tabaka birim varyansları .....                                                                                                                           | 166   |
| Çizelge 5.22. $P_h$ , $x_h / P_h$ hesaplamaları ve ${}_AW_h$ bileşik örnekleme ağırlıkları.....                                                                        | 168   |
| Çizelge 5.23. $R_h$ yanıt oranları ve $R_0$ denklemi için hesaplamalar .....                                                                                           | 169   |
| Çizelge 5.24. ${}_AW_h^*$ yanıtlamama için bileşik ağırlıklar.....                                                                                                     | 170   |
| Çizelge 5.25. Yaş gruplarına göre $n_{kR}$ ve $N_k$ dağılımı ile $\sum_{h=1}^H {}_AW_h {}_AW_h^* y_{hkR}$ ve $\sum_{h=1}^H {}_AW_h {}_AW_h^* x_{hkR}$ ağırlıkları..... | 171   |
| Çizelge 5.26. $X_k$ yığın toplamaları ve ${}_AW_k^{**}$ sonradan tabaka ağırlıkları.....                                                                               | 172   |
| Çizelge 5.27. $X_k$ yığın toplamaları ve ${}_AW_k^{**}$ sonradan tabaka ağırlıkları.....                                                                               | 172   |
| Çizelge 5.28. Bileşik oran tahmin edicileri sonuçlarının kıyaslanması .....                                                                                            | 173   |

## SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

| Simgeler       | Açıklama                                 |
|----------------|------------------------------------------|
| $E(.)$         | Beklenen değer                           |
| $V(.)$         | Varyans                                  |
| $\hat{V}(.)$   | Varyans tahmini                          |
| $AV$           | Asimptotik varyans                       |
| $\theta$       | Parametre                                |
| $\hat{\theta}$ | Tahmin edici                             |
| $R$            | İki değişkenin birbirine oranı           |
| $r$            | Örnek oranı                              |
| $\bar{Y}$      | Yığın ortalaması                         |
| $Y$            | Yığın toplamı                            |
| $f$            | Örnekleme kesri                          |
| $F$            | Genişletme faktörü                       |
| $\pi$          | Seçim olasılığı                          |
| $W$            | Örnekleme ağırlığı                       |
| $W'$           | Düzeltilmiş örnekleme ağırlığı           |
| $W^*$          | Yanıtlamama ağırlığı                     |
| $W^{**}$       | Ayarlama ağırlıklandırması ağırlığı      |
| $A^r_b$        | Ayhan [2003] bileşik oran tahmin edicisi |
| $A W_h$        | Ayhan [2003] önerdiği örnekleme ağırlığı |
| $\lambda$      | Lagrange çarpanı                         |
| $\varepsilon$  | Hata terimi                              |
| $w$            | Hájek tahmin edici yapısı                |
| $tb$           | Tabakalı tesadüfi örnekleme              |

**Simgeler****a****st****t****C****Reg****Açıklama**

Ağırlıklandırma sınıf düzeltmesi yöntemi

Sonradan tabakalama

Tarama

Ayarlama tahmin edicisi

Regresyon tahmin edicisi

**Kısaltmalar****Açıklama****BTÖ**

Basit tesadüfi örnekleme

**GREG**

Genelleştirilmiş regresyon tahmin edicisi

**HA**

Hájek

**HKO**

Hata kareler ortalaması

**HT**

Horvitz-Thompson

**KÖ**

Küme örnekleme

**ST**

Sonradan tabakalama

**TÖ**

Tabakalı tesadüfi örnekleme

**VIF**

Varyans şişme faktörü

## 1. GİRİŞ

Anket arařtırmalarının amacı, üzerinde inceleme yapılan birimlerin oluřturduėu yığında yapılan bir tamsayımdan veya yığından sınırlı sayıda seilen birimlerin oluřturduėu bir rnek zerinden gzlemler yapılmasıyla yığın hakkında genel ıkarsamalar yapmaktır.

Arařtırma tamsayım veya rnek birimleri zerinden yrtlyor olabilir. Tamsayım arařtırma konusu yığın birimlerinin tamamına uygulanan bir alıřma demektir. Ancak tm yığın birimlerinin belirlenemediėi, incelenemeyecek kadar ok sayıda birim olduėu, yığın birimlerini inceleme sresi ve maliyetinin ok yksek olduėu durumlarda tamsayım yapılamaz. Tamsayım yerine yığını temsil edecek nitelikte, yığının yapısına uygun olan bir rneklemeye yntemiyle daha kk geniřlikli bir rnek seilerek yığına iliřkin istenen zellikler belirlenmeye alıřılır [ıngı, 1994].

oėu arařtırmada, yığın hakkında bilgi edinme ve yığının zelliklerini ėrenme iřlemi rnek zerinden yapılır. rnek, arařtırmanın konusunu oluřturan yığının tm zelliklerini yansıtan yığının bir parasıdır. Yığını temsil ettiėi varsayılan rnekten yığının ilgilenilen zelliklerinin (parametrelerinin) tahminleri yapılır.

Yığının zelliklerini tanımlayan sayısal llere (karakteristiklere) parametre adı verilir. Parametre genel gsterimi iin  $\theta$  kullanılır.  $\bar{Y}$  yığın ortalaması,  $Y$  yığın toplamı,  $P$  yığında belli zelliėe sahip birimler oranı,  $R = Y / X$  iki deėiřkenin birbirine oranı nemli yığın parametreleridir.

Yığın parametrelerinin rnekten hesaplanan tahminlerine tahmin ediciler (estimators) denir. Tahmin edicilerin genel gsteriminde  $\hat{\theta}$  kullanılır. İyi bir tahmin edicinin tutarlılık (consistency), yansızlık (unbias) ve etkinlik (efficiency) zelliklerini saėlaması gerekir. İyi bir tahmin ediciyle daėılımı, tahmin edilecek yığın parametresine yakın civarda yoėunlařan bir tahmin edici kastedilir [Esin ve ark., 2001].

Tahmin edicilerin etkin olması istenir. En etkin tahmin edici en az hataya sahip tahmin edici demektir. Etkinliğin ölçüsü hata kareler ortalamasıdır.  $\hat{\theta}$ , tahmin edicisinin hata kareler ortalaması  $HKO(\hat{\theta}) = E[\hat{\theta} - \theta]^2 = V(\hat{\theta}) + (Yan)^2$  şeklinde tanımlanmıştır. Hata kareler ortalaması en küçük olan tahmin edici en etkin ve en iyi tahmin edici demektir.

$V(\hat{\theta})$ ,  $\hat{\theta}$  tahmin edicinin varyansıdır.  $V(\hat{\theta})$ , değişken hatası olarak adlandırılır ve yalnızca örnekleme hataları nedeniyle ortaya çıkar [Kish, 1965]. Varyansın tersi duyarlılık (precision) ölçüsüdür. Tahmin edicilerin duyarlı olması istenir. Yansız tahmin ediciler arasında en küçük varyansı olan en duyarlı ve en etkin tahmin edicidir.

HKO'nun diğer bileşeni yan, örnekten elde edilen ortalama değer gerçekteki yığın değerinden sapmasıdır. Yetersiz örnekleme çerçevesinin kullanılması, örnek seçim sürecinin yanlış uygulanması, verilerin toplanma ve değerlendirilme işlemlerinin hatalı olması, yanıt alınamayan birimlerin olması vb. nedenler yanlış tahminlere yol açabilir [Cochran, 1977].

Hata kareler ortalamasının minimum olması istenir ve araştırma sonuçlarını etkileyen hataların ölçülmesi ve azaltılması için çaba harcanması gerekir. Araştırmalarda toplam hata; örnekleme hataları ve örnekleme dışı hatalar olarak ikiye ayrılabilir [Verma, 2002].

Araştırma istatistiği (tahmin edici) gerçekteki değerden (parametreden) iki nedenle farklı olabilir. Bilgi tam sayımdan değil, örneklemeyle dayalı araştırmadan elde edildiği için; bu tür hatalar örnekleme hatalarıdır. Ölçüm sürecinde hatalar yapıldığı için; bu tür hatalar örnekleme dışı hatalardır [Kish, 1965].

Yığının yalnızca bir altkümesine ilişkin bilgi olması nedeniyle ortaya çıkan hatalara örnekleme hataları denilir. Bu hataların etkileri azaltılabilmekle birlikte hiçbir zaman

tam olarak ortadan kaldırılamaz. Diğer bir ifadeyle, hedef yığında bulunan tüm birimler araştırmaya dahil edilmediği sürece örnekleme hataları kaçınılmazdır.

Uygulamalı araştırmalarda örnekleme süreciyle ilişkisi olmayan hatalara örnekleme dışı hatalar denir. Örnekleme dışı hatalar tahmin edicilerin yanlı olmasına yol açar. Üç tür örnekleme dışı hata vardır:

- Kapsam hatası (coverage error),
- Yanıtlanmama (nonresponse),
- Ölçüm hatası (measurment error).

Yığın örnekleme çerçevesinde tam ve doğru bir şekilde temsil edilmediğinde, kapsam hatası yapılmış demektir. İyi bir araştırmada öncelikle hedef yığın ve çerçeve doğru belirlenerek kapsam hatasından kaçınılmalıdır [Verma, 2002].

Hedef yığındaki bazı birimlerin örnekleme çerçevesine dahil edilmediğinde düşük kapsanma (under coverage) var demektir. Örnekleme çerçevesindeki bazı birimler hedef yığına ait olmadığında üstünde kapsama (over coverage) olur. Yığındaki bazı birimler çerçevede birden çok kez görüldüğünde tekrar (duplication) var demektir.

En çok karşılaşılan sorun düşük kapsanma olmasıdır. Listede güncelleme yapılarak listede olmayan birimler çerçeveye dahil edilebilir. Düşük kapsanmış olma çerçevedeki bir başarısızlık olup hedef yığına ait tüm birimlere ulaşılmamasına neden olur. Ulaşılamayan birimler seçilemeyecek, daha az gözlenecek ve sıfır seçilme olasılığıyla örnekte yer alacaktır. Araştırmanın amacı dışındaki birimler çalışmaya dahil edildiğinde üstünde kapsama sorunuyla karşılaşılır [Verma, 2002].

Kapsam hatalarının boyutunu ve etkisini tahmin etmek kolay değildir. Örnek dışında yığın veya dışsal kaynaklı bilgi olmasını gerektirir. Toplam kapsam hatasının etkisi ve örnek dağılımının yığın dağılımından sapmaları, örnek dağılımında ağırlıklandırma yapılmasıyla azaltılabilir [Verma, 2002].

Yanıtlanmama olması yana neden olur çünkü genellikle yanıt alınamayanlar yanıtlayanlardan farklıdır ve yanıtlanmamadan kaynaklanan yanı ihmal etmek zordur [Rubin, 1987]. Yanıt alınamayanların yanıtlayanlardan farklı olması nedeniyle yanıt alınanlar örneğinin yığını temsil etmesi gücü düşüktür. En iyi çözüm en yüksek cevap oranını elde etmektir. Önlenemeyen yanıtlanmama için örnekleme ağırlıkları düzeltilerek yanıtlanmamadan kaynaklanan yan azaltılmaya çalışılabilir.

Anket araştırmalarda değişkenlerin, fikirlerin sınanması için toplanacak verilerin önemli bir özelliği de ölçülebilir olmasıdır. Kullanılan ölçek türü ve sorulara verilen yanıt alternatifi sayısı kişilerin soruları cevaplamasında etkili olmaktadır. Ölçüm hatasına yol açabilecek önyargılı ve karmaşık sorulardan kaçınılmalı, sorular dikkatli hazırlanmalı ve önceden test edilmelidir.

Örnek çapı büyüdükçe örnekleme hataları azalırken örnekten elde edilen tahmin değeri parametre değerine yakın bir değer alır. Örnekleme hataları örnek çapı arttırılarak azaltılabilir, ancak örnekleme dışı hatalar örnek çapı arttırıldığında azalmaz, aksine örnek çapı büyüdükçe artış eğilimi gösterir. Örnekleme dışı hatalar, hem örnek araştırmalarında hem de tam sayımlarda vardır ve tam sayımlarda daha fazla olmaktadır [Churchill, 1996].

Yığın ve örnek arasındaki farklılıklar örnekten yapılan tahminlerin yanlı olmasına neden olur. Ağırlıklandırma yanın düzeltilmesinde kullanılacak önemli bir yöntem olarak tanımlanmıştır. Ağırlıklar, yanlı bir örnekten yansız tahminleri elde etmek amacıyla uygulanır [Ayhan, 1981].

Ağırlıklandırma, örnekleme ve örnekleme dışı hataları düzeltmek, tahminlerin etkinliğini arttırmak için araştırma verisine uygulanır. Araştırmalarda parametre tahminlerinin kalitesini arttırmak için bazı ağırlıklandırma türleri yürütülür.

Veriye ağırlıklar uygulamanın amacı örnek ile yığın arasında ayrılıklara ve yana yol açabilecek; birimlerin eşit olmayan olasılıklarla seçilmesi, yanıtlanmama olması ve

yığının doğru kapsanmamış olması durumlarında yapılacak tahminlerin yığının tamamını temsil edecek nitelikte olmasını sağlamaktır.

Araştırmacı veriyi seçim olasılığı, yanıtlanmama, tabakalama ve diğer amaçlar için ardışık olarak düzeltirse ağırlıklandırma çok adımlı bir süreç olabilir. Genel olarak ağırlıklandırmayı aşağıdaki adımlarda gruplamak mümkündür [Verma, 2007a]:

- Örnekleme (Tasarım) ağırlıklandırması,
- Yanıtlanmama ağırlıklandırması,
- Ayarlama ağırlıklandırması,
- Kırpma ve ölçekleme.

Ağırlıklandırma birbirini izleyen adımlarda uygulanır. Her adımda hesaplanan ağırlıklar bir önceki adımdaki ağırlıklarla çarpılır. İlk adımda verideki gözlem değerlerine örnekleme ağırlıkları atanır. İkinci adımda yanıtlanmama ağırlıklandırması için bir önceki adımda gözlem değerlerine atanan örnekleme ağırlıklarına yanıtlanmama ağırlıkları çarpan olarak atanır. Ağırlıklandırma bu şekilde adımsal olarak yürütülür.

Araştırma verisinde ağırlıklandırma yapılmasına, çoğunlukla eşit olasılıklarla seçim yapılmamış olmasından dolayı her bir örneklenen birime tasarım (örnekleme) ağırlıkların verilmesiyle başlanılır.

Veride ağırlıklandırma gelişmesinin ikinci aşaması çoğu kez birim veya toplam yanıtlanmamayı gidermek için uygulanır. Araştırmalarda örnekleme çerçevesini ve örneğe seçilen birimlerden yanıt alınmayanların nedenini açıklamak için yanıtlayanları ayrımsal olarak ağırlıklandırmak faydalı olur.

Ağırlıklandırmanın son aşamasında örnek verisinde ilave ağırlıklandırma yapılmasıyla, örnekteki birimlerin dağılımının araştırmanın önemli karakteristikleri (cinsiyet, bölge, yaş, eğitim vb.) için yığın dağılımına uyumu sağlanabilir. Ağırlıklar

yığının kapsanmaması durumunda ortaya çıkabilecek yanı ve örnek ile yığın arasındaki diğer ayrılıkları gidermek amacıyla araştırma verisine uygulanır. Bu ağırlıklandırma türü ayarlama ağırlıklandırması başlığı altında incelenebilir.

Kırpma (trimming) ve ölçekleme (scaling), ağırlıkların dağılımındaki uç değerleri önleyen düzeltmelerdir. Bu tür düzeltmeler bazı yanlara yol açıyor olmasına karşın, uç ağırlıklardan dolayı varyansta oluşabilecek yüksek artışları önlemek amacıyla yapılabilir. Ağırlıklandırma işlemlerinin hiç biri ağırlıkların uç değerleri olmayacağını garanti etmez. Ağırlık değerlerindeki büyük değişimler araştırma tahminlerinin varyansını arttırır [Kish, 1992].

Genel kabul görmüş bir yaklaşım uç değerlerin kırılmasıdır. Ancak kırpmanın sınırlarını belirlemede genel bir kısıt yoktur [Verma ve ark., 2007]. Kırpma ve ölçekleme çalışma kapsamında incelenmemektedir. Bu konuda, DHS (2005), Verma, Betti ve Ghellini (2007) ve Ünüvar (1997) çalışmalarına başvurulabilir.

Yanıtlanmama ve ayarlama ağırlıklandırması yapmada çeşitli ağırlıklandırma yöntemleri kullanılır. Yardımcı değişken veya değişkenlere ilişkin yığın hakkında yeterli bilgi varsa, ağırlıklandırma yöntemleriyle gözlemlere ağırlıklar uygulanmasıyla parametre tahminlerinin etkinliği arttırılabilir.

Göze ağırlıklandırması, ağırlıklandırma sınıf düzeltmesi, sonradan tabakalama, tarama, lineer ağırlıklandırma, lojistik regresyon ağırlıklandırması, genelleştirilmiş regresyon tahmin edicisi (GREG), ayarlama tahmin edicisi ve daha başka yöntemler ağırlıklandırma yöntemleri olarak kullanılmaktadır.

Kish (1990, 1992) çalışmaları ağırlıklandırmanın temel teşkil eden başvuru kaynaklarıdır. Bu çalışmalarda ağırlıklandırmanın neden, ne zaman ve nasıl yapılacağı üzerine detaylı bilgiler verilmektedir. Kalton ve Flores-Cervantes (2003) çalışması ağırlıklandırma yöntemlerinin incelenmesi bakımından önemlidir. Son yıllarda ağırlık düzeltmeleri yapmada geniş çapta yardımcı bilgiyi kullanan

ağırlıklandırma yöntemlerinin kullanımında artış ve teorik gelişmeler olduğu görülmektedir.

Ağırlıklandırma uygulamalarında zorluklar yaşanmaktadır. İyi yazılımlı bilgisayar programları olmasına karşın, sistemlerde hatalar olduğu ve karmaşık analizlerde hataların boyutunun arttığı görülmektedir. Teorik çalışmalar hızla geliştirilmektedir ancak bu çalışmalar çoğu araştırmacıya ulaşamamakta veya genel ve yeterli düzeyde basit olmadıklarından yararlı olamamaktadır [Kish, 1992].

Çoğu enstitü ve araştırma kurumları hanelerin çeşitli karakteristiklerini kullanarak yanıtlarını ağırlıklandırmakta ve genel sonuçlar çıkarsamaktadır. Bununla birlikte, bazı kurumlar ağırlıklandırmayı kullanmamaktadır. Ancak, tüm hanelerin eşit seçilme şansı olduğu, farklılık gösteren cevaplanma durumları olmadığı, yığının doğru kapsandığı durumlarda veride ağırlıklandırma yapılmasına gerek olmayacağı söylenebilir. Ağırlıklandırma gerektiren veri ağırlıklandırılmamışsa, sonuç tahminleri yığınla ilgili yanlış yorumlamalar yapılmasına neden olacaktır.

Çalışma kapsamında, örnekleme araştırmalarında ağırlıklandırma süreci ele alınmaktadır. Örnekleme, yanıtlanmama ve ayarlama ağırlıklandırması ayrı bölüm başlıkları altında incelenmektedir. Bunun için, günümüze kadar yapılan çalışmalar incelenmiş ve detaylı bir çalışma yapılması amaçlanmıştır.

Literatürde ağırlıklandırma aşamaları ve ağırlıklandırma yöntemlerinin Horvitz-Thompson (1952) yapısında tanımlamaları vardır. Çalışmada ağırlıklandırmanın hem Horvitz-Thompson (1952) hem de Hájek (1971) tahmin edicisi yapısında tanımlamaları verilmeye çalışılacaktır.

Ağırlıklandırma yöntemlerinden ayarlama tahmin edicisinin incelenmesine özel bir önem verilmiştir. Ayarlama tek başına bir ağırlıklandırma yöntemidir. Bununla birlikte, literatürde sonradan tabakalama, tarama, GREG ağırlıklandırmasının ve daha başka yöntemlerin de birer ayarlama ağırlıklandırması yöntemi olduğu ispatlanmıştır.

Ağırlıklandırma sınıf düzeltmesi, sonradan tabakalama ve tarama yöntemlerinde ağırlıklar, yardımcı değişkenlere dayalı oluşturulan çapraz sınıflamalardaki yığın ve örnek çaplarına dayalı belirlenir. Ayarlama tahmin edicisiyle ağırlıkları belirlemede hem yığın çapı ve örnek çapı bilgileri kullanılıyorken hem de yardımcı bilginin örnekten gözlenen değerleri,  $\hat{X}$  örnek toplamı ve  $X$  yığın toplamı bilgileri kullanılır. Ayarlama tahmin edicisinde ağırlıkları belirlemede yardımcı bilgidен daha fazla yararlanılır.

Ayhan (2003)'ün önerdiği bileşik oran tahmin edicisi yöntemiyle ağırlıklar yardımcı değişkenlerin gözlem değerlerine ve  $X$  yığın toplamı bilgilerine dayalı tanımlanmıştır. Bilinen bileşik oran tahmin edicisine kıyasla bu yöntemle ağırlıkları belirlemede yardımcı değişkenlere ilişkin daha çok bilgi kullanılmaktadır.

Ayarlama tahmin edicisi ve Ayhan (2003)'ün önerdiği bileşik oran tahmin edicisinde ağırlıkları belirlemede daha çok bilgi içeriliyor olmasının parametre tahminlerinin doğruluğunda ilave artış sağlayacağı beklenebilir.

Ayarlama tahmin edicisinde örnekleme tasarımı ve yardımcı değişkenlere dayalı bir çalışma modeli altında etkin tahminler yapılması amaçlanır.  $y$  ve  $x$ 'ler arasındaki ilişkiye dayalı modelle, yardımcı bilginin tahmin aşamasında en iyi nasıl kullanılabileceği saptanmaya çalışılır [Wu ve Sitter, 2001].

Uygulamalı araştırmalarda genellikle  $y$  ilgilenilen değişkeniyle ilişkili çok sayıda yardımcı değişkene ait bilgiler derlenir. İlgilenilen değişkeni açıklama niteliğine sahip çok sayıda değişken varken, değişkenlerin bağımsız olması şartı nedeniyle tahmin modeline çok az sayıda değişken alınabilmektedir. Modele alınacak değişken seçimi koşullu olarak yapılabilir. Koşullu stratejileri tahminin doğruluğunu artırır ve uç değerler olmasını engeller [Wywiał, 2003].

Pek çok araştırmada genellikle bir değişkene göre dağılımı incelenen örnek verisinin, başka değişkenlere dayalı incelemeleri de yapılmak istenilebilir. Böyle durumlarda

koşullu dağılım kullanımı ve değişkenlerin tahmin modeline aşamalı eklenmesi istenilen tahminlere daha iyi ulaşmamızda katkı sağlayacaktır. Bilinenden yola çıkarak bilinmeyene ulaşmak gerçek sonuçlara biraz daha yaklaşmamızı sağlayacaktır.

Ayarlama tahmin edicisi yardımcı değişkenlerin modele aşamalı eklenmesine olanak sağlayarak koşullu (aşamalı) ağırlıklandırmayla parametrelerin daha iyi tahminlerinin yapılması olanaklı olacaktır.

Çalışmanın bir sonraki bölümü olan ikinci bölümde örnekleme ağırlıklandırması incelenmektedir. Olasılı örnekleme yöntemlerinden basit tesadüfi örnekleme, tabakalı tesadüfi örnekleme ve küme örnekleme tasarımlarında ağırlıklandırma ve parametre tahminlerinin Horvitz-Thompson (1952) ve Hájek (1971) tahmin edicileri yapıları verilmektedir. Ayhan (2003)'ün önerdiği bileşik oran tahmin edicisine yer verilmektedir.

Üçüncü bölümde, yanıtlanmama ağırlıklandırması üzerinde durulmaktadır. Yanıtlanmama ağırlıklandırması yöntemlerinden en yaygın kullanılan ağırlıklandırma sınıf düzeltmesi yöntemi incelenmektedir.

Dördüncü bölümde, ağırlıklandırmanın son aşaması olan ayarlama ağırlıklandırması üzerinde durulmaktadır. Ayarlama ağırlıklandırması yapmak amacıyla geliştirilen ağırlıklandırma yöntemlerinden sonradan tabakalama, tarama, ayarlama tahmin edicisi, genelleştirilmiş regresyon (GREG) tahmin edicisi yöntemleri ve aşamalı ayarlama tahmin edicisi incelenmektedir.

Çalışmanın beşinci ve son bölümü olan uygulamada Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü tarafından gerçekleştirilen 2003 Türkiye Nüfus ve Sağlık Araştırmasının (TNSA-2003) doğurgan yaşlardaki 15-49 yaşında evlenmiş kadınlar verisi kullanılmıştır. Araştırma verisinde örnekleme ağırlıklandırması, yanıtlanmama ağırlıklandırması ve ayarlama ağırlıklandırması incelenerek elde edilen tahmin sonuçları kıyaslanmaktadır. Ayrıca aşamalı ayarlama ve Ayhan (2003)'ün önerdiği

bileşik oran tahmin edicisi incelenmesiyle çalışma tamamlanmıştır. Bu uygulama çalışmasının gelecekte yapılacak çalışmalara faydalı olacağı beklenmektedir.

## 2. ÖRNEKLEME AĞIRLIKLANDIRMASI

Ağırlıklandırmanın birinci aşaması eşit olmayan seçim olasılıkları için ağırlıklandırma yapmaktır. Örneğe seçilen (yanıt alınan veya yanıt alınmayan) her bir birime örneğe seçim olasılığının tersi örnekleme (tasarım) ağırlığı olarak atanır. Olasılı örneklemeyle seçim olasılıkları biliniyordur ve örnekleme ağırlıkları genellikle kolayca hesaplanır. [Kalton ve Flores-Cervantes, 2003]. Ancak bazen örnekleme çerçevesi problemleri nedeniyle örnekleme ağırlıklarını belirlemede zorluklar yaşanabilmektedir.

Örneklemede, tamsayım yerine yığını temsil edecek nitelikte, yığının yapısına uygun bir örnekleme yöntemiyle daha küçük genişlikli bir örnek seçilerek yığına ilişkin istenen özellikler belirlenmeye çalışılır. Çoğu araştırmada, yığın hakkında bilgi edinme ve yığının özelliklerini öğrenme işlemi örnek üzerinden yapılır.

Örnekleme kuramı, sonlu  $N$  sayıda yığın birimi içeren yığınlardan,  $n$  büyüklüğünde tesadüfi örnekler seçme ve seçilen örneklerden tahminler yapma yöntemlerini inceler ( $n < N$ ). Bir başka deyişle, örnekleme kuramının konusu, yığından, yığının yapısına en uygun örnekleme yöntemiyle, örnek seçme süreci ve örnekten yığının özelliklerinin tahmin edilmesi sürecidir. Seçim sürecinde kullanılan yöntemle göre yığın parametreleri tahmin edilir [Çingir, 1994].  $\bar{Y}$ , yığın ortalaması;  $Y$ , yığın toplamı;  $P$ , istenen özelliğe sahip birimler oranı veya  $R$ , yığın oranı parametreleri tahmin edilmek istenebilir.

Bilimselliği olan araştırmalarda, örnekler gelişigüzel seçilmez ya da araştırmaya sadece gönüllü birimler alınmaz. Seçimler bilimsel bir tekniğe dayandırılmalı ve yığındaki her bir birim ölçülebilir seçilme şansına sahip olmalıdır. Ancak bu yolla, örnekten elde edilen tahmin sonuçları güvenilir bir şekilde büyük yığınlara uyarlanabilir.

Yığındaki her bir birimin bilinen ve sıfır olmayan olasılıklarla örneğe seçildiği, örnekleme yöntemlerine olasılı örnekleme yöntemleri adı verilir. En temel olasılı

örnekleme yöntemleri basit tesadüfi örnekleme, sistematik örnekleme, tabakalı örnekleme ve küme örnekleme yöntemleridir.

Örnekleme yöntemleri birimlere farklı seçim olasılıklarını atayarak birimlerin örneğe seçilmesine yol açıyor olabilir. Tabakalı örnekleme, küme örnekleme, çok aşamalı örnekleme birimlerinin örneğe eşit olmayan olasılıklarla seçildiği olasılı örnekleme yöntemleridir. Olasılı örneklemede seçim olasılıkları bilindiği için seçim olasılıklarının tersi olan örnekleme ağırlıkları (tasarım ağırlıkları) genellikle kolayca belirlenir. Bu yöntemlerde eşit olmayan seçim olasılıklarını telafi etmek için her bir örneklenen birime seçim olasılığının tersi örnekleme ağırlığı olarak atanır.

$N$  çaplı yığından eşit olasılıkla yerine koymaksızın  $n$  çaplı bir örneğin seçim olasılığı  $f = \pi_i = n/N$  olacaktır ( $f$ , örnekleme kesridir (sampling fraction)). Yığındaki herhangi bir toplam tahmini, örnekteki toplam değerinin  $N/n$  genişletme faktörüyle (expansion factor) çarpılmasıyla elde edilecektir. Ancak, seçim olasılıkları değiştiğinde  $N/n$  genişletme faktörü (daha genel ifadeyle ağırlık olarak da adlandırılır) birimden birime farklılık gösterecektir [ORC Macro, 1996].

Yığındaki bir  $i$  birimi  $\pi_i$  olasılığıyla örneğe seçildiğinde,  $i$  inci birimin örnekleme (tasarım) ağırlığı  $W_i = 1/\pi_i$  şeklinde hesaplanır ( $\pi_i = 1/F_i$ ,  $i$  inci birimin  $n$  çaplı örnekte içerilmesi olasılığı ve  $F_i$ , genişletme faktörüdür). Mesela, birim yığından  $1/300$  olasılığıyla örneğe seçilmişse, bu birim yığındaki 300 birimi temsil ediyor demektir. Dolayısıyla  $W_i = 1/\pi_i = 1/f_i = F_i$  örnekleme ağırlıkları, genişletme faktörleri olarak rol oynar ve araştırma yığınınındaki birimlerin sayısını temsil eder.

$W_i$  örnekleme ağırlıkları, Deville ve Särndal (1992), Verma (2007b), Verma (2008, syf. 231), DHS (2005) TMSD(2006) ve Verma (2007b) tarafından  $W_i = 1/\pi_i = 1/f_i$  şeklinde seçim olasılığının tersiyle; Kish (1992) ve Kalton ve Flores-Cervantes (2003) tarafından  $W_i \propto 1/\pi_i$  seçim olasılıklarına orantısal olacak şekilde tanımlanmıştır.

Örnek verisi her bir birime eşit seçilme olasılığı atanması yönteminden gelmemişse ağırlıklandırma yapılarak yanıtlayıcıların eşit olmayan olasılıkla seçimi ve seçilen birimlerin grupları arasındaki ayrımsal yanıt oranları düzeltilmelidir. Böylece her bir birimin örnekte eşit düzeyde temsili sağlanmalıdır. Tabakalı örnekleme, küme örnekleme ve çok aşamalı örnekleme yöntemleri gözlem birimlerine ağırlıkların atanmasını gerektiren yöntemlerdir.

Örnekteki her birimin eşit örnekleme ağırlığı olduğunda örnekleme tasarımına kendinden ağırlıklı tasarım denir. Kendinden ağırlıklı örnekleme (self weighting sampling) tasarımlarında parametre tahminleri için gözlem değerlerine ağırlık atanmasına gerek yoktur [Kish, 1990]. Basit tesadüfi örnekleme, sistematik örnekleme ve çapa orantılı olasılıklı örnekleme yöntemleri kendinden ağırlıklı tasarımlardır.

Bu tasarımlarda tahminler yapılırken ağırlıklar kullanılmaz. Kendinden ağırlıklı örnekler çoğunlukla tercih edilir, çünkü işlem kolaylığında, varyans azaltmada ve sağlamlıkta<sup>1</sup> (robustness) önemli düzeyde avantajlar sağlarlar [Kish, 1990].

Yığından örnek seçim sürecinde kullanılan örnekleme yöntemine göre örnek çapı belirlenir. Yığından seçilen örnek üzerinden parametrelerin tahminleri kullanılan örnekleme yöntemlerine göre yapılır. Örnekleme yöntemlerinde parametre tahminleri

---

<sup>1</sup> Esas dağılım varsayılan dağılımdan (normal dağılımlı olma gibi) ayrılık gösterdiğinde sağlam (robust) olma kurallarına gereksinim duyarız, dağılımın sağlanmaması durumlarında etkinlik ve tahmin edicilerin dağılımı çok etkilenmemesi durumudur. Yığın ortalamasının bir tahmin edicisi olarak  $\bar{y}_{HT} = \bar{y}$  örnek ortalaması sağlam bir istatistik değildir. Mesela, gözlemlerden bir tanesi çok farklı bir değer aldığında örnek ortalamasının değeri bir anda çok farklı olabilmektedir, uç değer olması gibi. Ama ortanca tek bir değer değişmesinden çok da fazla etkilenmemektedir.  $\theta$ 'nın bir  $\hat{\theta}$  tahmin edicisi varsayılan bir dağılım için tam etkinse sağlamdır denir. Tam etkin tahmin edici geniş (large)  $n$  için yansızdır ve varyansı Cramer-Rao minimum varyans sınırına eşittir [Oral, 2007]. Sağlam tahmin edicinin uç değerlere karşı tepkisi düşüktür, düşük varyanslı ve düşük yanlılığa sahiptir [Hulliger, 1999].

Horvitz ve Thompson (1952) (HT) ve Hájek (1971) (HA) tahmin edicileri yapısında formüle edilmiş olabilir.

Örnekleme kuramıyla ilgili temel kaynaklarda  $\bar{Y}$ ,  $Y$  ve  $P$  parametrelerinin tahmin edicileri Horvitz-Thompson yapısına dayalı tanımlanmıştır. Cochran (1977), Çıngı (1994), Esin ve ark. (2001), Horvitz-Thompson'a dayalı formüle edilmiş tahmin edicileri incelemiştir. Tillé (2005)'in çalışmasında örnekleme kapsamında verilen tahmin edicilerin HT oldukları detaylı olarak incelenmiştir. Särndal, Swenson ve Wretman(1992), yerine koymaksızın basit tesadüfi örneklemede ve tanım kümelerinde (domains) yığın ortalamasının HA tahmin edicisi yapısını incelemiştir.  $R = Y / X$ , iki değişkenin oranının tahmin edicisi ise iki HT tahmin edicisi oranı şeklinde tanımlanmıştır.

Bu çalışmada olası örnekleme yöntemlerinden basit tesadüfi örnekleme, tabakalı tesadüfi örnekleme ve basit (iki aşamalı) küme örnekleme yöntemi incelenecektir. Bu yöntemlerde örnekleme ağırlıkları ile Horvitz-Thompson ve Hájek tahmin edicileri incelenecektir. Bu bölümdeki formüller yığından seçilen tüm örnek birimlerinden yanıt alınması durumu için verilmiştir.

Bu bölümde öncelikli olarak eşit olası örnekleme ve eşit olmayan olası örnekleme yöntemlerinden bahsedilecektir. Sonrasında örnekleme yöntemlerinde ağırlıklandırma incelenecektir. Ancak örnekleme yöntemlerinde yığın ortalaması ve yığın toplamı tahminlerin Horvitz-Thompson ve Hájek tahmin edici yapılarında tanımları yapılacağından öncelikli olarak Horvitz-Thompson ve Hájek tahmin edicisi üzerinde durulacaktır.

Olasılı örnekleme genellikle,

- Eşit Olasılı Örnekleme
- Eşit Olmayan Olasılı Örnekleme

olarak sınıflanır [Kish, 1965].

### 2.1. Eşit Olasılı Örneklem

Seçim aşamasında örnek birimlerine eşit seçilme şanslarının verilmiş olması örneklem kuramında çok önemlidir. Yığındaki birimlerin örneğe seçilme şansları aynı olduğunda eşit olasılıkla seçilmiş örneklem tasarımı var demektir.

### 2.2. Eşit Olmayan Olasılı Örneklem

Uygulamada çoğunlukla farklı birimlere farklı seçim olasılıkları atanması için nedenler vardır. Yeterli doğruluğu sağlamak amacıyla orantısız olmayacak biçimde bazı tabakalardan, tanım kümelerinden, bölgelerden veya belirli özelliklere sahip birimlerden daha çok sayıda birim seçmek daha iyi tahminler yapılmasına imkan veriyor olabilir.

Bazı tabaka veya tanım kümelerindeki birimler diğer tabakalara (veya tanım kümelerine) kıyasla daha yüksek (veya daha düşük) olasılıkla seçilir. Mesela, tabakalı örneklem yöntemiyle örneklem birimleri büyük çaplı tabakalardan daha yüksek seçim olasılıklarıyla, küçük çaplı tabakalardan daha düşük seçim olasılıklarıyla örneğe seçilebilir. Örneklem birimleri, geniş çaplı tabakalardan daha büyük olasılıklarla seçilerek veya değişkenliğin yüksek olduğu tabakalardan daha fazla sayıda seçilerek istenilen parametre tahminlerinin yeterli doğruluğu sağlanmak istenir.

Bir ülkenin kentsel alanın veya daha küçük alanların istenilen sonuçlar üzerine katkılarının daha çok olduğu düşünülüyorsa, bu alanlardan fazla sayıda birim seçilmek istenir. Örneğin, Türkiye Nüfus ve Sağlık Araştırması 2003'de İstanbul ve Güney Doğu Anadolu Projesi (GAP) bölgeleri sahip oldukları özel konumları nedeniyle örneklem tasarımında büyük gözlem sayılarıyla temsil edilmişlerdir.

Birimlerin seçim olasılıklarındaki değişimler tahmin aşamasında giderilmelidir. Bu, her bir birime seçim olasılığının tersinin ağırlık olarak atanması şeklinde gerçekleştirilebilir [Verma, 2002].

Birimlerin seçim olasılıklarının tersiyle ağırlıklandırılması oldukça eskiye dayanmaktadır. Hansen ve Hurwitz (1943), Horvitz ve Thompson (1952) çalışmalarında olduğu gibi ve ağırlıklandırma günümüzde yaygın olarak benimsenecek hale gelmiştir. Ters olasılık ağırlıkları, 1'e eşit veya 1'den büyük değerlidir. Gözlenen bir birim kendini ve daha başka sayıdaki gözlenmemiş olan diğer birimleri temsil ediyor demektir.

Yığından örnek seçimi yerine koyarak veya yerine koymaksızın yapılabilir. Hansen ve Hurwitz (1943), yerine koyarak eşit olmayan olasılıklı örnekleme yapma fikrini öne sürmüştür.

Yerine koymaksızın eşit olmayan olasılıklı örnekleme ilk kez Madow (1949) tarafından ve daha sonra Narian (1951) tarafından önerilmiştir ancak teorik bir çerçeve verilmemiştir [Shahbaz ve Hanif, 2003]. Horvitz ve Thompson (1952) yerine koymaksızın eşit olmayan olasılıklı örneklemeyle ilgili ilk teorik yapıyı vermiştir.

### 2.3. Horvitz-Thompson Tahmin Edicisi

Horvitz ve Thompson (1952), yerine koymaksızın eşit olmayan olasılıkla örneklemenin genel bir teorisini oluşturmuştur. Horvitz-Thompson (1952) tahmin edicisi literatürde ters-olasılık ağırlıklandırması (invers-probability weighting) veya  $\pi$ -ters ağırlıklandırması ( $\pi$ -invers weighting) adlarıyla da kullanılmaktadır.

$U$ , yığını  $N$  sayıda  $u_1, u_2, \dots, u_N$  birimlerinden oluşuyor olsun.  $n$  çaplı örnek yerine koymaksızın örnekleme yöntemiyle seçiliyor olsun.  $\pi_i = P(u_i)$ ,  $i$  inci birimin  $n$  çaplı örneğe seçilmesi olasılığıdır ( $i = 1, \dots, N$ ) ve

$$\sum_{i=1}^N P(u_i) = \sum_{i=1}^N \pi_i = n. \quad (2.1)$$

Sonlu  $N$  çaplı bir yığından  $n$  çaplı örnekleri yerine koymaksızın örneklemede  $\binom{N}{n}$  farklı şekilde seçilir.

a) Yığın toplamının HT tahmin edicisi

$Y$  yığın toplamı,

$$Y = \sum_{i=1}^N y_i$$

ve yığın toplamının yansız bir tahmin edicisi Horvitz ve Thompson (1952) tarafından,

$$\hat{Y}_{HT} = \sum_{i=1}^n \frac{y_i}{\pi_i} \quad (2.2)$$

önerilmiştir.

Horvitz ve Thompson (1952),  $\hat{Y}_{HT}$  yansız tahmin edicisinin varyansını,

$$V(\hat{Y}_{HT}) = E[\hat{Y}_{HT} - Y]^2 \quad (2.3)$$

$$V(\hat{Y}_{HT}) = \sum_{i=1}^N \frac{y_i^2}{\pi_i} + \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N \frac{\pi_{ij}}{\pi_i \pi_j} y_i y_j - Y^2$$

$$V(\hat{Y}_{HT}) = \sum_{i=1}^N \frac{1 - \pi_i}{\pi_i} y_i^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} y_i y_j \quad (2.4)$$

şeklinde belirlemiştir. Burada tüm  $i$  ve  $j$ 'ler için  $\pi_i > 0$  ve  $\pi_{ij} > 0$  dır. Särndal, Swenson ve Wretman(1992),

$$V(\hat{Y}_{HT}) = \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N (\pi_{ij} - \pi_i \pi_j) \frac{y_i}{\pi_i} \frac{y_j}{\pi_j} \quad (2.5)$$

eşitliğiyle vermiştir.

$\hat{Y}_{HT}$  'nin varyansının yansız bir tahmin edicisi,

$$\hat{V}(\hat{Y}_{HT}) = \sum_{i=1}^n \frac{1-\pi_i}{\pi_i^2} y_i^2 + \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j \pi_{ij}} y_i y_j \quad (2.6)$$

olarak belirlenmiştir [Horvitz ve Thompson, 1952]. Särndal, Swenson ve Wretman(1992),

$$\hat{V}(\hat{Y}_{HT}) = \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \left( \frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} \right) \frac{y_i}{\pi_i} \frac{y_j}{\pi_j} \quad (2.7)$$

eşitliğiyle tanımlamıştır.

Sen (1953) ve bağımsız olarak Yates ve Grundy (1953),  $\hat{Y}_{HT}$  'nin varyansının yansız bir tahmin edicisini,

$$\hat{V}(\hat{Y}_{HT})_{SYG} = \frac{1}{2} \sum_{i=1}^n \sum_{\substack{j=1 \\ i \neq j}}^n \frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}} \left( \frac{y_i}{\pi_i} - \frac{y_j}{\pi_j} \right)^2 \quad (2.8)$$

belirlemiştir [Shahbaz ve Hanif, 2003].  $\pi_i$ ,  $i$  inci birimin örneğe seçilme olasılığı (birinci sıra içirme olasılığı) ve  $\pi_{ij}$ ,  $i$  ve  $j$  inci birimlerin birlikte örnekte içerilmesi olasılığı – ikinci sıra içerilmesi olasılığıdır.  $i = j$  olduğunda  $\pi_{ii} = \pi_i$  dir [Särndal ve ark., 1992]. Pek çok araştırmacı HT tahmin edicilerinin varyans ve varyans tahmin edicileri için çeşitli formüller önermiştir [Bkz. Shahbaz, 2003].

#### b) Yığın ortalamasının HT tahmin edicisi

$\bar{Y}$  yığın ortalaması,

$$\bar{Y} = \sum_{i=1}^N y_i / N$$

dır ve yığın ortalamasının yansız bir tahmin edicisi,

$$\bar{y}_{HT} = \frac{1}{N} \sum_{i=1}^n \frac{y_i}{\pi_i} \quad (2.9)$$

$$\bar{y}_{HT} = \frac{1}{N} \hat{Y}_{HT}$$

şeklinde tanımlanır.  $\bar{y}_{HT}$  tahmin edicisinin varyansı,

$$V(\bar{y}_{HT}) = \frac{1}{N^2} V(\hat{Y}_{HT}), \quad (2.10)$$

$\bar{y}_{HT}$  'nin varyansının tahmin edicisi,

$$\hat{V}(\bar{y}_{HT}) = \frac{1}{N^2} \hat{V}(\hat{Y}_{HT}) \quad (2.11)$$

eşitliğinden belirlenir.

Horvitz-Thompson tahmin edicileri yansız (tasarım-yansız) tahmin edicilerdir. Yansızlığı için Tillé [2005] çalışmasına bakılabilir. Bununla birlikte, küçük veya orta çapta örnekler için yüksek hata kareler ortalamasına sahip olabilmektedir [Little, 1981]. Basu (1971), Horvitz-Thompson tahmin edicisinin bazı durumlar için iyi olmayan bir tahmin edici olduğunu fillerle ilgili verdiği meşhur örneğiyle göstermiştir. Bunun üzerine Hájek (1971), Basu'nun çalışmasına dayalı olarak bir ağırlıklandırılmış tahmin edici olan Hájek tahmin edicisini önermiştir.

## 2.4. Hájek Tahmin Edicisi

Basu (1971)'in vermiş olduğu örnek üzerine Hájek (1971),

$$\bar{y}_{Haj} = \frac{\sum_{i=1}^n \frac{y_i}{\pi_i}}{\sum_{i=1}^n \frac{1}{\pi_i}} \quad (2.12)$$

tahmin edicisini önermiştir. Hájek tahmin edicisi iki Horvitz-Thompson tahmin edicisinin oranıdır. Burada,  $\pi_i$ , i inci birimin örneğe seçilmesi olasılığıdır ve  $\pi_i = 1 / F_i$  dir.

a) Yığın ortalamasının Hájek (ağırlıklandırılmış) tahmin edicisi

$$\begin{aligned} \bar{y}_{Haj} &= \frac{\sum_{i=1}^n \frac{y_i}{\pi_i}}{\sum_{i=1}^n \frac{1}{\pi_i}} = \frac{\sum_{i=1}^n \frac{y_i}{\pi_i}}{\hat{N}} \\ \bar{y}_{Haj} &= \frac{\sum_{i=1}^n W_i y_i}{\sum_{i=1}^n W_i} = \bar{y}_w \end{aligned} \quad (2.13)$$

b) Yığın toplamının Hájek (ağırlıklandırılmış) tahmin edicisi

$$\hat{Y}_{Haj} = N \bar{y}_w = \hat{Y}_w \quad (2.14)$$

ile tanımlanır. Burada  $W_i = 1 / \pi_i = F_i$ , örnekleme ağırlığıdır.  $\hat{N} = \sum_{i=1}^n 1 / \pi_i$ , N'in HT tahmin edicisidir.

Literatürde Hájek, Hájek türü ağırlıklandırılmış tahmin ediciler veya sadece ağırlıklandırılmış tahmin ediciler olarak da adlandırılmaktadır [Kalton, 1983; Kish, 1992].  $\bar{y}_{Haj}$  tahmin edicisinin  $\bar{y}_w$  şeklinde yaygın gösterimi kullanılmaktadır.  $\bar{y}_w$  tahmin edicisi yansız (tasarım-yansız) bir tahmin edici değildir ancak Särndal (1980) asimptotik olarak yansız olduğunu göstermiştir.

$\bar{y}_w$  tahmin edicisi doğrusal değildir ve yaklaşık yansızdır ve tam bir varyans eşitliği verilememektedir.  $\bar{Y}$  'yı iki yığın toplamının oranı şeklinde kabul edersek,  $\bar{Y} = Y / X$  ve burada  $x_i = 1$ ,  $i = 1, \dots, N$  için kabul edilirse  $\bar{y}_w$  bir oran tahmin edicisi olur ve  $R = Y / X$  yığın oranı için tanımlanan sonuçlar kullanılabilir.

Taylor linerizasyonu yakınsaması kullanımıyla Hájek tahmin edicisinin yaklaşık varyansı,

$$AV(\bar{y}_w) = \frac{1}{N^2} \sum_{j=1}^N \sum_{\substack{i=1 \\ i \neq j}}^N (\pi_{ij} - \pi_i \pi_j) \left( \frac{y_i - \bar{Y}}{\pi_i} \right) \left( \frac{y_j - \bar{Y}}{\pi_j} \right) \quad (2.15)$$

olarak verilir ve bir varyans tahmin edicisi,

$$\hat{V}(\bar{y}_w) = \frac{1}{\hat{N}^2} \sum_{j=1}^n \sum_{\substack{i=1 \\ i \neq j}}^n \left( \frac{\pi_{ij} - \pi_i \pi_j}{\pi_{ij}} \right) \left( \frac{y_i - \bar{y}_w}{\pi_i} \right) \left( \frac{y_j - \bar{y}_w}{\pi_j} \right) \quad (2.16)$$

eşitliğinden elde edilir [Särndal ve ark., 1992].

HT ve HA tahmin edicisinin her ikisi de kolay uygulanabilir ve istenilen istatistiksel özellikleri taşıyor olmaları nedeniyle araştırma tahminlerindeki kullanımları uzun geçmişe dayanmaktadır [Opsomer, 2008]. Ancak uygulamalı araştırmalarda Hájek tahmin edicilerinin yaygın, Horvitz-Thompson tahmin edicilerin nadir kullanılıyor olduğu görülmektedir [Hulliger, 1999].

Kalton (1983), Kish (1992), Hulliger (1999), Gelman (2007), daha başka yazarlar, ORC Macro (1996), DHS (2005), TMSD (2006) gibi çalışmalar ve SPSS yazılım programı uygulamalı araştırmalar için Hájek tahmin edicilerinin kullanımını önermişlerdir.

Hájek (ağırlıklandırılmış) tahmin edici kullanılması hesaplama kolaylığı açısından daha avantajlıdır. Mesela, ağırlıklandırılmış veriler için standart yazılım programlarına başvurulabilmektedir [Kalton, 1983]. Smith (1991) karmaşık örnekleme tasarımlarında sonradan tabakalama için ağırlıklandırılmış tahmin edici kullanımını önermiştir.

Särndal, Swenson ve Wretman (1992),  $\bar{y}_w$  tahmin edicisinin genellikle  $\bar{y}_{HT}$  tahmin edicisinden daha iyi olduğunu belirtmişlerdir.  $N$  yığın çapı biliniyor olsun veya olmasın araştırmalarda genellikle kullanılacak tahmin edicinin  $\bar{y}_w$  olduğu belirtilmiştir.

Horvitz ve Thompson (1952),  $N$  çaplı yığından alınan  $n$  çaplı örnekte  $\hat{Y} = \sum_{i=1}^n \alpha_i y_i$  yapısındaki tüm tahmin ediciler arasında sadece  $\hat{Y}_{HT} = \sum_{i=1}^n y_i / \pi_i$  Horvitz-Thompson tahmin edicisinin  $Y$  yığın toplamı için yansız bir tahmin edici olduğunu göstermiştir.

HT tahmin edicisi verilen seçim olasılıkları altında yanı azaltıyor olmasına karşın, seçim olasılıklarının iyi belirlenemediği durumlarda iyi olmayan tahminlere yol açabilmektedir. Buna karşın, HT tahmin edicisinin faydalı olduğu ispatlanmıştır, örneğin  $R$  oranının tahminindeki kullanımı gibi [Einbeck ve ark., 2007].

Särndal, Swenson ve Wretman (1992); i)  $y_i$  görel olarak homojen olduğunda, ii) örnek çapı sabit olmadığında veya iii)  $\pi_i$ 'ler  $y_i$ 'lerle zayıf veya negatif ilişkili olduğu durumlarda, HA tahmin edicisinin görel olarak HT tahmin edicisinden

çoğunlukla daha iyi bir tahmin edici olduğunu göstermişlerdir. [Dorfman ve Valliant, 1997].

Bu üç durum için HA ve HT tahmin edicileri çeşitli örneklerle kıyaslanmıştır. Basu (1971)'in fillerle ilgili örneğiyle ve Tillé (2005)  $y_i = C$  gibi bir sabit sayıya eşit olması– Bernoulli örneklemesi böyle bir tasarıma örnektir - durumunda HT tahmin edicisinin iyi olmayan özellikleri olduğu belirtilmiştir.

Horvitz ve Thompson (1952),  $\pi_i = nY_i / Y$  olduğunda  $\hat{Y}_{HT}$  tahmin edicisinin varyansının sıfır olduğunu ve yapılan örneklemenin en iyi olduğunu belirtmiştir. Rao (1994),  $\pi_i$  seçim olasılıkları  $Y_i$ 'lerle ilişkili olmadığında HT tahmin edicisinin anlamsız sonuçlara yol açabileceğini belirtmiştir [Einbeck ve ark., 2007]. Araştırmalarda  $\pi_i$ 'lerin  $Y_i$ 'lerle ilişkili olması koşullunu sağlamanın genellikle mümkün olamayacağı açıktır.

Little (1981),  $\bar{Y}$  yığın ortalamasının ağırlıklandırılmamış örnek ortalamasını (YM), Horvitz-Thompson (HT), Hájek (HA) ve sonradan tabakalama (ST) tahmin edicilerini kıyaslamıştır. Hata kareler ortalamaları için  $ST < HA < YM < HT$  küçüklük sıralamasının olduğunu belirlemiştir. Little'ın simülasyon sonuçlarından HA tahmin edicisinin HT tahmin edicisinden daha etkin bir tahmin edici olduğu belirlenmiştir. Yığın ortalamasının tahmin edicileriyle ilgili çıkarımları ise şöyle verilmiştir:

- i) Horvitz-Thompson tahmin edicisi yansızdır ancak sağlam (robust) istatistik değildir. Özellikle çarpık yığınlarda çok yüksek hata kareler ortalamasıyla sonuçlanır.
- ii) Ağırlıklandırılmamış örnek ortalamasının düşük varyansı vardır ancak ağırlıklandırılmamış ortalama tahminleri yan'a neden olur ve yan örnek çapı arttığında artar. Yanlı tahminler örnek araştırmalarında kötü tahmin edici olarak kabul edilir.

- iii) Yığın tabaka çapı bilgisini kullanan tahmin ediciler (ST gibi) kullanmayanlardan daha iyidir. Eğer tabaka çapları biliniyorsa bu bilginin kullanılması tercih edilmelidir.

Burada temel örnekleme yöntemlerinden basit tesadüfi, tabakalı tesadüfi ve iki aşamalı küme örnekleme yöntemlerinde,  $\bar{Y}$  ve  $Y$  'nin Horvitz-Thompson ve Hájek tahmin edicileri ve örnekleme ağırlıkları ile  $R = Y / X$  yığın oranının tahmini incelenecektir. Tabakalı örneklemede yığın oranı için Ayhan (2003)'ün önermiş olduğu bileşik oran tahmin edicisi verilecektir.

## 2.5. Basit Tesadüfi Örnekleme

Basit tesadüfi örnekleme (BTÖ) yönteminde yığındaki her bir birimin örnekte yer alması olasılığı aynıdır. Yığın parametre tahminlerini hesaplamada gözlem değerlerine atanacak ağırlık aynı olduğu için ağırlıklandırma yapılmasına gerek yoktur. Bu yöntemle elde edilecek istatistiksel hesaplamalar ağırlıksız yapıldığı için işlemler kolaylıkla yapılır.

Yerine koymaksızın BTÖ tasarımında,  $N$  çaplı yığındaki  $i$  inci örnekleme biriminin  $n$  çaplı örneğe seçilmesi olasılığı  $\pi_i = n / N$ , tüm örnekleme birimleri için aynıdır ( $i = 1, \dots, N$ ).

Seçim olasılıkları,

$$\text{Birinci sıra seçim olasılığı; } \pi_i = \frac{n}{N} = f \quad i = 1, \dots, N$$

$$\text{İkinci sıra seçim olasılığı; } \pi_{ij} = \frac{n}{N} \frac{n-1}{N-1} \quad i \neq j = 1, \dots, N.$$

Örnekleme birimlerinin örneğe seçilme olasılıkları  $\pi_i = n / N = f$  dir.  $W_i$ , örnekleme ağırlıkları seçim olasılıklarının tersine eşittir ve tüm örnek birimleri için  $W_i = 1 / \pi_i = N / n = F$  aynıdır,  $i = 1, \dots, n$ .

### 2.5.1. Yığın ortalaması ve yığın toplamının Horvitz-Thompson tahmin edicisi

Basit tesadüfi örneklemede  $\bar{Y}$  yığın ortalaması,

$$\bar{Y} = \sum_{i=1}^N y_i / N$$

ve  $Y$  yığın toplamı,

$$Y = \sum_{i=1}^N y_i$$

ile tanımlanır.

#### a) Yığın ortalamasının HT tahmin edicisi

Yerine koymaksızın BTÖ tasarımında yığın ortalamasının tahmini,

$$\bar{y} = \frac{\sum_{j=1}^n y_j}{n}$$

dır.  $\bar{y}$ ,  $\bar{Y}$  'nin bir Horvitz-Thompson tahmin edicisidir.  $\bar{y}$  'nin bir HT tahmin edicisi olduğunu aşağıdaki gibi gösterebiliriz.

$$\bar{y}_{HT} = \frac{1}{N} \sum_{i=1}^n \frac{y_i}{\pi_i}$$

$$\bar{y}_{HT} = \frac{1}{N} \sum_{i=1}^n W_i y_i$$

$$= \frac{1}{N} \sum_{i=1}^n y_i \frac{N}{n}$$

$$\bar{y}_{HT} = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y} \quad (2.17)$$

$W_i$ , örnekleme ağırlıklarıdır ve  $W_i = 1/\pi_i = N/n = F$

b) Yığın toplamının HT tahmin edicisi

Yığın toplamının basit tesadüfi örnekleme yöntemiyle tahmini,

$$\hat{Y} = N\bar{y}$$

ile elde edilir.  $\hat{Y}$ ,  $Y$ 'nin Horvitz-Thompson tahmin edicisidir:

$$\hat{Y}_{HT} = \sum_{i=1}^n \frac{y_i}{\pi_i} = \sum_{i=1}^n y_i \frac{N}{n} = \frac{N}{n} \sum_{i=1}^n y_i = N \bar{y}_{HT} = N \bar{y} = \hat{Y} \quad (2.18)$$

BTÖ kendinden ağırlıklı tasarım olduğu için tüm birimlere atanacak ağırlıklar eşittir. Tahmin edicileri ve varyanslarını hesaplamada ağırlıklandırma yapmaya gerek yoktur.

### 2.5.2. Yığın ortalaması ve yığın toplamının Hájek tahmin edicisi

a) Yığın ortalamasının HA tahmin edicisi

Yerine koymaksızın BTÖ tasarımında yığın ortalamasının HA tahmin edicisi,

$$\bar{y}_w = \frac{\sum_{i=1}^n W_i y_i}{\sum_{i=1}^n W_i} \quad W_i = \frac{N}{n} \quad (2.19)$$

$$\bar{y}_w = \frac{\sum_{i=1}^n \frac{N}{n} y_i}{\sum_{i=1}^n \frac{N}{n}} = \frac{\frac{N}{n} \sum_{i=1}^n y_i}{\frac{N}{n} \sum_{i=1}^n 1} = \frac{\sum_{i=1}^n y_i}{n} = \bar{y}$$

ile elde edilir. Burada  $\bar{y}_{HT} = \bar{y}_w = \bar{y}$  olduğu görülmektedir.

#### b) Yığın toplamının HA tahmin edicisi

Yerine koymaksızın BTÖ tasarımında yığın toplamının HA tahmin edicisi,

$$\hat{Y}_w = N \bar{y}_w = \hat{Y} \quad (2.20)$$

ile tanımlanır.

### 2.6. Tabakalı Tesadüfi Örneklem

Yığın birimleri incelenen özellikleri bakımından önemli farklılık gösteriyorsa aynı özellikteki birimler bir grupta (tabakada) toplanacak şekilde yığın homojen tabakalara ayrılabilir. Yığın bölgeler, yerleşim yeri türleri veya gelir grupları gibi tabakalara ayrılmak istenebilir. Tabakalı örneklem yöntemine tabakalar arasında incelenen özellikler bakımından farklılıklar olması durumunda başvurulur. Tabakalardaki birimlerin kendi içinde benzer (homojen), tabakalar arasında ise farklı (heterojenliğin) olması istenir. Tabakalar ilgilenilen değişkenlere göre homojen olduğunda duyarlılıkta artış sağlanır ve tabakalar ilgilenilen değişkenlere veya yanıt olasılıklarına göre homojen olduğunda yan azaltılmış olur.

Tabakalı tesadüfi örneklem (TÖ), örneklem birimlerinin tabakalardan yerine koymaksızın basit tesadüfi örneklem ile seçildiği bir örneklem yöntemidir. Buna göre, tabakadan tabakaya incelenecek özellikleri bakımından önemli farklılıklar gösteren  $N$  çaplı yığın  $N_1, N_2, \dots, N_H$  çaplı tabakalara ayrılır ( $N = N_1 + N_2 + \dots + N_H$ ). Yığını temsil eden bir örnek elde etmek için her tabakadan

sırasıyla  $n_1, n_2, \dots, n_H$  çaplı örnekler basit tesadüfi örnekleme yöntemiyle seçilir. Yığın parametrelerini tahmin etmek için tabakalardan seçilen örnekler tek bir örnekte birleştirilir.  $n_1 + n_2 + \dots + n_H = n$  çaplı örnekten derlenen bilgiler kullanılarak yığının parametreleri için tahminler yapılır.

Tabakalı örnekleme yönteminde her bir tabakadan örnek seçimi küme örnekleme, sistematik örnekleme, çok aşamalı örnekleme veya başka örnekleme yöntemleriyle de yapılabilir. Her bir tabakaya farklı örnekleme yöntemleri uygulanabilir. Tabakalamanın asıl amacı örnekleme tasarımında esneklik sağlamak, yığının farklı kısımlarına ayırmak ve tasarımın istatistiksel etkinliğini arttırmaktır [Verma, 2002].

Yığından tabakalı tesadüfi örnekleme ile seçilen  $n$  çaplı örnekte,  $i$  inci birim  $h$  ıncı tabakaya ait ise  $i$  inci birim seçim olasılığı (birinci sıra seçim olasılığı),

$$\pi_i = \frac{n_h}{N_h} = \pi_h = f_h$$

ile tanımlanabilir.  $f_h$ ,  $h$  ıncı tabaka örnekleme kesridir,  $h = 1, \dots, H$ ,  $i = 1, \dots, n_h$  ve  $i \in h$ . Seçim olasılıkları tabakadan tabakaya farklılık göstermektedir.

Örnekleme ağırlığı seçim olasılığının tersi olarak tanımlanmıştır. Buradan, TÖ de  $i$  inci birim  $h$  ıncı tabakaya ait ise  $i$  inci birim örnekleme ağırlığı,

$$W_i = \frac{1}{\pi_i} = \frac{N_h}{n_h} = W_h = F_h$$

ile tanımlanabilir.  $F_h$ ,  $h$  ıncı tabaka genişletme faktörü ve  $W_h$ ,  $h$  ıncı tabaka örnekleme ağırlığıdır,  $h = 1, \dots, H$ ,  $i = 1, \dots, n_h$  ve  $i \in h$ .  $h$  ıncı tabakadaki tüm örnekleme birimlerinin örnekleme ağırlıkları aynı olup  $W_i = W_h$  ile gösterilir.

Literatürde  $h$  ıncı tabaka ağırlığı için  $W_h = N_h / N$  gösterimi yaygındır. Ancak çalışma kapsamında örnekleme ağırlıkları seçim olasılığının tersi şeklindeki tanımlandığı için  $W_h = N_h / n_h$  gösteriminin kullanımı tercih edilmiştir. Pek çok kaynak ve uluslararası araştırmada  $W_h = N_h / n_h$  yaygın kullanımı vardır.

Tabakalı tesadüfi örneklemede yığın ortalaması,

$$\bar{Y} = \sum_{h=1}^H \sum_{i=1}^{N_h} y_{hi} / N$$

ve yığın toplamı,

$$Y = \sum_{h=1}^H \sum_{i=1}^{N_h} y_{hi} ,$$

olarak tanımlıdır.

### 2.6.1. Yığın ortalaması ve yığın toplamının Horvitz-Thompson tahmin edicisi

Tabakalı tesadüfi örneklemede yığın ortalamasının ve yığın toplamının Horvitz-Thompson tahmin edicisi aşağıdaki gibi belirlenir.

#### a) Yığın ortalamasının HT tahmin edicisi

TÖ tasarımında yığın ortalamasının HT tahmin edicisi,

$$\bar{y}_{HT,tb} = \frac{1}{N} \sum_{h=1}^H \sum_{i=1}^{n_h} \frac{y_{hi}}{\pi_i}$$

$$\bar{y}_{HT,tb} = \frac{1}{N} \sum_{h=1}^H \sum_{i=1}^{n_h} W_i y_{hi}$$

$$W_i = \frac{1}{\pi_i} = \frac{N_h}{n_h} = W_h$$

$$\bar{y}_{HT,tb} = \frac{1}{N} \sum_{h=1}^H \sum_{i=1}^{n_h} \frac{N_h}{n_h} y_{hi} = \bar{y}_{tb}$$

$$\begin{aligned}
&= \frac{1}{N} \sum_{h=1}^H \frac{N_h}{n_h} \sum_{i=1}^{n_h} y_{hi} \\
\bar{y}_{HT, tb} &= \frac{1}{N} \sum_{h=1}^H W_h y_h \quad W_h = \frac{N_h}{n_h}
\end{aligned} \tag{2.21}$$

ile elde edilir.  $N_h$ ,  $h$  inci tabaka çapı,  $n_h$ ,  $h$  inci tabakadan seçilen örnek çapı ve  $y_h = \sum_{i=1}^{n_h} y_{hi}$ ,  $h = 1, \dots, H$ . Örnekleme kuramında  $\bar{y}_{HT, tb} = \bar{y}_{tb}$  gösterimi vardır. [Bkz., Cochran, 1977; Çıngı, 1994; Esin ve ark., 2001].

b) Yığın toplamının HT tahmin edicisi

TÖ tasarımında yığın toplamının HT tahmin edicisi,

$$\begin{aligned}
\hat{Y}_{HT, tb} &= N \bar{y}_{HT, tb} \\
\hat{Y}_{HT, tb} &= \sum_{h=1}^H W_h y_h \quad W_h = \frac{N_h}{n_h}
\end{aligned} \tag{2.22}$$

şeklinde elde edilir.

## 2.6.2. Yığın ortalaması ve yığın toplamının Hájek tahmin edicisi

a) Yığın ortalaması için Hájek tahmin edicisi

Tabakalı örneklemede yığın ortalamasının HA tahmin edicisi,

$$\bar{y}_{w, tb} = \frac{\sum_{h=1}^H \sum_{i=1}^{n_h} W_h y_{hi}}{\sum_{h=1}^H \sum_{i=1}^{n_h} W_h} \quad W_h = \frac{N_h}{n_h} \tag{2.23}$$

$$\begin{aligned}
&= \frac{\sum_{h=1}^H \sum_{i=1}^{n_h} \frac{N_h}{n_h} y_{hi}}{\sum_{h=1}^H \sum_{i=1}^{n_h} \frac{N_h}{n_h}} \\
&= \frac{\sum_{h=1}^H \sum_{i=1}^{n_h} \frac{N_h}{n_h} y_{hi}}{\sum_{h=1}^H \frac{N_h}{n_h} \frac{n_h}{N_h}} = \frac{\sum_{h=1}^H \frac{N_h}{n_h} y_h}{\sum_{h=1}^H \frac{N_h}{N_h}} = \frac{\sum_{h=1}^H \frac{N_h}{n_h} y_h}{N} \\
\bar{y}_{w,tb} &= \frac{1}{N} \sum_{h=1}^H W_h y_h = \bar{y}_{HT,tb} \tag{2.24}
\end{aligned}$$

eşitliğinden bulunabilir. Buradan  $\bar{y}_{w,tb}$  ile  $\bar{y}_{HT,tb}$  eşdeğer olduğu görülmektedir.

#### b) Yığın toplamının Hájek tahmin edicisi

TÖ tasarımımda yığın toplamının HA tahmin edicisi,

$$\hat{Y}_{w,tb} = N \bar{y}_{w,tb} \tag{2.25}$$

ile elde edilebilir.

### 2.7. İki Aşamalı Küme Örneklemesi

Yığını oluşturan birimlerin tam olarak listelenemediği, özellikle ülke çapında yapılan araştırmalarda örneğe alınacak birimlere ulaşmanın güç olduğu, yığın çapının çok büyük ve örnekleme birimlerinin geniş bir coğrafi alana yayılmış olduğu durumlarda küme örnekleme yönteminden yararlanılabilir.

Yığın, yerleşim yerlerine veya coğrafi bölgelere göre belirli sayıda kümelere ayrılıyor olabilir. Küme örneklemesinde kümelerin her biri yığını temsil edecek nitelikte olacak şekilde belirlenmeye çalışılır. Küme içi mümkün olduğunca yığındaki tüm farklı özellikteki birimleri içerecek şekilde heterojen ve kümeler arası

ise homojen yapıda olması istenir. Böylece her bir küme yığını temsil edecek nitelikte olacağından, kümelerden bir kaçının örneğe seçilmesiyle yığının temsili sağlanmış olur.

Küme örneklemesinde yığın  $M$  adet benzer özellikler gösteren kümelere ayrılır. İki aşamalı küme örneklemesinin birinci aşamasında,  $M$  adet kümeden (birincil örnekleme birimlerinden: böl'lerinden) tesadüfi olarak  $m$  tane böl seçimi yapılır. Seçilen  $m$  böl'leri (kümeleri)  $N_1, N_2, \dots, N_m$  genişliklidir. Daha sonra ikinci aşama olarak, ( $N_1, N_2, \dots, N_m$  genişlikli)  $m$  sayıdaki her bir böl'den  $n_1, n_2, \dots, n_m$  tesadüfi örnekleri seçilerek seçilen örnekler  $n = n_1 + n_2 + \dots + n_m$  birleştirilir.  $n$  çaplı örnekten istenilen kitle parametrelerinin tahminleri yapılır. Örnekleme birimleri bu şekilde seçildiğinde, sadece  $M$  küme listesine ve seçilen  $m$  kümesindeki birimlerin listesine ihtiyaç vardır. [Esin ve ark., 2001].

İki aşamalı küme örneklemesiyle tahmin iki aşamada yapıldığı için örnekleme birimlerini ağırlıklandırma da iki aşamada yapılır. Birinci aşamada birincil örnekleme birimleri,  $\frac{m}{M}$  olasılığıyla seçilmiş, ikinci aşamada ise böl'lerden böl'lerin seçimi  $\frac{n_j}{N_j}$  olasılıklarıyla olmuştur. Dolayısıyla,  $N$  çaplı yığındaki herhangi bir örnekleme birimini seçme olasılığı her aşamadaki seçim olasılıkları çarpımına eşittir:

$$\pi_j = \frac{m}{M} \frac{n_j}{N_j}$$

Örnekleme ağırlıkları ise seçim olasılıklarının tersidir ve

$$W_j = \frac{1}{\pi_j} = \frac{M}{m} \frac{N_j}{n_j} \quad (2.26)$$

olarak belirlenir.

İki aşamalı küme örneklemesinde yığında birim başına ortalama ve yığın toplamı sırasıyla aşağıdaki gibi tanımlıdır,

$$\bar{\bar{Y}} = \frac{1}{N} \sum_{j=1}^M \sum_{i=1}^{N_j} y_{ji}$$

$$Y = \sum_{j=1}^M \sum_{i=1}^{N_j} y_{ji}$$

### 2.7.1. Yığın ortalaması ve yığın toplamının Horvitz-Thompson tahmin edicisi

a) Yığın toplamının HT tahmin edicisi

$$\begin{aligned} \hat{Y}_{HT,K\ddot{O}} &= \sum_{j=1}^m \sum_{i=1}^{n_j} \frac{1}{\pi_j} y_{ji} \\ &= \sum_{j=1}^m \sum_{i=1}^{n_j} \left( \frac{1}{\frac{M}{m} \frac{n_j}{N_j}} \right) y_{ji} = \sum_{j=1}^m \sum_{i=1}^{n_j} \frac{M}{m} \frac{N_j}{n_j} y_{ji} \\ \hat{Y}_{HT,K\ddot{O}} &= \sum_{j=1}^m \sum_{i=1}^{n_j} W_j y_{ji} = \hat{Y}_{K\ddot{O}} \quad W_j = \frac{M}{m} \frac{N_j}{n_j} \end{aligned} \quad (2.27)$$

olarak belirlenir.  $y_{ji}$ ,  $j$ . kümedeki  $i$ . birim gözlem değeridir.

b) Yığında birim başına ortalamasının HT tahmin edicisi

İki aşamalı küme örneklemesiyle yığın ortalamasının HT tahmin edicisi,

$$\hat{\bar{\bar{Y}}}_{HT,K\ddot{O}} = \frac{\hat{Y}_{K\ddot{O}}}{N} = \hat{\bar{\bar{Y}}}_{K\ddot{O}} \quad (2.28)$$

hesaplanır.

### 2.7.2. Yığın ortalaması ve yığın toplamının Hájek tahmin edicisi

#### a) Yığında birim başına düşen ortalamanın HA tahmin edicisi

İki aşamalı küme örneklemesinde yığında birim başına düşen yığın ortalaması Hájek tahmin edicisi,

$$\begin{aligned}
 \hat{\bar{Y}}_{w,KÖ} &= \frac{\sum_{j=1}^m \sum_{i=1}^{n_j} W_j y_{ji}}{\sum_{j=1}^m \sum_{i=1}^{n_j} W_j} & W_j &= \frac{M}{m} \frac{N_j}{n_j} \\
 &= \frac{\sum_{j=1}^m \sum_{i=1}^{n_j} \frac{M}{m} \frac{N_j}{n_j} y_{ji}}{\sum_{j=1}^m \sum_{i=1}^{n_j} \frac{M}{m} \frac{N_j}{n_j}} = \frac{\frac{M}{m} \sum_{j=1}^m \frac{N_j}{n_j} \sum_{i=1}^{n_j} y_{ji}}{\frac{M}{m} \sum_{j=1}^m \frac{N_j}{n_j} \sum_{i=1}^{n_j} 1} \\
 &= \frac{\sum_{j=1}^m \frac{N_j}{n_j} y_j}{\sum_{j=1}^m \frac{N_j}{n_j} n_j} \\
 \hat{\bar{Y}}_{w,KÖ} &= \frac{\sum_{j=1}^m \frac{N_j}{n_j} y_j}{\sum_{j=1}^m N_j}
 \end{aligned} \tag{2.29}$$

ile belirlenir. Burada  $y_j = \sum_{i=1}^{n_j} y_{ji}$ ,  $n_j$  çaplı j inci kümedeki örneğe seçilen birimlerin gözlem değerleri toplamıdır.  $\hat{\bar{Y}}_{w,KÖ} \neq \hat{\bar{Y}}_{KÖ}$  olduğu görülmektedir.

#### b) Yığın toplamının HA tahmin edicisi

$$\hat{\bar{Y}}_{HT,KÖ} = \frac{\hat{Y}_{KÖ}}{N} = \hat{\bar{Y}}_{KÖ} . \tag{2.30}$$

Yığın ortalaması ve yığın toplamının HT ve Hájek tahmin edicilerinin formülleri örnekleme birimlerinin tümünden yanıt alınmış olduğu durum ve ilave ağırlıklandırma işlemi yapılmadığı durum için tanımlanmıştır. Yanıtlanmama ve ilave ağırlıklandırma yapılması durumlarında tahmin edicilerin formülleri sonraki bölüm aşamalarında incelenecektir.

## 2.8. İki Değişkenin Birbirine Oranı $R$ 'nin Tahmini

Yardımcı değişkeni tahmin aşamasında kullanmanın amacı daha iyi tahminler elde etmektir. Yüksek düzeyde etkin tahmin stratejileri güçlü yardımcı bilgiyi içerir [Särndal ve ark., 1992].  $N$  çaplı bir yığında örnekleme birimlerinin ilişkili  $x_i$  ve  $y_i$  ile gösterilen iki ölçümü elde edilebiliyorsa, bu değişkenlerden biri yardımıyla ( $x$  yardımcı değişkeniyle) diğerinin ( $y$  ilgilenilen değişkeninin) tahmin edilmesi, tahminin duyarlılığını arttırabilmek için başvurulan bir yöntemdir.

Oran tahmini  $r$ 'yi dikkate almak için çok sayıda neden vardır. Bir tanesi orana ilişkin payda teriminin örnek çapı  $n$  olmayıp bir tesadüfi değişken olması durumudur. Ayrıca birçok durumda örnekleme birimleri temel birimlerden farklıdır [Esin ve ark., 2001].

$R = Y / X$  oranının tahmini iki amaca yönelik kullanılır:

1. İki değişkenin birbirine oranıyla ilgileniyor olabilir. Mesela,  $Y$  yığındaki sigara içen erkek öğrencilerin sayısını gösteriyorken,  $X$  erkek öğrencilerin sayısını gösteriyor olabilir. Bu durumda,  $R$  erkek öğrenciler arasında sigara içenlerin oranıdır.
2. İlgilenilen değişken  $y$  ile yardımcı değişken  $x$  yüksek ilişkili olduğunda,  $y$ 'nin tahmininin doğruluğunda artış sağlamak için oran tahmini kullanılabilir.

Oran tahmin edicisi,  $y$  ile  $x$  arasında yüksek ilişki olduğunda standart hatalarda büyük azalma sağlayarak tahminin doğruluğunu artırır. Oran tahmin edicisi tutarlı ancak yanlı bir tahmin edicidir. Geniş-örneklerde (large-samples) yanlılık ihmal edilebilir. Çoğu uygulamalı araştırmada, yanlı bir tahmin edici olması, örnekleme hatasını azaltmada sağlamış olduğu avantajın yanında oldukça önemsiz kalmaktadır. Örnek çapı yeteri kadar büyük olduğunda, oran tahmin edicisi yaklaşık normal dağılım gösterir ve varyans için geniş-örnek (large-sample) formülleri geçerli olur. Geniş-örnek sonuçları örnek çapı 30'u aştığında ( $n > 30$  olduğunda) kullanılabilir [Cochran, 1977].

İki değişkenin oranı  $R = Y / X$  şeklinde tanımlanır ve örnek üzerinden tahmini,

$$r = \frac{\sum_{i=1}^n (y_i / \pi_i)}{\sum_{i=1}^n (x_i / \pi_i)} \quad (2.31)$$

ile verilir.  $r$  oran tahmin edicisi iki Horvitz-Thompson tahmin edicisi oranı şeklinde belirlenir ve bir Hájek türü tahmin edicidir.

### 2.8.1. Basit tesadüfi örneklemede oran tahmini

Yığından basit tesadüfi örnekleme ile seçilen  $n$  çaplı örnekte, örnek birimlerinin birbiriyle ilişkili olduğu bilinen iki ölçümü  $y_i$  ve  $x_i$  olsun,  $i = 1, \dots, n$ . Buradan, örnek oranı  $r$  hesaplanarak  $R$  yığın oranı tahmin edilmeye çalışılacaktır. BTÖ ile oran tahmininde ve tahminin varyansında ağırlıklandırma kullanılmamaktadır.

Değişkenlerden biri  $y_i$  diğeri  $x_i$  ile gösterildiğinde yığın değişkenler oranı,

$$R = \frac{\sum_{i=1}^N y_i}{\sum_{i=1}^N x_i} = \frac{Y}{X} \quad (2.32)$$

ve örnekten tahmini örnek oranı  $r$  ile gösterilir ve

$$r = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i} = \frac{y}{x} \quad (2.33)$$

olarak elde edilir.  $x$  ve  $y$  değerleri örnekten örneğe değişir ve tesadüfi değişkenlerdir.  $r$  iki tesadüfi değişken oranıdır ve BTÖ ile ağırlıksız olarak hesaplanır.

### 2.8.2. Tabakalı tesadüfi örneklemede oran tahmini

Tabakalı tesadüfi örnekleme yönteminin kullanıldığı durumda  $R$  yığın oranının tahmininde iki yol vardır: Birincisi bileşik oran tahmini, ikincisi ise ayrı oran tahminidir. Araştırmalarda yaygın olarak kullanılan bileşik oran tahmin edicisi incelenecektir.

Tabakalı tesadüfi örneklemede  $Y$  ve  $X$  yığın toplamalarının tahmini yapıldıktan sonra, tahmin edicilerin oran tahmininde kullanılması bileşik oran tahmini olarak adlandırılır.  $R$  yığın oranı için bileşik oran tahmin edicisi,

$$r_b = \frac{\bar{y}_{tb}}{\bar{x}_{tb}} = \frac{\hat{Y}_{tb}}{\hat{X}_{tb}} \quad (2.34)$$

şeklinde iki ortalama veya iki toplam tahmininin oranı olarak tanımlanabilir. Buradan bileşik oran tahmin edicisi,

$$r_b = \frac{\hat{Y}_{tb}}{\hat{X}_{tb}} = \frac{\sum_{h=1}^H \frac{N_h}{n_h} y_h}{\sum_{h=1}^H \frac{N_h}{n_h} x_h}$$

$$r_b = \frac{\sum_{h=1}^H W_h y_h}{\sum_{h=1}^H W_h x_h} \quad W_h = \frac{N_h}{n_h} = F_h \quad (2.35)$$

elde edilir.  $W_h$ ,  $h$  ıncı tabaka örnekleme ağırlığı ve  $F_h$ ,  $h$  ıncı tabaka genişletme faktörüdür.

Yığın oranın tahmini için formüller örnekleme birimlerinin tümünden yanıt alınmış ve ilave ağırlıklandırmanın yapılmamış olduğu durumlar için tanımlanmıştır. Yanıtlanmama ve ilave ağırlıklandırma durumu ileriki bölümlerde incelenecek olan yığın ortalamasının Hájek türü ağırlıklandırılmış tahmin edicisi hesaplamasına benzer şekilde yapılır.

### 2.8.3. İki aşamalı küme örneklemesinde oran tahmini

Basit küme örneklemesinde iki değişkenin birbirine oranı  $R$ 'nin oran tahmin edicisi,

$$r_{KÖ} = \frac{\hat{Y}_{KÖ}}{\hat{X}_{KÖ}}$$

$$r_{KÖ} = \frac{\sum_{j=1}^m \sum_{i=1}^{n_j} \frac{M}{m} \frac{N_j}{n_j} y_{ji}}{\sum_{j=1}^m \sum_{i=1}^{n_j} \frac{M}{m} \frac{N_j}{n_j} x_{ji}}$$

$$r_{KÖ} = \frac{\sum_{j=1}^m \sum_{i=1}^{n_j} W_j y_{ji}}{\sum_{j=1}^m \sum_{i=1}^{n_j} W_j x_{ji}} \quad W_j = \frac{M}{m} \frac{N_j}{n_j} \quad (2.36)$$

olarak belirlenir.

#### 2.8.4. Ayhan [2003]'ün önerdiği bileşik oran tahmin edicisi

Tabakalı örnekleme, küme örnekleme veya çok aşamalı örnekleme gibi karmaşık örnekleme tasarımlarında  $R = Y / X$  yığın oranı için Ayhan (2003)'ün önerdiği bileşik oran tahmin edicisi incelenebilir. Ayhan (2003), bileşik oran tahmin edicisi için,

$${}_A r_b = \frac{\sum_{h=1}^H {}_A W_h y_h}{\sum_{h=1}^H {}_A W_h x_h} \quad {}_A W_h = \frac{P_0}{P_h} \quad (2.37)$$

formülünü önermiştir.  ${}_A W_h$ , Ayhan [2003] tarafından önerilen h'inci tanım kümesi (veya tabaka) örnekleme ağırlığıdır.  ${}_A W_h$  ağırlıkları şöyle tanımlanmıştır:

$${}_A W_h = P_0 / P_h = \left[ \sum_{h=1}^H x_h / \sum_{h=1}^H (x_h / P_h) \right] / [(X/x) p_h] \quad (2.38)$$

$$P_0 = \sum_{i=1}^I x_i / \sum_{i=1}^I (x_i / P_i)$$

$$P_h = (X/x) p_h, \quad p_h = n_h / N_h$$

${}_A W_h$  : Örnekleme (tasarım) ağırlığı

$P_0$  : Ağırlıklandırılmış ve ağırlıklandırılmamış örnek çaplarının eşit olmalarının sağlanması için hesaplanan bir sabit

$p_h$  : h. tanım kümesindeki birimlerin örneğe seçilmesi olasılığı

Ağırlıkların hesaplanmasına  $x$  yardımcı değişken bilgisi de dahil edilmiştir.  ${}_A W_h$  ağırlıklarında  $P_0$  standartlaştırması yapılarak  $\sum_{h=1}^H {}_A W_h x_h = x$  olması sağlanmıştır.

${}_A r_b$  tahmin edicisinde, örnekleme ağırlıklandırması, yanıt alınamama ve ilave ağırlıklandırma durumlarını inceleyebiliriz. Ağırlıklandırma, eşit olasılıkla seçim yapılmamış olması durumunu, yanıtlanmamayı gidermek ve tahminlerin

doğruluğunda artış sağlayabilmek amacıyla (ayarlama, sonradan tabakalama, vb.) yapılır. Ağırlıklandırma yapısı Ayhan (2003) tarafından Çizelge 2.1’de verilmiştir.

Çizelge 2.1. Tanım kümeleri için ağırlıklandırma

| Tasarım Ağırlıkları<br>Ayrılmış sınıflarda |       |       | ⇒ | Yanıtlanmama<br>Ağırlıkları<br>Ayrılmış sınıflarda |         |         | ⇒ | Sonradan Tabaka<br>Çapraz/Birleşik |       |
|--------------------------------------------|-------|-------|---|----------------------------------------------------|---------|---------|---|------------------------------------|-------|
| $W_1$                                      |       |       |   | $W_1^*$                                            |         |         |   | $T_1$                              |       |
|                                            | $W_i$ |       |   |                                                    | $W_i^*$ |         |   |                                    | $T_i$ |
|                                            |       | $W_I$ |   |                                                    |         | $W_I^*$ |   |                                    | $T_I$ |

Örnekte yanıt alınmayan birimler olması durumunda tasarım ağırlıklarında yanıtlanmama için ağırlıklandırma düzeltilmesi yapılmalıdır. Tasarım ağırlıklarında,

$${}_A W_h^* = R_0 / R_h \quad (2.39)$$

$R_h$ ,  $h$  ıncı tabaka yanıt oranıdır,  $h = 1, \dots, H$ . Tasarım için  $R_0$ , toplam yanıt oranı

$$R_0 = \sum_{h=1}^H ({}_A W_h x_h) / \sum_{h=1}^H ({}_A W_h x_h / R_h) \quad (2.40)$$

şeklinde hesaplanabilir. Burada  $R_0$  örnek çaplarını düzelterek aynı olmalarını sağlar,

$$\sum_{h=1}^H ({}_A W_h {}_A W_h^* x_h) = x \quad [\text{Ayhan, 2003}].$$

Son olarak ise, sonradan tabakalama yapılmak istenilebilir. Sonradan tabakalama için bileşik ağırlıklandırma,

$$T_h = ({}_A W_h {}_A W_h^* X_h) / X \quad (2.41)$$

şeklinde olur. Burada,  $\sum_{h=1}^H ({}_A W_h {}_A W_h^* T_h x_h) = x$  toplam örnek düzeltme işlemidir [Ayhan, 2003].

## 2.9. Yığın Ortalaması Tahmin Edicileri için Genel Varyans Tahmini

Araştırmalarda parametre tahmin edicilerinin varyanslarının belirlenmesi ve güven aralıklarının oluşturulması oldukça önemlidir. Tek değişkene dayalı olan ve doğrusal tahmin ediciler için kullanılan basit varyans formülleri literatürde yaygın olarak verilmektedir. Ancak birden çok sayıda değişkene dayalı olan veya doğrusal olmayan karmaşık tahmin edicilerin (örn., oran veya ayarlama tahmin edicisi) varyanslarının tahmininde daha karmaşık yapılardaki çeşitli varyans tahmin yöntemlerine başvurmak gerekir. Bu varyans tahmin yöntemleri,

- Doğrusallaştırma (Linerizasyon: Taylor-serileri yöntemi),
- Rasgele Grup Yöntemleri (Random Group Methods),
- Dengelenmiş Tekrarlı Yineleme (Balanced Repeated Replication)
- Yeniden Örnekleme Teknikleri (Resampling Techniques)
  - Jacknife
  - Bootstrap

yaklaşımlarıdır [Hardin ve Gedif, 2006].

$\theta = \bar{Y} = \sum_{i=1}^N y_i / N$  parametresi tahmin edilmek istenmektedir. Ancak, Horvitz-Thompson tahmin edicisi (ters-olasılık ağırlıklandırması), Hájek tahmin edicisi, sonradan tabakalama, tarama, ayarlama tahmin edicisi yöntemleri uygulamada kullanılıyor olmalarına karşın, bu yöntemlerle belirlenen tahmin edicilerin varyans tahminlerinin formülasyonlarını vermek oldukça zordur.

Lu ve Gelman (2000, 2003), Horvitz-Thompson tahmin edicisi (ters-olasılık ağırlıklandırması) varsayımı altında, ağırlıklandırmanın çeşitli düzeylerinin yürütüldüğü araştırmalar için yığın ortalaması tahmin edicilerinin varyanslarının tahmini için genel bir varyans eşitliği tanımlamışlardır. Amaç farklı yapıdaki ağırlıklandırma düzeltmelerinin tümünü aynı genel bir gösterimle vermek olmuştur.

Lu ve Gelman (2000, 2003) genel bir yolla tasarım-tabanlı analitik ve Taylor-serileri varyans tahmin edicilerini çıkarsamıştır. Tüm ağırlıklandırma yöntemlerini birlikte dikkate alarak (çoğu gerçek araştırmada bu şekilde olmaktadır) Taylor-serileri yöntemi doğrultusunda genel bir varyans tahmin edicisi tanımlamıştır.

Tez çalışması kapsamında farklı ağırlıklandırma aşamaları ve yöntemleri incelenmiştir. Dolayısıyla,  $\theta$ 'nın tüm ağırlıklandırma yapılarıyla elde edilecek tahmin edicileri için tek bir genel varyans tahmini kullanılması ve Lu ve Gelman (2000, 2003)'ın önerdiği varyans tahmin edicisine başvurulması uygun olacaktır.

İlgilenilen istatistik (parametre tahmin edicisi) doğrusal olmadığı, oran tahmin edicisi veya başka tahmin ediciler olduğu varsayılın. Taylor-serileri yöntemi (Linerizasyon) doğrusal olmayan istatistikten aynı asimptotik varyanslı doğrusal bir istatistik çıkarsaması yapmayı içerir.  $\hat{\theta}$  tahmin edicisi için varyans tahmini,

$$\hat{V}(\hat{\theta}) \approx \hat{V}\left(\sum_{i=1}^n W_i z_i\right) \quad (2.42)$$

olarak tanımlanır.  $z_i$ , yeni değişkeni  $\hat{\theta}$ 'ya dayalı olan ve  $\hat{\theta}$ 'nin “doğrusallaştırılan” değişkeni olarak adlandırılır.  $\hat{\theta} = \frac{\bar{y}}{\bar{x}}$  iken  $z_i = y_i - \hat{\theta} x_i$ ,  $i = 1, \dots, n$  [Osier ve Museux, 2006].

Lu ve Gelman (2000, 2003), ters olasılık ağırlıklandırmasında  $\theta$ 'nın genel bir varyans tahmin edicisini önermiştir. Ağırlıklandırmanın tamamıyla ters olasılık olduğu ve bağımsız örneklemeyle  $i$  birimlerinin yığından  $\pi_i = 1/W_i$  olasılığına orantısal olarak seçildiği doğal olan basit varsayımı yapılmaktadır. Ters olasılık ağırlıklandırmasının varyansını uygun bir şekilde hesaplayabilmek için ağırlıklandırılmış ortalamanın oran yapısı kabul edilmelidir:

$$\hat{\theta} = \frac{\sum_{i=1}^n W_i y_i}{\sum_{i=1}^n W_i} \quad (2.43)$$

Bu eşitliğin paydası 1'dir. Bu sadece  $W_i$  ağırlıkları,  $W_i / \sum_{i=1}^n W_i$  olacak şekilde yeniden belirlendiğinde ancak böyle olabilir,  $i = 1, \dots, n$  [Lu ve Gelman, 2000; Lu ve Gelman, 2003].

Yığın değerleri  $\mu_u = E(W_i y_i)$ ,  $\mu_w = E(W_i)$ ,  $u_i = W_i y_i$  ve  $\theta = \mu_u / \mu_w$ ,  $\bar{u} = \sum_{i=1}^n W_i y_i / n$ ,  $\bar{w} = \sum_{i=1}^n W_i / n$  şeklinde tanımlanırsa, bu durumda  $\theta$ 'nın tahmin edicisi,

$$\hat{\theta} = \frac{\bar{u}}{\bar{w}} = \frac{\mu_u (1 + \varepsilon_{\bar{u}})}{\mu_w (1 + \varepsilon_{\bar{w}})} \quad (2.44)$$

şeklinde yazılabilir,  $\varepsilon_{\bar{u}} \sim N(0, V(\bar{u}) / \mu_u^2)$  ve  $\varepsilon_{\bar{w}} \sim N(0, V(\bar{w}) / \mu_w^2)$  olduğu varsayılmıştır. Taylor genişletmesi,

$$\hat{\theta} = \theta(1 + \varepsilon_{\bar{u}})(1 - \varepsilon_{\bar{w}} + \varepsilon_{\bar{w}}^2 - \dots)$$

eşitliğini sağlar. Buradan varyansı,

$$\begin{aligned} V(\hat{\theta}) &\approx \theta^2 V(\varepsilon_{\bar{u}} - \varepsilon_{\bar{w}}) = \frac{1}{\mu_w^2} V(\bar{u} - \theta \bar{w}) \\ &\approx \frac{1}{\bar{w}^2} V(\bar{u} - \hat{\theta} \bar{w}) = \frac{1}{\left(\sum_{i=1}^n W_i\right)^2} V\left(\sum_{i=1}^n z_i\right) \end{aligned}$$

olarak tahmin edilebilir,  $z_i = u_i - \hat{\theta}w_i$ ,  $i = 1, \dots, n$ . Bağımsız örnekleme varsayımı altında,  $V\left(\sum_{i=1}^n z_i\right)$  varyansı,  $E\left(\sum_{i=1}^n z_i\right) \approx 0$  olduğunda  $\sum_{i=1}^n z_i^2$  ile tahmin edilebilir. Buradan  $\hat{\theta}$ 'nin varyansı,

$$\hat{V}_{HT}(\hat{\theta}) = \sum_{i=1}^n W_i^2 (y_i - \hat{\theta})^2 \quad (2.45)$$

şeklinde tahmin edilebilir, varyans formülünde  $\sum W_i = 1$  dir [Lu ve Gelman, 2003].

## 2.10. Yığın Oranı Tahmin Edicileri için Genel Varyans Tahmini

$\theta = R = Y / X = \sum_{i=1}^N y_i / \sum_{i=1}^N x_i$  yığın oranı parametresi tahmin edilmek istenilmektedir. Yığın oranı tahmini için önerilen farklı  $\hat{\theta}$  tahmin edicileri vardır.

$y$  ve  $x$  değişkenlerinin örnek toplamları  $y = \sum_{i=1}^n y_i$  ve  $x = \sum_{i=1}^n x_i$  için doğrusal bağıntı  $y = bx$  ile verilebilir. Örnek ortalamalarıyla da benzer şekilde  $\bar{y} = b\bar{x}$  tanımlanabilir.

Ağırlıklandırılmış verilerde yığın oranının tahmin edicilerin varyans tahminleri için genel bir formül önerilebilir. Bunun için Taylor-serileri yönteminden yararlanılabilir.

Yığın oranı tahmin edicilerinin varyans tahminleri için genel bir eşitlik belirlemeye çalışalım. Yeni bir değişken  $z_i = y_i - \hat{\theta}x_i$  şeklinde tanımlansın, burada  $\hat{\theta} = r = y/x$  dir. Bu değişkenin örnek ortalaması  $\bar{z} = \bar{y} - \hat{\theta}\bar{x}$  ve yığın toplamı tahmini  $N\bar{z}$  olacaktır. Bu değişkeninin yığın ortalaması  $\bar{Z} = \bar{Y} - R\bar{X} = 0$  dir. Buna göre ağırlıklandırılmış bir örnekleme tasarımında  $\bar{z}_{tb}$  ortalamasının varyans tahmini,

$$\hat{V}(\bar{z}) = \frac{1}{N^2} \sum_{i=1}^n W_i^2 s_{zi}^2 \quad (2.46)$$

$$\hat{V}(\bar{z}) = \frac{1}{N^2} \sum_{i=1}^n W_i^2 [(y_i - \bar{y}) - r(x_i - \bar{x})]^2 \quad (2.47)$$

olur  $((1-f)$  ihmal edilmiştir). Buradan,  $z_i = y_i - \hat{\theta}x_i$  şeklinde tanımlandığında,  $\hat{\theta} = r$  alınır ve  $r$ 'nin varyans tahmini,

$$\hat{V}(r) = \frac{1}{X^2} \sum_{i=1}^n W_i^2 s_{zi}^2 = \frac{1}{X^2} \sum_{i=1}^n W_i^2 [(y_i - \bar{y}) - r(x_i - \bar{x})]^2 \quad (2.48)$$

ile elde edilir.  $X = \sum_{i=1}^N x_i$ ,  $x$  değişkeninin toplam değeridir. Eş. 2.48 varyans tahmini, tabakalı örnekleme, küme örnekleme veya aşamalı örnekleme gibi karmaşık örnekleme tasarımlarında ve ağırlıklandırılmış verilerde yığın oranı tahmin edicilerin varyans tahminleri için genel bir formül olarak önerilebilir.

Tabakalı örnekleme, küme örnekleme veya aşamalı örnekleme gibi karmaşık örnekleme tasarımlarında örnek birimlerine seçim olasılıklarının tersi olan ağırlıklar atanarak yığın parametre tahminleri yapılır. Yanıtlanmama ve ayarlama ağırlıklandırması yapılması durumunda veride ilave ağırlıklar atanır. Ağırlıklandırılmış verilerde yığın oranının tahmin edicilerin varyans tahminleri için genel bir formül önerilebilir.

Tabakalı tesadüfi örnekleme tasarımında yığın oranı tahmin edicisinin varyans tahmini üzerinden Taylor-serileri yöntemine dayalı genel eşitliğini belirlenebilir. Yeni değişken  $z_{hi} = y_{hi} - \hat{\theta}x_{hi}$  şeklinde tanımlanır, burada  $\hat{\theta} = r = y/x$  dir. Bu değişkenin örnek ortalaması  $\bar{z}_{tb} = \bar{y}_{tb} - \hat{\theta}\bar{x}_{tb}$  ve yığın toplamı tahmini  $N \bar{z}_{tb}$  olacaktır. Bu değişkeninin yığın ortalaması  $\bar{Z} = \bar{Y} - R\bar{X} = 0$  dir. Buna göre ağırlıklandırılmış bir örnekleme tasarımında  $\bar{z}_{tb}$  ortalamasının varyans tahmini,

$$\hat{V}(\bar{z}_{tb}) = \frac{1}{N^2} \sum_{h=1}^H W_h^2 s_{zh}^2 \quad W_h = \frac{N_h}{n_h} \quad (2.49)$$

$$\hat{V}(\bar{z}_{tb}) = \frac{1}{N^2} \sum_{h=1}^H W_h^2 \sum_{i=1}^{n_h} [(y_{hi} - \bar{y}_h) - r_{tb}(x_{hi} - \bar{x}_h)]^2 \quad (2.50)$$

olur  $((1 - f_h))$  'lar ihmal edilmiştir). Buradan,  $z_i = y_i - \hat{\theta}x_i$  şeklinde tanımlandığında,  $\hat{\theta} = r_{tb}$  alınır ve  $r_{tb}$  'nin varyans tahmini,

$$\hat{V}(r_{tb}) = \frac{1}{X^2} \sum_{i=1}^n W_h^2 s_{zh}^2 = \frac{1}{X^2} \sum_{h=1}^H W_h^2 \sum_{i=1}^{n_h} [(y_{hi} - \bar{y}_h) - r_{tb}(x_{hi} - \bar{x}_h)]^2 \quad (2.51)$$

ile elde edilir.  $X = \sum_{i=1}^N x_i$ ,  $x$  değişkeninin toplamıdır. Tabakalı tesadüfi örnekleme için önerilen bu oran tahmin edicisinin varyans tahmini diğer örnekleme tasarımları için de temel alınabilir.

### 3. YANITLANMAMA AĞIRLIKLANDIRMASI

Tüm arařtırmalarda, yanıtlanmamanın bazı düzeyleri olan soru yanıtlanmama, birim yanıtlanmama veya her ikisiyle de karřılařılmaktadır. Soru yanıtlanmama sorunu genellikle ikame, birim yanıtlanmama sorunu ise genellikle ağırlıklandırma yöntemleriyle giderilmeye çalışılır. Yanıtlanmama sorunun giderilmesinde başarısız olunması arařtırma tahminlerinin yanlılığında artışa yol açabilir. Bu potansiyel yanı en aza indirmede kullanılan genel bir yaklařım yanıtlayıcıların örnekleme ağırlıklarını düzeltmeyi içerir. Böylece, yanıtlayan birimlerin yanıt alınamayan birimleri temsili sağılanmaya çalışılır [Creel ve Fahimi, 2005].

Örnekleme ağırlıklandırması yapıldıktan sonra, cevap alınamayan birimlerden kaynaklanan yanı gidermek için örnekleme ağırlıklarında düzeltme yapılır. Yanıtlanmama için ağırlıklandırma düzeltmesi, birim yanıtlanmamayı ele almada kullanılan popüler bir yöntemdir. Hem örnekleme çerçevesini hem de seçilenlerden cevap alınamama nedenini açıklamak için yanıtlayıcıların ayrımsal olarak ağırlıklandırılması faydalı olur [Kalton ve Flores-Cervantes, 2003].

Yanıtlanmama ağırlıkları örneğin karakteristiklerinde veya farklı kısımlarında gözlenen cevap oranları arasındaki farklılıkların etkisini azaltmak için veriye uygulanır. Bu ağırlıklar hem yanıtlayan hem de yanıtlanmayan birimler için bilinen karakteristiklere bağılı bir şekilde belirlenebilir. Yanıtlanmama için ağırlıklandırma yapılması, yanıtlanmama oranları örneğin bir bölümünden diğeri bölümüne daha yüksek ve değıřken olduğıunda çok önemli bir hal alır [Verma, 2008]. Veri yanıtlanmama için ağırlıklandırılarak en azından birimlerin çeřitli karakteristiklerinin örnek dağılımlarındaki büyük çarpıklıkları giderilir [Verma, 2002].

Genel strateji hem yanıt alınan hem de yanıt alınamayanlar için mevcut olan yardımcı bilgi veya bilgilere dayalı olarak yanıtlanmayanlara benzer olan yanıtlayıcıları belirlemek, sonrasında da yanıtlayıcıların tasarım ağırlıklarını arttırarak benzer yanıtlanmayanları temsil etmektir. Mevcut yardımcı bilgiyle yanıtlayan ve yanıtlanmayanlar ağırlıklandırma sınıflarına ayrılır ve her bir sınıftaki

yanıtlayanların örnekleme ağırlıkları bir çarpım faktörüyle arttırılır. Böylece yanıtlayanlar yanıt alınamayanları bu sınıflarda temsil etmiş olur [Kalton ve Flores-Cervantes, 2003].

Araştırmada yanıt alınamayan birimlerin olması halinde öncelikli olarak tasarım ağırlıkları cevap oranının tersiyle düzeltilir. Cevap oranları tasarım ağırlıklarıyla çarpılarak ağırlıklar arttırılır. Böylece yanıtlayanların yanıtlamayanları temsil etmesi sağlanır. Ancak bu düzeltmeler yanıtlayanlar ile yanıtlamayanlar arasındaki sistematik farklılıklardan kaynaklanacak yanı gidermez [ORC Macro, 1996]. Yanıtlayanlar ile yanıtlamayanlar arsında farklılık olup olmadığı mutlaka araştırılmalıdır. Dolayısıyla ağırlıklandırma sınıfları (düzeltme gözeleri) oluşturulmasına ihtiyaç vardır.

Ağırlıklandırma yöntemleriyle yanıtlanmamadan kaynaklanan yanı gidermek için yanıtlayanlarla benzer özellik gösteren yanıtlamayanların belirlenmesi gerekir. Bunun için yardımcı değişkenlerle ağırlıklandırma sınıfları oluşturulmasına gereksinim vardır. Ağırlıklandırma sınıfları bir yardımcı değişkene veya birden çok sayıda yardımcı değişkenin çapraz sınıflamasına dayalı oluşturulabilir [ORC Macro, 1996; Vaughn, 2001; Verma, 2008].

Yanıtlamama düzeltmesi için, ağırlıklandırma sınıflarını oluşturmada kullanılacak yardımcı değişkenlere karar verildikten sonra, bu yardımcı değişkenlere göre ağırlıklar geliştirilen ağırlıklandırma yöntemlerinden herhangi biriyle belirlenmelidir. Böylece, bu sınıflarda yanıtlayanların tasarım ağırlıkları yanıtlamama ağırlıklarıyla arttırılarak benzer yanıtlamayanlar temsil edilir.

Ağırlıklandırma sınıfları nitel, kategorik veya kategorik hale dönüştürülmüş yardımcı değişkenlere dayalı oluşturulur. Bu durumda yanıtlamama düzeltmeleri için kullanılan en yaygın ağırlıklandırma yöntemleri, ağırlıklandırma sınıf düzeltmesi, sonradan tabakalama, tarama yöntemleridir.

Yanıtlanmama için ağırlıklandırma düzeltmesinde kullanılacak yardımcı değişkenler nicel yapıda da olabilir. Bu durumda genelleştirilmiş regresyon tahmin edicisi (GREG) veya ayarlama tahmin edicisi gibi farklı ağırlıklandırma yöntemlerine başvurulabilir. Bununla birlikte, nicel yapıdaki değişkenler kategorik hale dönüştürülerek ağırlıklandırma sınıfları oluşturulabilir. Ancak değişkenlerdeki bilgi kaybı dikkate alınmalıdır.

Yanıtlamama düzeltmeleri için ağırlıklandırma sınıf düzeltmesi ve sonradan tabakalama yöntemleri yıllardır yaygın bir biçimde kullanılmaktadır. Daha karmaşık düzeltme yöntemleri olan tarama (raking), ayarlama (calibration), genelleştirilmiş regresyon tahmini (GREG), lojistik regresyon modellemesi gibi yöntemlerin daha fazla yardımcı bilgi içeriyor olabilmesi yeni gelişmelerdir. Bu yöntemler hem yanıtlanmamadan kaynaklanan yanın giderilmesinde, hem de örnek dağılımlarının bilinen yığın dağılımlarına uyumunu sağlamada kullanılır [Kalton ve Flores-Cervantes, 2003].

Bu bölümde öncelikli olarak ağırlıklandırma sınıflarının nasıl oluşturulması gerektiği üzerinde durulacaktır. Daha sonra, BTÖ, TÖ ve iki aşamalı KÖ tasarımlarında yanıtlamama düzeltmesi için ağırlıklandırma sınıf düzeltmesi yönteminin uygulanması incelenecektir. *(Yanıtlanmama ağırlıklandırması kapsamında ağırlıkların genel gösteriminde \* üst indis kullanımı tercih edilmiştir.)*

Sonradan tabakalama, tarama, GREG ve ayarlama tahmin edicisi yöntemleri ayarlama ağırlıklandırması bölümü başlığı altında incelenecektir. Ancak yöntemlerin yanıtlamama düzeltmelerinde kullanılması da yine aynı şekilde olmaktadır.

### **3.1. Ağırlıklandırma Sınıflarının Belirlenmesi**

Hem yanıtlayan hem de yanıtlamayanlar için mevcut yardımcı bilgi (yardımcı değişken) veya bilgilere dayalı olarak yanıt alınamayanlara benzer olan yanıtlayıcılar belirlenerek ağırlıklandırma sınıfları (düzeltme gözeleri) oluşturulur. Ağırlıklandırma sınıflarında yanıtlamama ağırlıkları belirlenir. Yanıtlayanların örnekleme

ağırlıklarına yanıtlamama ağırlıkları çarpan olarak atanarak yanıtlayanların ağırlıkları arttırılır. Böylece, bu sınıflarda yanıtlayanların yanıt alınamayanları temsili sağlanmış olur.

Her bir ağırlıklandırma sınıfındaki yanıtlayanlar ve yanıt alınamayanların karakteristikleri aynıdır varsayımı vardır [Crockett, 2008]. Ağırlıklandırma sınıflarında belirli sayıda yanıtlayıcıların (mesela en az 20 kişinin) olması istenir ve göze içerisinde yanıtlayanların yanıtlamayanlara belirli bir oranın sağlanması beklenir [Carlson ve Williams, 2001]. Eğer ki herhangi bir sınıf için yanıtlamayan birim yoksa örnekleme ağırlığında değişme olmaz.

Ağırlıklandırma sınıfları oluşturmada yanıtlayan ve yanıtlamayanlar için bilinen yardımcı değişkenler olarak en yaygın kullanılan, nüfus sayımı veya yerel kayıtlardan coğrafik bölge, yerleşim bölgesi ve büyüklüğü, yaşadığı semt ve ev koşullarının özellikleri olabilir. Demografik özelliklerden yanıtlayıcıların cinsiyeti, yaşı, medeni durumu; sosyal ekonomik göstergelerden aile mesleği, eğitim düzeyi olabilir. Anketörlerin yanıt alınamamış birimlere ait konut türü, evin türü, sosyal çevre özellikleri kayıtları ve komşulardan edindiği bilgiler olabilir.

Kish ve Anderson (1978) ağırlıklandırma sınıflarını tek bir değişkene dayalı oluşturmaktansa daha çok sayıda değişkene dayalı oluşturma tercih edilir olduğunu, bununla birlikte değişken seçim işleminin subjektif olduğunu belirtmişlerdir [Vaughn, 2001].

Yardımcı değişken yanıt/yanıtlamama düzeyiyle ve ilgilenilen araştırma değişkeniyle ilişkili olmalıdır [Rizzo ve ark., 1994; Little, 1984; Little, 1986]. Kalton ve Flores-Cervantes (2003), yanıtlamama düzeltmesi için yardımcı değişkenlerin seçiminde genel olarak yanıt olasılıklarını kestiriyor olmasına dikkat etmenin önemli olduğunu belirtmişlerdir.

Çok sayıda ağırlıklandırma sınıfı kullanılarak göreceli olarak küçük ve homojen ağırlıklandırma kategorilerinde yanıtlayan ve yanıtlamayanların karakteristiklerinin

aynı olduğu varsayımı yapılabilir. Ancak diğer yandan çok sayıda ağırlıklandırma sınıfı kullanılması uygulamada örnek tahminlerinin varyansını önemli ölçüde arttırarak, büyük ve değişken ağırlıklar elde edilmesine yol açabilir. Dolayısıyla yardımcı değişken seçimi, yanıt oranlarının örneğin farklı sınıflarında ne kadar değişken olduğuna ve bu değişkenliğin birimlerin özellikleriyle ne kadar ilişkili olduğuna bağlıdır [Verma, 2002; Verma, 2008].

Araştırmalarda yardımcı değişkenlerin sayısı çok fazla olduğunda, değişkenlerden hangilerinin seçileceğine ve ağırlıklandırma sınıf sayılarının nasıl azaltılabileceğine karar vermede Little (1984, 1986) iki yöntem önermiştir:

- Yanıt eğilimi tabakalanması,
- Kestirim ortalama tabakalanması.

Yanıt eğilimi tabakalanması (response propensity stratification) ve kestirim ortalama tabakalanması (predicted mean stratification) çok sayıda yardımcı değişken bilgisinin daha az sayıda olmasına ve daha iyi temsil edilecek şekilde sınıflandırılmasına olanak sağlar. Birincisi yanıt eğilimine göre sınıflama yapıyorken ikincisi ilgilenilen değişkeninin kestirimlerine göre sınıflama yapar. Böylece bu sınıflarda yanıt oranı hesaplandıktan sonra, her bir sınıfın örnekleme ağırlığı yanıt oranının tersiyle çarpılarak yanıtlayanların yanıtlamayanları temsili sağlanır.

Yanıtlamayanlarla ilgili bilgi arttığında ve yanıt oranı düştüğünde yanıt eğilimi ve kestirim ortalama tabakalanması modellerini kullanmanın çok daha cazip olacağı belirlenmiştir. Bu yöntemlerle ağırlıklandırma sınıf düzeltmesinde olduğu gibi yanıtlayan ve yanıtlamayanlar için bilinen değişkenlere dayalı sınıflama yapılması kısıtlaması vardır. Ancak sınıfların çapı veya oranı ile değişkenlerin sayısı ve türleri için herhangi bir kısıtlama yoktur [Carlson ve Williams, 2001].

### 3.1.1. Yanıt eğilimi tabakalanması

Yanıt eğilimi tabakalanması yanıtlamamayı düzeltmede kullanılan alternatif ve gittikçe popüler olan bir yöntemdir. Yanıt eğilimi tabakalanması yöntemi, bir konu yanıtlanır mı yoksa yanıtlanmaz mı, şeklindeki yanıt çıktısıyla bir lojistik model kurarak yanıt olasılığının tahminini araştırmaktadır [Vaughn, 2001].

Hem yanıtlayanlar hem de yanıtlamayanlar için bilinen yardımcı değişkenlerle her bir örnekleme biriminin yanıt verme eğilimini (yanıt olasılığını) kestirmede lojistik regresyon modelleri kullanır ve yanıt eğiliminin tersi ağırlıklar için düzeltme faktörleri olur [Siegel ve ark., 2005]. Bu yaklaşım Rosenbaum ve Rubin (1983)'ün eğilim skoru teorisinin bir genişletmesidir [Little, 1986].

Rosenbaum ve Rubin (1968) yardımcı değişkenlerin sayısı arttığında sınıfların sayısı üstel olarak artacağını belirtmişlerdir. Araştırmalarda yanıt eğilimi skoru verilen yardımcı değişkenler altında bir birimin yanıtının koşullu olasılığıdır [Wun ve ark., 2004].

Yanıt eğilimi tabakalanmasıyla yanıtlanmamaya karşı yanıtın olabilirliğini kestirmeye çalışan bir lojistik regresyon modeli düzenlenir [Carlson ve Williams, 2001]. Böylece modelden önemli yardımcı değişkenler kullanılarak ağırlıklandırma sınıfları oluşturulur. Ağırlıklandırma sınıf düzeltmesine benzer olarak modelden belirlenen yanıtlayıcının kestirim yanıt olasılığının tersi yanıtlayıcı için bir düzeltme olarak kullanılır [Rao, 2001].

### 3.1.2. Kestirim ortalama tabakalanması

Little (1986) kestirim ortalama tabakalanması olarak adlandırılan bir başka ağırlıklandırma sınıflaması tanımlamıştır. İlgilenilen değişkenin kestirimini yapabilmek amacıyla  $x$  yardımcı değişkenlerine dayalı bir regresyon modeli kullanılır. Örnekten  $y$  ilgilenilen değişken değerleri hesaplanır ve ağırlıklandırma sınıfları olarak kullanılmak üzere gruplara ayrılır.

Kestirim ortalama tabakalanması  $x$  yardımcı değişkenleri verilmişken  $y$ 'nin dağılımının modellenmesidir. Regresyon modelleri, kayıp veriye neden olan her bir  $y$  ilgilenilen değişkenleri için oluşturulur. Bu regresyonların her biri farklı kestirim ortalama tabakalanmalarını gerektirir. Dolayısıyla da farklı ağırlıklandırma düzeltmelerinin yapılması gerekir [Little, 1986].

Yanıt eğilimine göre tabakalama büyük örnek yanını kontrol altında tutarken kestirim ortalamasına göre tabakalama hem yanı hem de varyansı kontrol altında tutar. Ancak kestirim ortalama tabakası her bir  $y$  ilgilenilen değişkeni için ayrı modellerin kurulmasını ve yanıtlamama düzeltmeleri yapılmasını gerektirir [Little, 1984].

Yanıt eğilimi ve kestirim ortalama tabakalanması üzerine detaylı bilgi Little (1982, 1984), Vartivarian ve Little (2003), Vartivarian (2004), Little ve Vartivarian (2005) çalışmalarından ve literatürdeki çalışmalardan incelenebilir. Tez kapsamında bu yöntemler üzerinde ayrıca durulmayacaktır.

Yanıtlanmama durumu için ağırlıklandırma sınıflarını oluşturmada kullanılan veri genellikle araştırma değişkenleriyle çok zayıf ilişkilidir ve düzeltmenin araştırma değişkenlerinin tümü için aynı anda sağlanması istenir. Dolayısıyla sınıfları belirlemede önem verilecek husus, ağırlıklandırma sınıflarının farklı yanıt oranları gözlenecek şekilde oluşturulmasıdır [Kalton, 1984].

BTÖ, TÖ ve KÖ tasarımlarında, yanıtlanmama ağırlıklandırma düzeltmeleri için yaygın kullanılan ağırlıklandırma sınıf düzeltmesi yöntemi incelenecektir.

### **3.2. Ağırlıklandırma Sınıf Düzeltmesi**

Yanıtlamama düzeltmesi yapmak amacıyla yaygın olarak kullanılan ve yanıt olasılıklarına başvuran yöntemlerden biri ağırlıklandırma sınıf düzeltmesidir. Bu yöntem, örneği ağırlıklandırma sınıflarına (düzeltme gözelerine) ayıran ve her bir

sınıfa yanıt oranının tersini ağırlık olarak atayan bir yanıtlamama düzeltmesidir. Ağırlıklandırma sınıfları oluşturulduktan sonra her sınıftaki yanıtlayıcıların örnekleme ağırlıkları o sınıfın yanıt oranının tersi ile çarpılır. Böylece o sınıftaki yanıt alınamayanların oranı telafi edilmiş ve yanıtlamama düzeltmesi yapılmış olur. Bu yöntemle ağırlıklandırma sınıfları, ilgilenilen değişkene göre mümkün olduğunca sınıf içerisinde benzer olacak ve yanıt oranları sınıftan sınıfa farklılık gösterecek şekilde belirlenmelidir.

Ağırlıklandırma sınıf düzeltmesi yanıtlamama için kullanıldığında yöntem, yanıtlanmama için göze ağırlıklandırması yöntemi olur. Yanıtlanmama için göze ağırlıklandırmasındaki temel varsayım bir sınıftaki (gözedeki) yanıtlayanlar o sınıftaki yanıtlamayanları temsil ediyor olmasıdır. Yanıt olasılık modeli kullanarak bir sınıftaki örnekleme birimleri aynı olasılıkla yanıt verme durumuna sahip oluyorsa bu varsayım kabul edilebilirdir [Kalton ve Maligalig, 1991].

Bu, Little ve Rubin (2002)'nin rasgelelikte kayıp (missing at random) varsayımına karşılık gelmektedir [Kalton ve Flores-Cervantes, 2003]. Göze ağırlıklandırması yıllarca yanıtlamamayı gidermek amacıyla kullanılmıştır ve yöntem üzerine geniş bir literatür vardır [bkz., örneğin, Oh ve Scheuren, 1983; Kalton, 1983; Little, 1986; Kalton ve Flores-Cervantes, 2003].

Kalton (1983), ağırlıklandırma sınıflarında, sınıf içi homojen olacak biçimde seçildiğinde yani sınıf içerisinde yığın ortalaması  $\bar{Y}_{YANIT} = \bar{Y}_{YANITSIZ}$  olduğunda, ağırlıklandırma sınıf düzeltmesi tahmin edicisi yanının sıfıra gideceğini belirtmiştir.

Ağırlıklandırma sınıf düzeltmesi yöntemi bir veya daha fazla sayıda yardımcı değişkene dayalı incelenebilir. Çalışmada bir yardımcı değişken kullanılması durumuyla incelenecektir. Yöntemin daha fazla sayıdaki yardımcı değişken kullanımı içinde benzer tanımlamaları söz konusudur.

C düzeye sahip olduğu bilinen bir yardımcı değişkene göre örnek birimleri ağırlıklandırma sınıflarına ayrılıyor olsun. Örnek C sayıda  $n_c$  çaplı düzeltme gözelerine ayrılır ve her bir gözede  $n_{cR}$  yanıtlayan örnek birimleri vardır,  $c = 1, \dots, C$ . Ağırlıklandırma sınıf düzeltmesi yönteminin uygulanmasında, tüm  $c = 1, \dots, C$  'ler için  $n_{cR} > 0$  ve  $n_{cR} < n_c$  olduğu kabul edilmiştir. Ağırlıklandırma sınıf düzeltmesi yöntemiyle c inci sınıf yanıtlamama ağırlığı,

$$W_c^* = \frac{n_c}{n_{cR}} \quad (3.1)$$

olarak tanımlanır.  $W_c^*$ , c inci sınıf yanıtlamama ağırlığıdır ve c inci sınıf yanıt oranının tersine eşittir.

Ağırlıklandırma sınıf düzeltmesi yöntemi sırasıyla BTÖ, TÖ ve iki aşamalı KÖ tasarımlarında incelenecektir.

### 3.3. Basit Tesadüfi Örneklemde Yanıtlanmama Ağırlıklandırması

N çaplı sonlu yığından BTÖ ile seçilen  $n$  çaplı örnekteki birimlerin bazılarında yanıt alınamamış,  $n_R$  sayıda birimden yanıt alınmış olsun ( $n_R < n$ ). Yanıtlanmama durumunda yanıtlanmamadan kaynaklanabilecek yanıtın giderilmesi için önerilen ağırlıklandırma sınıf düzeltmesi yöntemi incelenecektir.

Yığından BTÖ ile seçilen örnekleme birimlerinden yanıt alınamayan birimler olması durumunda, yanıtlanmamadan kaynaklanan bilgi kaybını gidermek amacıyla ağırlıklandırma yapılacağına karar verilmiş olsun. Yanıtlanmama düzeltmesinde hem yanıtlayan hem de yanıtlamayanlar için bilinen bir veya birden çok sayıda yardımcı değişkene göre örnek ağırlıklandırma sınıflarına ayrılır.

Yanıtlayıcılar ağırlıklandırma sınıflarına ayrıldıktan sonra yanıtlayıcıların örnekleme ağırlıklarına, hesaplanan yanıtlamama ağırlıkları çarpan olarak atanır. Her bir

sınıftaki yanıtlamama ağırlıkları yanıtlayıcıların gözlem değerlerine atanarak yanıtlanmama düzeltmesi yapılmış olur.

Burada bir yardımcı değişkene dayalı ağırlıklandırma sınıflarının oluşturulması ve yanıtlamama ağırlıklarının belirlenmesi incelenmiştir. Daha fazla sayıdaki yardımcı değişken içinde benzer tanımlamalar ve hesaplamalar söz konusudur. Bununla birlikte hiç ağırlık yapılmaksızın yalnızca yanıtlayanlara dayalı elde edilen yığın ortalamasının tahmin edicisi olan yanıtlayıcı ortalaması yöntemine de yer verilmiştir.

Basit tesadüfi örneklemede tüm birimlerden yanıt alınmış olması durumunda yığın ortalamasının HT tahmin edicisi,

$$\bar{y}_{HT} = \frac{1}{N} \sum_{i=1}^n W_i y_i \quad W_i = \frac{N}{n}$$

eşitliği tanımlıdır. Ancak, örnekte yanıt alınmayan birimler olması durumunda öncelikli olarak  $W_i$  örnekleme ağırlıkları  $n_R/n$  yanıt oranının tersiyle çarpılarak yanıtlamama düzeltmesi yapılarak örnekleme ağırlıkları arttırılır.

Yanıtlamama düzeltmesi yapılmış örnekleme ağırlıkları,

$$W'_i = W_i \frac{n}{n_R} = \frac{N}{n} \frac{n}{n_R} = \frac{N}{n_R} \quad (3.2)$$

ile tanımlanır.

### 3.3.1. Yanıtlayıcı ortalaması

Yığından tesadüfi olarak alınan  $n$  çaplı örnekte  $n_R$  sayıda yanıtlayan birim olduğunda yığın ortalamasının yalnızca yanıtlayan birimlere dayalı HT tahmin edicisi,

$$\bar{y}_{HT} = \frac{1}{N} \sum_{i=1}^{n_R} W'_i y_i \quad W'_i = \frac{N}{n_R} \quad (3.3)$$

$$= \frac{1}{N} \sum_{i=1}^{n_R} \frac{N}{n_R} y_i$$

$$\bar{y}_{HT} = \frac{1}{n_R} \sum_{i=1}^{n_R} y_i \quad (3.4)$$

olur.  $W'_i$ , düzeltilmiş örnekleme ağırlıklarıdır. Bu, yalnızca yanıtlayanlar üzerinden derlenen veriden düzeltilme yapılmaksızın elde edilen yığın ortalamasının bir tahmin edicisidir. Böylece yanıtlayanlar yanıtlamayanları temsil etmiş olur. Ancak bu düzeltmeler yanıtlayanlar ile yanıtlamayanlar arasındaki sistematik farklılıklardan kaynaklanacak yanı azaltmaz [ORC Macro, 1996]. Dolayısıyla ağırlıklandırma sınıfları oluşturulmasına ihtiyaç vardır.

Yanıtlanmamadan kaynaklanan yanı gidermek amacıyla yanıtlamama için ağırlıklandırma yöntemlerinden ağırlıklandırma sınıf düzeltmesi kullanılabilir.

### 3.3.2. Ağırlıklandırma sınıf düzeltmesi

Basit tesadüfi örneklemede ağırlıklandırma sınıf düzeltmesi yönteminin uygulanmasına öncelikli olarak, C düzeye sahip olduğu bilinen bir yardımcı değişkene göre yanıtlayanlar örneği C sayıda ağırlıklandırma sınıfına ayrılır. Ağırlıklandırma sınıf düzeltmesi yöntemiyle c inci sınıf yanıtlamama ağırlığı,

$$W_c^* = \frac{n_c}{n_{cR}} \quad (3.5)$$

olarak belirlenir,  $c=1,...,C$  ve c inci sınıf yanı oranının tersine eşittir [Oh ve Scheuren, 1983; Little, 1986]. c inci sınıftaki yanıtlayanların gözlem değerlerine  $W_c^*$  ağırlıkları çarpan olarak eklenir.

Ağırlıklandırma sınıf düzeltmesi yöntemiyle  $\bar{Y}$  yığın ortalamasının HT tahmin edicisi,

$$\bar{y}_a = \frac{1}{N} \sum_{c=1}^C \sum_{i=1}^{n_R} W_c^* W_i' y_i \quad W_i' = \frac{N}{n_R} \quad W_c^* = \frac{n_c}{n_{cR}} \quad (3.6)$$

$$\begin{aligned} \bar{y}_a &= \frac{1}{N} \sum_{c=1}^C \sum_{i=1}^{n_R} \frac{n_c}{n_{cR}} \frac{N}{n_R} y_{ci} \\ \bar{y}_a &= \sum_{c=1}^C \frac{n_c}{n_{cR}} \frac{1}{n_R} \sum_{i=1}^{n_R} y_{ci} \\ \bar{y}_a &= \sum_{c=1}^C \frac{n_c}{n_R} \bar{y}_{cR} \end{aligned} \quad (3.7)$$

olarak elde edilir. Burada  $\bar{y}_{cR} = \sum_{i=1}^{n_{cR}} y_{ci} / n_{cR}$ , c inci ağırlıklandırma sınıfının yanıtlayanlar örnek ortalamasıdır.  $n_{cR}$ , c inci ağırlıklandırma sınıfındaki yanıtlayan örnek birimlerinin sayısıdır.  $W_c^*$ , c inci ağırlıklandırma sınıfı yanıtlamama ağırlığıdır ve yanıt oranının tersine eşittir.

Ağırlıklandırma sınıf düzeltmesi yöntemiyle  $\bar{Y}$  yığın ortalamasının HA tahmin edicisi,

$$\bar{y}_{w,a} = \frac{\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* W_i' y_{ci}}{\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* W_i'} \quad W_i' = \frac{N}{n_R} \quad W_c^* = \frac{n_c}{n_{cR}} \quad (3.8)$$

eşitliğinden belirlenir.

$Y$  yığın toplamının ağırlıklandırma sınıf düzeltmesi yöntemiyle sırasıyla HT ve HA tahmin edicisi,  $\hat{Y}_a = N \bar{y}_a$  ve  $\hat{Y}_{w,a} = N \bar{y}_{w,a}$  ile tanımlıdır.

### 3.4. Tabakalı Tesadüfi Örneklemde Yanıtlanmama Ağırlıklandırması

Tabakalı tesadüfi örneklem yönteminde, incelenecek özellikleri bakımından önemli farklılıklar gösteren  $N$  çaplı sonlu yığın  $H$  sayıda homojen  $N_1, N_2, \dots, N_H$  çaplı tabakalara ayrılır,  $N = N_1 + N_2 + \dots + N_H$ . Her tabakadan sırasıyla  $n_1, n_2, \dots, n_H$  çaplı örnekler basit tesadüfi örneklemle alınır ve yığından istenilen özelliklerin tahmini  $n = n_1 + n_2 + \dots + n_H$  çaplı örnek üzerinden yapılmaya çalışılır. Ancak yanıtlamayan örnek birimleri olduğunda tabakalardaki yanıtlayan örnek birimlerin sayısı  $n_{1R}, n_{2R}, \dots, n_{HR}$  olur ( $n_{hR} < n_h, h = 1, \dots, H$ ).  $n$  örnek çapının  $n_R$ 'ye düştüğü gözlenir ( $n_R < n$  ve  $n_R = n_{1R} + n_{2R} + \dots + n_{HR}$ ).

#### 3.4.1. Tabakalarda yanıtlamama düzeltmesi

Tabakalı örneklemle seçilen örnekte yanıtlamayan birimlerin olması durumunda yığın ortalaması ve toplamı tahmin edilmek istenilmektedir. Bunun için yanıtlayanlar örneğine yanıtlanmamadan kaynaklanabilecek yanı gidermek amacıyla ağırlıklandırma uygulanacaktır. Örneklem ağırlıklarına yanıtlamama ağırlıklarının yansıtılması gereklidir.

Tabakalı tesadüfi örneklemle seçilen örnekte yanıtlamayanlar için ağırlıklandırma düzeltmesi tabakalarda yapılabilir. Buradan, tabakalarda yanıtlamayan birimleri için tabaka örnek oranının tersi ( $n_h / n_{hR}$ ) örneklem ağırlıklarına çarpan olarak atanır.

Yığın ortalamasının HT tahmin edicisi,

$$\begin{aligned} \bar{y}_{ib} &= \frac{1}{N} \sum_{h=1}^H \frac{N_h}{n_h} \frac{n_h}{n_{hR}} \bar{y}_{hR} & W_h &= \frac{N_h}{n_h} \\ \bar{y}_{ib} &= \frac{1}{N} \sum_{h=1}^H \frac{N_h}{n_{hR}} \bar{y}_{hR} = \frac{1}{N} \sum_{h=1}^H W'_h \bar{y}_{hR} & W'_h &= \frac{N_h}{n_{hR}} \end{aligned} \quad (3.9)$$

şeklinde hesaplanır.  $W'_h$ , h'nci tabaka yanıtlamama düzeltmesi yapılmış örnekleme ağırlığıdır  $\bar{y}_{hR} = \sum_{i=1}^{n_{hR}} y_{hi} / n_{hR}$ , h'nci tabaka yanıtlayanların ortalamasıdır.

Yığının toplamı tahmin edicisi ise,

$$\hat{Y}_{tb} = N \bar{y}_{tb}$$

şeklinde hesaplanır.

Tabaka örnekleme ağırlıklarında yanıtlamama düzeltme yapılmasıyla  $\bar{Y}$  yığın ortalaması ve  $Y$  yığın toplamının HA tahmin edicisi,

$$\bar{y}_{w,tb} = \frac{\sum_{h=1}^H \sum_{i=1}^{n_{hR}} W'_h y_{hi}}{\sum_{h=1}^H \sum_{i=1}^{n_{hR}} W'_h} \quad W'_h = \frac{N_h}{n_{hR}} \quad (3.10)$$

$$\hat{Y}_{w,tb} = N \bar{y}_{w,tb}$$

formülüyle elde edilir.

Bu yanıtlamama düzeltmesi örnekleme ağırlıklarında artış sağlayarak yanıtlayanların yanıtlamayanları temsil etmesini sağlar.

Tabakalı tesadüfi örnekleme uygulanmış veride yanıtlayan ve yanıtlamayanlar için mevcut  $C$  düzeyli bir yardımcı değişkene göre ağırlıklandırma sınıfları oluşturulmasına karar verilmiş olsun. Ağırlıklandırma sınıf düzeltmesi yöntemiyle yanıtlanmama için ağırlıklandırma düzeltmesi yapılıyor olsun.

### 3.4.2. Ağırlıklandırma sınıf düzeltmesi

Ağırlıklandırma sınıf düzeltmesi yönteminde öncelikli olarak,  $H$  sayıda tabakalara ayrılmış örnek  $C$  sayıda ağırlıklandırma sınıflarına ayrılır.  $C$  ağırlıklandırma sınıflarına ayrılan tabakalar için  $H \times C$  çapraz sınıflaması yapılmış olur.  $h$  ıncı tabaka  $c$  inci düzeltme gözesinde  $n_{chR}$  sayıda yanıtlayan örnek birimleri olduğu gözleniyor olsun,  $c = 1, \dots, C$ ,  $h = 1, \dots, H$ .

Ağırlıklandırma sınıf düzeltmesi yöntemiyle yanıtlanmama düzeltmesi yapılacaktır.  $hc$  tabaka-ağırlıklandırma sınıfındaki gözlem değerlerine  $h$  ıncı tabaka örnekleme ağırlığı ve  $c$  inci ağırlıklandırma sınıfı yanıtlanmama ağırlığı çarpan olarak atanır.

Tabaka örneklemede yanıtlanmama için ağırlıklandırma sınıf düzeltmesi yöntemi uygulanmasıyla  $\bar{Y}$  yığın ortalamasının HT tahmin edicisi,

$$\begin{aligned} \bar{y}_{tb,a} &= \frac{1}{N} \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W'_h y_{chi} & W'_h &= \frac{N_h}{n_{hR}} & W_c^* &= \frac{n_c}{n_{cR}} \\ &= \frac{1}{N} \sum_{c=1}^C \sum_{h=1}^H W_c^* W'_h y_{chR} \end{aligned} \quad (3.11)$$

ve Hájek türü ağırlıklandırılmış tahmin edicisi,

$$\bar{y}_{w,tb,a} = \frac{\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W'_h y_{chi}}{\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W'_h} \quad W'_h = \frac{N_h}{n_{hR}} \quad W_c^* = \frac{n_c}{n_{cR}} \quad (3.12)$$

şeklinde elde edilir.  $W'_h$ ,  $h$  ıncı tabaka yanıtlanmama düzeltilmesi yapılmış örnekleme ağırlığı,  $W_c^*$ ,  $c$  inci ağırlıklandırma sınıfı yanıtlanmama ağırlığıdır.

### 3.5. İki Aşamalı Küme Örneklemesinde Yanıtlanmama Ağırlıklandırması

Küme örnekleme yöntemiyle seçilen örnek için de yanıtlanmama ağırlıklandırması BTÖ ve TÖ yöntemlerindeki gibi yapılır. Yanıtlanmama olması durumunda öncelikle örnekleme ağırlıklarında düzeltme yapılır.  $W_j$ , örnekleme ağırlıklarında  $n_i / n_{iR}$  düzeltmesi yapılarak yanıtlayanların yanıtlamayanları temsili sağlanır.  $W'_j$ , düzeltilmiş örnekleme ağırlıkları

$$W'_j = \frac{M}{m} \frac{N_j}{n_j} \frac{n_j}{n_{jR}} = \frac{M}{m} \frac{N_j}{n_{jR}} \quad (3.13)$$

olur. KÖ yanıtlamama için ağırlıklandırma sınıf düzeltmesi yöntemi uygulanmasıyla  $\bar{Y}$  yığın ortalamasının HT tahmin edicisi,

$$\hat{\bar{Y}}_{HT,KÖ,a} = \frac{1}{N} \sum_{c=1}^C \sum_{j=1}^m \sum_{i=1}^{n_j} W_c^* W'_j y_{ji} \quad W'_j = \frac{M}{m} \frac{N_j}{n_{jR}} \quad W_c^* = \frac{n_c}{n_{cR}} \quad (3.14)$$

ve Hájek türü ağırlıklandırılmış tahmin edicisi,

$$\bar{y}_{w,KÖ,a} = \frac{\sum_{c=1}^C \sum_{j=1}^m \sum_{i=1}^{n_j} W_c^* W'_j y_{ji}}{\sum_{c=1}^C \sum_{j=1}^m \sum_{i=1}^{n_j} W_c^* W'_j} \quad W'_j = \frac{M}{m} \frac{N_j}{n_{jR}} \quad W_c^* = \frac{n_c}{n_{cR}} \quad (3.15)$$

şeklinde elde edilir.  $W'_j$ , j inci küme yanıtlanmama düzeltilmesi yapılmış örnekleme ağırlığı,  $W_c^*$ , c inci ağırlıklandırma sınıfı yanıtlamama ağırlığıdır.

Ağırlıklandırma sınıf düzeltmesi için getirilen bir eleştiri araştırmacı için mevcut bilginin tamamını kullanmakta başarısız olmasıdır. Bir başka deyişle araştırmacının

ilgilenilen deęiřkenle iliřkili yardımcı deęiřkenleri, hatta yanıtla doęrudan iliřkili deęiřkenleri dahil etmekte başarısız olmasıdır [Särndal, 1980].

Yanıtlamayanlar hakkında ne kadar çok bilgiye sahip olursak aęırlıklarda o kadar etkin düzeltme yaparız. Aęırlıklandırma sınıf düzeltmesi aęırlıklandırma sınıfları için nispeten az sayıda deęiřken olduęunda akla uygun bir yöntem olarak gözükmektedir [Carlson ve Williams, 2001]. Yardımcı deęiřken sayısı çok olması durumunda farklı aęırlıklandırma yöntemlerine başvurulması önerilebilir.

Yanıtlanmamayla başa çıkmada aęırlıklandırma yaklaşımları verinin rasgelelikte kaybolduęu varsayımına dayanmaktadır. Bu yanıt eęiliminin sadece yardımcı deęiřkenlere dayalı olduęunu ifade etmektedir. Aęırlıklandırma sınıf düzeltmeleri ve ayarlama aęırlıklandırması yöntemleri bir sınıftaki tüm yanıtlayıcıların aynı yanıt eęilimine sahip olduęunu varsayar. Bu sınıflar bir veya daha fazla sayıda yardımcı deęiřkenle oluşturulabilir [Creel ve Fahimi, 2005].

#### 4. AYARLAMA AĞIRLIKLANDIRMASI

Ayarlama ağırlıklandırması yığının kapsanmaması durumunda ortaya çıkabilecek yanı ve örnek ile yığın arasındaki diğer ayrılıkları gidermek amacıyla araştırma verisine uygulanır. Ayarlama ağırlıklandırması yapılan örnek verisiyle, yığının parametreleri gerçeğe daha yakın tahmin edilmeye çalışılır [Verma, 2007a].

Kapsanmama ve araştırma tahminlerinin duyarlılığını arttırmada kullanılan ağırlıklandırma yöntemleri ayarlama başlığı altında incelenebilir. Yığınla ilgili dış kaynaklara örnek verisinin uyumunu sağlayan ağırlıklandırma düzeltmeleri ayarlama ağırlıklandırması olarak adlandırılır. Veriye kapsanmama ve örnek ile yığın arasındaki diğer ayrılıkları gidermek amacıyla ayarlama ağırlıklarının uygulanması, örnekleme ve yanıtlanmama ağırlıklandırması yapıldıktan sonra uygulanır.

Ağırlıklandırmanın son adımı olarak örnek verisinde ilave ağırlıklandırma yapılarak örnekteki birimlerin dağılımının araştırmanın önemli karakteristik özellikleri (cinsiyet, bölge, yaş, eğitim, ık vb.) bakımından yığın dağılımına uyumunun sağlanması amaçlanır. Bu ağırlıklandırma parametrelerin çok daha iyi tahminlerinin yapılmasını sağlar, yanıtlanmama yanını azaltır ve duyarlılığı artırır.

Bir araştırmayı yürütürken yığını temsil eden bir örnek seçilmiş olması çok önemlidir. Ancak bazen yapılan tüm çabalara rağmen bilinçsizce ya da kasıtlı bir şekilde seçilen örnekte bazı özellikteki birimler daha az, diğerleri ise daha çok temsil edilmiş olabilir. Bu durum kapsanmama sorununu ortaya çıkarabilir. Birimlerin belirli bir karakteristiklerine ilişkin örnekteki dağılım yığındaki dağılımından farklı olabilir.

Örneğin, yığındaki oranı %21 olan bayanların örneğe %30 oranında seçilmiş olması gibi. Bu durum, bayanların gerçekte olduğunun üzerinde temsil edilmesine ve bayanlara yüksek ağırlıklar atanmasına yol açacaktır. Dolayısıyla yanlış tahminler yapılmasına neden olabilecektir. Bu sorunu gidermek amacıyla örnek birimleri cinsiyete göre tabakalanabilir. Bayan ve erkeklere ilişkin yığın hakkında toplam

düzeyinde bilgi varsa, örnek verisine ağırlıklandırma uygulanarak cinsiyete göre örnek dağılımının yığın dağılımına uyumu sağlanabilir.

Ağırlıkları hesaplayabilmek için örnek verisinin kıyaslanabileceği yığınla ilgili yardımcı veriye ihtiyaç vardır. Mesela, Ankara’da oturanlarla ilgili bir araştırma yürütülüyorsa Ankara’da yaşayanların karakteristik özelliklerini gösteren yaşayan nüfusun verisine veya yığın araştırma verisine ihtiyaç vardır. Böylece karakteristik özelliklerin dağılımının yardımcı veriye benzerliğinden emin olmak amacıyla örnek verisi yığın verisiyle kıyaslanır.

Örnek verisi yardımcı değişkenlere ait yığın verisiyle kıyaslanarak karakteristik özelliklerin dağılımının yardımcı veriye benzerliğinden emin olmaya çalışılır. Eğer dağılımlar yeteri kadar yakınsa veride ilave ağırlıklandırma yapılmasına gerek yoktur. Ancak, yüzdelik noktalarında farklılık gösteriyorlarsa ağırlıkları hesaplamak gerekir.

Araştırmalarda ağırlıkların belirlenmesi ilgilenilen değişken  $y_i$ ’ye bağlı değildir, yalnızca yardımcı değişken(ler)e bağlıdır. Dolayısıyla, herhangi bir araştırmada ağırlıkları hesaplamada yalnızca yardımcı değişken bilgisine ihtiyaç vardır. Ağırlıklar örneği dengeler, bu nedenle yardımcı değişkenlerin örnek dağılımı yığın dağılımına uyar [Bethlehem ve Keller, 1987].

Araştırmacılar yardımcı değişkenler için mükemmel tahminler sağlayan ağırlıkların, ilgilenilen değişken içinde iyi bir tahmin sağlayacağına inanmaktadırlar. Ağırlıkların uygulanmasıyla yansız tahminler elde edilecektir denilemez, ama yansızlığa çok yakın olunacağı söylenebilir [Deville ve Särndal, 1992]. Araştırma verisinde yardımcı bilginin etkili kullanımı hem tahmin edilecek yığın değerlerine hem de ilgilenilen değişkeni ile yardımcı değişkenler arasındaki gerçek ilişkiye bağlıdır.

Yardımcı değişken yoksa ayarlama veya yanıtlanmama ağırlıklandırması yapılamaz. Rizzo, Kalton ve Brick (1994), yardımcı değişken seçiminin ağırlıklandırma metodolojisi seçiminden muhtemelen çok daha önemli olduğunu belirtmiştir.

Yardımcı değişken seçiminde çoğunlukla yardımcı değişkenlerin araştırma değişkeniyle ilişkili olması durumu göz önünde bulundurulur. Bunun nedeni, ilgilenilen araştırma değişkenleriyle yakından ilişkili olan dış kaynaklarla kıyaslama yapılmasının, ilgilenilen değişkenlerin parametre tahminlerinin duyarlılığında artış sağlamasıdır. Yanıtlanmama düzeltmeleri için yardımcı değişkenlerin seçiminde ise temel olarak genellikle yanıt olasılıklarını kestiriyor olmaları dikkate alınır. [Kalton ve Flores-Cervantes, 2003].

Kapsanmama ve örnek ile yığın arasındaki diğer ayrılıkları gidermek amacıyla uygulanan ağırlıklandırma yöntemlerinden sonradan tabakalama, tarama, lineer ağırlıklandırma, lojistik regresyon ağırlıklandırması, genelleştirilmiş regresyon tahmin edicisi, ayarlama tahmin edicisi ve daha başka yöntemleri ayarlama ağırlıklandırması başlığı altında incelemek uygun görülmektedir.

Ayarlama, yardımcı değişkenler nitel ya da nicel yapıda olduğunda kullanılabilen genel yapıya sahip bir yöntemdir. Ağırlıklandırma yöntemleri olarak yaygın kullanılan sonradan tabakalama, tarama, lineer ağırlıklandırma, lojistik regresyon ağırlıklandırması, GREG ağırlıklandırmasının birer ayarlama yöntemi olduğu literatürde farklı kaynaklarda gösterilmiştir. Deville, Särndal ve Sautory (1993), Ardilly ve Tillé (2006), Vanderhoeft (2001) kaynaklarına başvurulabilir.

Ayarlama tek başına bir ağırlıklandırma yöntemidir. Bununla birlikte sonradan tabakalama, tarama, lineer ağırlıklandırma, lojistik regresyon ağırlıklandırması, GREG ağırlıklandırmasının ve daha başka yöntemleri de birer ayarlama ağırlıklandırması yöntemidir. Bu yöntemler yanıtlamama için ağırlıklandırma düzeltmesi olarak da kullanılmaktadır. Ağırlıklandırma yöntemlerinin ayarlama ağırlıklandırması olarak sınıflaması aşağıdaki gibi Çizelge 4.1'deki gibi verilebilir.

Çizelge 4.1. Ağırlıklandırma yöntemlerinin ayarlama ağırlıklandırması sınıflaması

Örnek araştırmalarında yardımcı değişken bilgisinin olması halinde, tahminlerin etkinliğini arttırmada ve varyansta geniş çaplı azalmaların olmasını sağlamada bir ağırlıklandırma yöntemi olan ayarlama tahmin edicisinin kullanımı önerilebilir.

Örnekteki birimlerin dağılımının çok daha güvenilir dışsal bilgiye uyumunu sağlamada kullanılan ağırlıklandırma düzeltme türü çoğu zaman sonradan tabakalama olarak adlandırılır [Kalton ve Flores-Cervantes, 2003]. Sonradan tabakalama, yalnızca yardımcı değişkenler nitel ya da kategorik olduğunda kullanılan bir düzeltme türüdür. Sonradan tabakalama yapabilmek için diğer bir yol ise nicel yapıdaki değişkenlerin kategorik hale dönüştürülmesidir. Ancak bu, bilgi kaybına yol açabilir.

Yardımcı değişkenler nicel yapıda olduğunda sonradan tabakalama yerine farklı ağırlıklandırma yöntemlerinin kullanımına başvurmak gerekir. Ayarlama, genelleştirilmiş regresyon (GREG) tahmin edicisi yöntemlerinin kullanımı tercih edilebilir. Dolayısıyla hem nicel, hem de nitel değişkenlerin her ikisini birden ağırlıklandırma aşamasında kullanan ayarlama ağırlıklandırması yöntemi başlığı altında sınıflamanın daha yerinde olacağı gözükmektedir.

Tezde yığının kapsanmaması durumu ve örnek ile yığın arasındaki diğer ayrılıkları gidermek amacıyla yaygın olarak kullanılan ayarlama ağırlıklandırma yöntemlerinden,

- Sonradan tabakalama
- Tarama
- Ayarlama tahmin edicisi
- Genelleştirilmiş regresyon (GREG) tahmin edicisi

yöntemleri incelenecektir. (*Ayarlama ağırlıklandırması kapsamında sonradan tabaka ve tarama ağırlıklarının genel gösteriminde \*\* üst indis kullanımı tercih edilmiştir.*)

#### 4.1. Sonradan Tabakalama

Hedef yığın birimlerinin yaş, cinsiyet, meslek, eğitim ve benzer özelliklere göre farklılıklar göstermesi nedeniyle yığında tabakalama yapılmak istenilebilir. Ancak bu bilgiler örnekleme öncesinde birim düzeyinde mevcut olmayabilir. Böyle bir durumda tabakalı örnekleme uygulamak mümkün olmaz.

Tabakalı örnekleme yöntemi yapılabilmesi için her bir tabakadaki yığın çaplarının ve hangi birimin hangi tabakaya ait olduğunun tabakalama öncesinde biliniyor olması gerekir. Tabakaların kendi içinde homojen, kendi aralarında ise heterojen yapıda olması istenir. Bir birimin hangi tabakaya ait olduğu örnekleme yapıldıktan sonra ancak belirlenebiliyorsa tabaka sonradan tabaka olarak adlandırılır [Guy, 1979].

Tabakalamada kullanılan değişkenlerin toplam düzeyindeki bilgileri dış kaynaklardan, nüfus sayımlarından, yerel kayıtlar, şirket kayıtlarından vb. sağlanabilir. Eğer örnek seçiminden sonra birimler bu değişken veya değişkenlere göre sınıflandırılıyorsa bu durumda bilinen  $N_k$ , tabaka çapları her bir gözede yığın toplamalarının tahminlerini elde etmek için kullanılabilir. Bu tahminler bir araya getirildiğinde toplam yığın tahminini sağlar ve bu işlem sonradan tabakalama (poststratification) olarak bilinir.

Sonradan tabakalama iyi bilinen ve çokça kullanılan bir ağırlıklandırma yöntemidir. Sonradan tabakalamanın (Bazen örnek seçiminden sonra yapılan tabakalama olarak da adlandırılır.) ardında yatan düşünce örnekten derlenen bilgilere göre yığını homojen tabakalara ayırmaktır [Bethlehem ve Kersten, 1985]. Tabakalar ilgili değişkene göre homojen olduğunda doğrulukta artış sağlar. Tabakalar ilgili değişkenlere veya yanıt olasılıklarına göre homojen olduğunda yanı azaltır [Bethlehem ve Keller, 1987].

Tabaka oluşturulurken ilgilenilen değişkenle güçlü ilişkili olan yardımcı değişkenlerin kullanılması tercih edilir. İlişki yeteri kadar güçlü ise araştırılacak özelliğin ağırlıklandırılmış örnek dağılımı yığın dağılımına uyacaktır. Bir tabakadaki

tüm birimlere aynı ağırlıklar verilir. Öyle ki yardımcı değişkenlerin ağırlıklandırılmış örnek dağılımı bu değişkenlerin yığın dağılımına uyar [Bethlehem ve Kersten, 1985].

Holt ve Smith (1979), sonradan tabakalamanın bilinen tabakalama yönteminden büyük ölçüde daha etkili olduğunu göstermiştir. Örneklemeye yapıldıktan sonra tabakalamayla duyarlılıkta kazancı maksimize edebilmek amacıyla değişkenlerin farklı kümeleri için farklı yollarla seçilebildiğinde sonradan tabakalama, örnek seçimi yapılmadan önce tabakalama yapmaktan (tabakalı örneklem yöntemi) muhtemelen daha etkin olacaktır. Duyarlılığı maksimize etmek için tabakalama değişkenleri örneklem yapıldıktan sonra farklı kaynaklardan seçilebilir. Sonradan tabakalama yöntemi, kişilerin yanıtları yaş, cinsiyet, meslek, eğitim, eyalet, ülke, ırk vb. özelliklerinde farklılıklar gösterebileceğinin beklendiği yerlerde araştırmalara kolaylıkla uygulanabilir bulunmuştur [Singh, 2003].

Sonradan tabakalama ağırlıkları veriye, uygulanan ve yürütülen diğer ağırlıklandırma türlerinden sonra atanır. Ağırlıklandırma düzeltmesinin bu aşaması iki amaca yöneliktir: kapsanmamayı gidermek ve araştırma tahminlerinin doğruluğunu güçlendirmektir. Ayrıca yanıtlamamadan kaynaklanan yanı gidermede de sıkça başvurulan bir yöntemdir [Kalton ve Flores-Cervantes, 2003].

Sonradan tabakalama yöntemiyle yığın çeşitli tabakalara ayrılarak her bir kategorideki yanıtların dağılımının tahmin edilmesi, yığın ile örnek arasında bilinen farklılıkların düzeltilmesi amaçlanır [Little, 1993; Gelman ve Carlin, 2000; Lu ve Gelman, 2003].

İki başlıca problem; boş tabaka ve yeterli yığın bilgisi olmaması, sonradan tabakalamanın uygulanmasını zorlaştırmaktadır. Eğer tabakalama birkaç sayıda değişkenin çaprazlanmasıyla elde edilmişse tabaka sayısı çok olabilir dolayısıyla bir veya daha fazla tabakanın hiç gözlemi olmaz.

Sonradan tabakalama ile boş tabakada, hiçbir tahmin yapılamaz ve yığın parametrelerinin gerçek tahminin yapılabilmesi mümkün olamaz. Bu sorun tabakanın

ıkarılmasıyla özömlenebilir. Ancak bu özellikle de büyük aplı örnekler için sıkıntılı ve zaman alıcı bir işlemdir. Bu soruna bir başka özüm ise tabaka oluştururken tüm mevcut deęişkenlerin kullanılmamasıdır. Ancak her bir deęişken yanın azaltılması ve duyarlılıkta önemli bir role sahip ise bu pek de tatmin edici bir özüm yolu olmayacaktır.

İkinci tabakalama problemi yığın bilgisi olup olmadığı ile ilgilidir; sonradan tabakalama uygulaması tabaka aplarının bilinmesini gerektirir. Tabakalama boş tabaka olmamış bir şekilde düzenlenmiş olsa dahi tabakaların yığın apları bilinmezse tabakalama yapılamaz.

Bazen yardımcı bilgi vardır ancak çok yönlü tabakalama yapmaya izin verecek kadar detaylandırılmamıştır. Örnek olarak, cinsiyet, yaş, medeni durum ve memleket ile çok yönlü tabakalama, cinsiyet, yaş, medeni durum ve memleketin düzeylerinin her bir kombinasyonu için yığın apı bilgisini gerektirir.

Her bir kombinasyon için yığın apları bilinmiyorsa, tam bir tabakalama mümkün olmaz. Bethlehem ve Kersten (1985) tam bir tabakalama olmaması durumu için, tüm kombinasyonların oluşturulması yerine tamamlanmamış tabakalama yapılmasını önermiştir.

Çok boyutlu bir tabloyla temsil edilmekte olan ayrıntılı bir ilişki farklı ağırlıklandırma yöntemlerine başvurularak ele alınabilir. Önerilen bir ağırlıklandırma yöntemi tarama veya ayarlama tahmin edicisi yöntemi olabilir.

Sonradan tabakalama yöntemiyle ilgili literatürde basit tesadüfi örnekleme tasarımındaki formülasyon ve teorisi üzerine detaylı bilgi vardır. Ancak daha karmaşık örnekleme tasarımlarındaki formülasyonu ve teorisi üzerine az sayıda başvuru kaynağı vardır.

Bu alışmada, basit tesadüfi örnekleme ve tabakalı örnekleme tasarımlarında sonradan tabakalama yöntemi ele alınacaktır. Yığın ortalaması ve yığın toplamı

parametrelerinin HT ve HA tahmin edicisi yapıları verilecektir. Sonradan tabakalama ağırlıklarının nasıl hesaplandığı ve ağırlıkların gözlem değerlerine atanması incelenecektir.

Sonradan tabakalama için yalnızca bir değişken kullanılması durumu incelenmiştir. Sonuçlar daha fazla sayıdaki tabakalama değişkeni olması durumuna genelleştirilebilmektedir.

$U$ , yığını  $N$  sayıda  $u_1, u_2, \dots, u_N$  birimlerinden oluşuyor olsun.  $y$  ilgilenilen değişken ve  $x$  yardımcı değişken olsun.  $N$  çaplı sonlu bir yığından belirlenen bir örnekleme yöntemiyle yerine koymaksızın alınan  $n$  çaplı örnek,  $x$  yardımcı değişkenine göre sonradan tabakalara ayrılıyor olsun.

$x$ , sonradan tabakalama değişkeni  $K$  düzeyli olsun. Yığın  $K$  sayıda sonradan tabakadan oluşur, yığın için  $U = \bigcup_{k=1}^K u_k$  gösterimi yazılabilir.  $s$ , örneğinde sonradan tabakalama yapıldığı için örnek  $(s_1, \dots, s_K)$  sonradan tabakalarına ayrılır,  $k=1, \dots, K$ .

$N_1, N_2, \dots, N_K$  sonradan tabaka çapları dışsal kaynaklardan biliniyordur ve  $n_1, n_2, \dots, n_K$  sonradan tabaka örnek çapları ise örnek seçimi yapıldıktan sonra belirlenebilmektedir. Sonradan tabakalamada  $f_k = n_k / N_k$ ,  $k$  inci sonradan tabaka örnekleme kesridir ve sonradan tabakalama ağırlığı,

$${}_{st}W_k^{**} = \frac{1}{f_k} = \frac{N_k}{n_k} \quad (4.1)$$

ile tanımlanabilir. Ancak literatürde sonradan tabakalama ağırlıkları için  $N_k / N$  gösterimi yaygındır ve formülasyonlar da bu tanımlamaya göre verilmektedir. Burada sonradan tabakalama ağırlıkları için  ${}_{st}W_k^{**} = N_k / n_k$  gösterimi tercih edilmiştir.

#### 4.1.1. Basit tesadüfî örneklemede sonradan tabakalama

Basit tesadüfî örnekleme yöntemiyle seçilen örnek birimleri üzerinden ilgilenilen değişkenler ve yardımcı değişkenlere göre veriler derlenmiş olsun.  $N$  çaplı sonlu bir yığından basit tesadüfî örnekleme yöntemiyle alınan  $n$  çaplı örnek, bir  $x$  yardımcı değişkenine göre sonradan tabakalara ayrılıyor olsun.

BTÖ  $(s, U)$  tasarımında  $K$  düzeyli bir  $x$  yardımcı değişkenine göre sonradan tabakalama yapılmasıyla örnek  $\{(s_1, u_1), \dots, (s_K, u_K)\}$  sonradan tabakalarına ayrılır. Buradan, BTÖ'de yığın ortalaması,

$$\bar{Y} = \sum_{k=1}^K \sum_{i=1}^{N_k} y_{ki} / N$$

ve yığın toplamının,

$$Y = \sum_{k=1}^K \sum_{i=1}^{N_k} y_{ki}$$

sonradan tabakalama tahmin edicisi incelenecektir.

##### a) Yığın ortalamasının sonradan tabakalama tahmin edicisi

$$\bar{y}_{st} = \sum_{k=1}^K \frac{N_k}{N} \bar{y}_k \quad (4.2)$$

$$= \frac{1}{N} \sum_{k=1}^K N_k \frac{y_k}{n_k}$$

$$= \frac{1}{N} \sum_{k=1}^K {}_{st}W_k^{**} y_k$$

$${}_{st}W_k^{**} = \frac{N_k}{n_k}$$

$$\bar{y}_{st} = \frac{1}{N} \sum_{k=1}^K \sum_{i=1}^{n_k} {}_{st}W_k^{**} y_{ki} \quad {}_{st}W_k^{**} = \frac{N_k}{n_k} \quad (4.3)$$

b) Yığın toplamının sonradan tabakalama tahmin edicisi

$$\hat{Y}_{st} = N \bar{y}_{st} = \sum_{k=1}^K \sum_{i=1}^{n_k} {}_{st}W_k^{**} y_{ki} \quad (4.4)$$

şeklinde tanımlıdır.  $\bar{y}_{st}$ , bir Horvitz-Thompson tahmin edicisidir.  $y_{ki}$ , k'nci sonradan tabaka i'nci birim gözlem değeri,  $\bar{y}_k = \sum_{i=1}^{n_k} y_{ki} / n_k$ , k'nci sonradan tabaka örnek ortalaması ve  $y_k = \sum_{i=1}^{n_k} y_{ki}$ , k'nci sonradan tabaka örnek toplamıdır.

Yığın ortalaması ve toplamının sonradan tabakalama tahmin edicisi Hájek ağırlıklandırılmış tahmin edicisi olarak da tanımlanabilir. Buna göre,

a) Yığın ortalamasının sonradan tabakalama HA tahmin edicisi

$$\bar{y}_{w,st} = \frac{\sum_{k=1}^K \sum_{i=1}^{n_k} {}_{st}W_k^{**} y_{ki}}{\sum_{k=1}^K \sum_{i=1}^{n_k} {}_{st}W_k^{**}} \quad {}_{st}W_k^{**} = \frac{N_k}{n_k} \quad (4.5)$$

b) Yığın toplamının sonradan tabakalama HA tahmin edicisi

$$\hat{Y}_{w,st} = N \bar{y}_{w,st} \quad (4.6)$$

ile elde edilir.  $W_k^{**}$ , k'nci sonradan tabaka ağırlığıdır,  $k = 1, \dots, K$ .

Sonradan tabakalama  $Y = \sum_{k=1}^K \sum_{i=1}^{N_k} y_{ki} = \sum_{k=1}^K Y_k$  yığın toplamını verir. Sonradan tabakalama tahmin edicisi  $\hat{Y}_{st} = \sum_{k=1}^K \hat{Y}_k$  yapısındadır ve bu  $Y_k$ 'nin tahminin  $n_k$ 'ya dayalı yapıyor olmasını ve  $n_1, n_2, \dots, n_K$  örnek çapları üzerinden toplam alınmasını içerir.  $\hat{Y}_k$ , k'nci sonradan tabaka toplamı tahmin edicisi  $N_k$ 'ya dayalı belirlenir.

$Y$  yığının toplamının sonradan tabakalama tahmin edicisinin belirlenmesi,  $\pi_i$  seçim olasılıklarının sonradan tabakalar içerisinde sabit olup olmadığına dayalı olarak değişir [Smith, 1991].  $\pi_i = \pi_k$ ,  $i \in N_k$  olduğunda, BTÖ tasarımı olduğu gibi,  $Y$ 'nin sonradan tabakalama tahmin edicisi,

$$\begin{aligned}\hat{Y}_{st} &= \sum_{k=1}^K \hat{Y}_k = \sum_{k=1}^K N_k \hat{\bar{Y}}_k = \sum_{k=1}^K N_k \bar{y}_k = \sum_{k=1}^K \sum_{i=1}^{n_k} \frac{N_k}{n_k} y_{ki} \\ \hat{Y}_{st} &= \sum_{k=1}^K \sum_{i=1}^{n_k} {}_{st}W_k^{**} y_{ki} \quad {}_{st}W_k^{**} = \frac{N_k}{n_k} \\ &= \sum_{k=1}^K \sum_{i=1}^{n_k} \frac{y_{ki}}{f_k} \\ \hat{Y}_{st} &= N \bar{y}_{st} = N \bar{y}_{w,st}\end{aligned}\tag{4.7}$$

şeklinde belirlenir [Zhang, 2000].

#### 4.1.2. Tabakalı tesadüfi örneklemede sonradan tabakalama

Tabakalı örnekleme yöntemiyle  $N$  çaplı yığın kendi içinde homojen  $N_1, N_2, \dots, N_H$  çaplı tabakalara ayrılmış ve her bir tabakadan sırasıyla  $n_1, n_2, \dots, n_H$  çaplı örnekler basit tesadüfi örneklemeyle seçilmiş olsun. TÖ yöntemiyle seçilen örnek üzerinden ölçülen ve ilgilenilen değişkenle ilişkili olduğu düşünülen kategorik yapıdaki bir yardımcı değişkene göre örnek sonradan tabakalara ayrılıyor olsun.

BTÖ yöntemine kıyasla tabakalı örnekleme veya küme örnekleme gibi karmaşık tasarımlarda,  $Y_k$ ,  $k$  ıncı sonradan tabaka yığın toplamının yansız bir tahmin edicisi Horvitz-Thompson tahmin edicisi ile,  $\ddot{Y}_k = \sum_{i \in S_k}^{n_k} y_i / \pi_i$  şeklinde verilebilir. Smith (1991), karmaşık örnekleme tasarımlarında sonradan tabakalama yapılması durumunda yığın parametrelerinin tahmin için Hájek türü bir tahmin edici kullanımını önermiştir.

Smith (1991), karmaşık örnekleme tasarımlarda tabaka içi tahmin edicisi,

$$\hat{Y}_k = N_k \left( \frac{\ddot{Y}_k}{\ddot{N}_k} \right) \quad (4.8)$$

$$= \frac{N_k \ddot{Y}_k}{\sum_{i \in s_k} \frac{1}{\pi_i}}$$

$$\hat{Y}_k = N_k \frac{\left( \sum_{i \in s_k} y_i / \pi_i \right)}{\left( \sum_{i \in s_k} 1 / \pi_i \right)} \quad (4.9)$$

şeklinde elde edileceğini vermiştir. Her bir  $U_k$  tabakası, içerisinde oran tahmin edicisine başvurulmuştur.  $i \in s_k$  için ağırlık  $(N_k / \pi_i) / (\sum_{i \in s_k} 1 / \pi_{ki})$  dır.  $1 / \pi_i$  örnekleme ağırlığıdır. TÖ de  $1 / \pi_i \frac{1}{\pi_i} = 1 / \pi_i = W_h = N_h / n_h$  demektir. Buradan,

$$\hat{Y}_k = N_k \frac{\left( \sum_{h=1}^H \sum_{i=1}^{n_{hk}} y_{hki} / (n_h / N_h) \right)}{\left( \sum_{h=1}^H \sum_{i=1}^{n_{hk}} (n_h / N_h) \right)}$$

$$\hat{Y}_k = N_k \frac{\left( \sum_{h=1}^H \frac{N_h}{n_h} \sum_{i=1}^{n_{hk}} y_{hki} \right)}{\left( \sum_{h=1}^H \sum_{i=1}^{n_{hk}} \frac{N_h}{n_h} \right)} = N_k \frac{\left( \sum_{h=1}^H \frac{N_h}{n_h} y_{hk} \right)}{\left( \sum_{h=1}^H \frac{N_h}{n_h} n_{hk} \right)}$$

$$\hat{Y}_k = N_k \frac{\left( \sum_{h=1}^H W_h y_{hk} \right)}{\left( \sum_{h=1}^H W_h n_{hk} \right)} \quad W_h = \frac{N_h}{n_h} \quad (4.10)$$

olarak tanımlanır [Smith, 1991].

TÖ de yığın toplamının sonradan tabakalama tahmin edicisi,

$$\hat{Y}_{tb,st} = \sum_{k=1}^K \hat{Y}_k \quad (4.11)$$

$$\hat{Y}_{tb,st} = \sum_{k=1}^K N_k \frac{\left( \sum_{h=1}^H W_h y_{hk} \right)}{\left( \sum_{h=1}^H W_h n_{hk} \right)} \quad (4.12)$$

ve TÖ de yığın ortalamasının sonradan tabakalama tahmin edicisi,

$$\bar{y}_{tb,st} = \frac{1}{N} \sum_{k=1}^K \hat{Y}_k \quad (4.13)$$

şeklinde belirlenir [Smith, 1991].

Bununla birlikte, TÖ örneklemede  $Y$  'nın sonradan tabakalama tahmin edicisi için

$$\bar{y}_{tb,st} = \frac{\sum_{k=1}^K \sum_{h=1}^H \sum_{i=1}^{nhk} {}_{st}W_k^{**} W_h y_{hki}}{\sum_{k=1}^K \sum_{h=1}^H \sum_{i=1}^{nhk} {}_{st}W_k^{**} W_h} \quad W_h = \frac{N_h}{n_h} \quad {}_{st}W_k^{**} = \frac{N_k}{n_k} \quad (4.14)$$

$$\bar{y}_{tb,st} = \frac{\sum_{k=1}^K {}_{st}W_k^{**} \sum_{h=1}^H W_h y_{hk}}{\sum_{k=1}^K {}_{st}W_k^{**} \sum_{h=1}^H W_h n_{hk}}$$

eşitliğinin kullanılabileceği önerilebilir.

#### 4.1.3. BTÖ de $\bar{y}_{st}$ tahmin edicisinin yanlılığı

Sonradan tabakalamada teorik bir soruyla karşı karşıya gelinir. Sonradan tabakalama bir tabakalı örnekleme tasarımı gibidir. Bu nedenle, acaba  $\bar{y}_{st}$  tahmin edicisinin özellikleri nasıl ele alınmalıdır? Orijinal örnekleme tasarımıyla basit tesadüfi örnekleme tasarımına mı, yoksa tabakalı örnekleme tasarımına mı başvurulmalıdır?

Eğer  $n$ 'in tüm mümkün değerleri üzerinden ortalama alınırsa basit tesadüfi örneklemede koşullu olmayan bir dağılım elde edilir. Sonradan tabakalama için önerilen, tabakalı örneklemeyle koşullu dağılım olarak, yani  $n' = (n_1, n_2, \dots, n_K)$  örnek tabaka çapları vektörüne koşullu olunacak şekilde başvurulmasıdır. [Holt ve Smith, 1979].

Holt ve Smith (1979), koşullu olmayan dağılımın tasarım aşamasında hangi örnekleme yönteminin ve tahmin stratejilerinin seçileceğine karar vermede rol oynayacağını, seçilen örnekten çıkarsamalar yapmada bir rolü olmadığını belirtmişlerdir. Sonradan tabakalama için BTÖ tasarımı altında koşullu çıkarsama çerçevesini uygulamayı önermişlerdir.

Royall (1971) ve Royall ve Eberhardt (1975), oran ve regresyon tahmin edicilerini kullanırken koşullu çıkarsamaları taraftarı olan hem teorik hem de ampirik altyapıları vermişlerdir. Oran ve regresyon tahmin edicileri için Jagers ve ark. (1985) ve Jagers (1986) geçerli bir çıkarsama vermiştir ve Rao (1985), araştırma örnekleme için koşullu çıkarsamaları dikkate almıştır [Smith, 1991].

$n$  üzerine koşullandırarak ve tabakalı örneklemenin sonuçlarını kullanarak,

$$\begin{aligned} E(\bar{y}_{st} | n) &= \sum_{k=1}^K \frac{N_k}{N} E(\bar{y}_k / n_k) \\ &= \sum_{k=1}^K \frac{N_k}{N} \bar{y}_k = \bar{Y} \end{aligned}$$

eşitliğine ulaşılır.  $\bar{y}_{st}$  koşullu olarak yansızdır.  $\bar{y}_{st}$ ,  $\bar{Y}$  için tasarım-yansız bir tahmin edici olduğu için  $HKO(\bar{y}_{st}) = V(\bar{y}_{st})$  olur.

$\bar{y}_{st}$ , tasarım-yansız bir tahmin edicidir ve ortalaması bilinen orantılı paylaştırılmış tabakalı tesadüfi örneklemedekiyle aynıdır. Dolayısıyla, sonradan tabakalama iyi yapılmışsa BTÖ'de hesaplanan  $\bar{y}$  örnek ortalamasında daha iyi bir tahmin edici olması beklenir [Cho ve Chiang, 2006].

Burada  $\bar{y}$  ağırlıklandırılmamış örnek ortalaması ele alınırsa,

$$\bar{y} = \sum_{k=1}^K \frac{n_k}{n} \bar{y}_k$$

$$E(\bar{y} | n) = \sum_{k=1}^K \frac{n_k}{n} \bar{Y}_k \neq \bar{Y}$$

koşullu olarak yanlı olduğu görülür. Ancak  $\frac{n_k}{n} = \frac{N_k}{N}$  olduğunda yansız bir tahmin edici olur,  $k=1, \dots, K$ . Dolayısıyla  $\bar{y}$ , BTÖ altında yansız bir tahmin edici iken koşullu tabakalı örnekleme yapısı altında yanlı bir tahmin edicidir [Cho ve Chiang, 2006].

#### 4.1.4. BTÖ de $\bar{y}_{st}$ sonradan tabakalama tahmin edicisinin varyansı

Sonradan tabakalama üzerine literatürde, koşullu mu yoksa koşullu olmayan varyansın mı kullanılacağı başlıca bir sorundur. İstatistikçiler arasında istatistiksel çıkarsama hakkında genel görüş birliği yoktur. Geleneksel örnekleme teorisi istatistikçileri arasında dahi koşullu varyansının sonradan tabakalanmış örneklerden çıkarsama için uygun olduğu hakkında görüş birliği yoktur. Hansen ve ark. (1953), Cochran (1963), Kish (1965) ve Des Raj (1972), koşullu olmayan varyans kullanımını savunmuşlardır [Holt ve Smith, 1979].

Yates (1960), koşullu varyansın kullanımını savunmaktadır. Bu iddia, teorik altyapıda elde edilen örnek çapı bir yardımcı istatistik olarak kullanılabileceğini tartışan Durbin (1969) tarafından desteklenmektedir ve böylece çıkarsamalar n'ye koşullu yapılabilir. Cox ve Hinkley (1974), benzer bir iddiayı benimsemiştir. Kalton ve Smith (1976) ile olan tartışmasında geleneksel bir görüş noktasından koşullu çıkarsamanın lehinde tartışmıştır [Smith, 1991].

Holt ve Smith (1979), çalışmalarında sonuçların değişik biçimlerini hem ampirik hem de teorik olarak kıyaslamış ve kesin kez koşullu çıkarsamalar lehine karar vermişlerdir. Bunun anlamı, koşullu olmayan varyansın örnek araştırmalarında hiçbir rolü yok demek olmadığını, yeni araştırmaları planlamada koşullu olmayan varyansın büyük önem taşıdığını belirtilmişlerdir. Koşullu olmayan varyansın örnek seçimi yapılmadan önceki örnekleme stratejilerini kıyaslamada kullanılabileceği ve örnek seçiminden sonraki çıkarsamalar için ise koşullu varyans kullanmanın uygun olduğu savunulmaktadır.

Koşullu dağılım için, verilen  $n$  altında  $\bar{y}_{st}$  'nin varyansı aşağıdaki şekildedir:

$$V(\bar{y}_{st} | n) = \sum_{k=1}^K \left( \frac{N_k}{N} \right)^2 \left( 1 - \frac{n_k}{N_k} \right) \frac{S_k^2}{n_k} \quad S_k^2 = \frac{\sum_{i=1}^{N_k} (y_{ki} - \bar{Y}_k)^2}{N_k - 1} \quad (4.15)$$

$n$  altındaki koşullu dağılımı tabakalardan bağımsız basit tesadüfi örnekleme olması halinde geçerlidir.  $S_k^2$ ,  $k$  ıncı sonradan tabakada birim başına düşen varyanstır.

Koşullu olmayan varyans ise  $n$  'nin mümkün olan tüm dağılımları üzerinden ortalama alınarak elde edilir. Koşullu olmayan varyans aşağıdaki gibi belirlenir.

$$\begin{aligned} V(\bar{y}_{st}) &= E[V(\bar{y}_{st} | n_1, n_2, \dots, n_K)] \\ &= E \left[ \sum_{k=1}^K \left( \frac{N_k}{N} \right)^2 \left( 1 - \frac{n_k}{N_k} \right) \frac{S_k^2}{n_k} \right] \\ &= \sum_{k=1}^K \left\{ \left( \frac{N_k}{N} \right)^2 E \left( \frac{1}{n_k} - \frac{1}{N_k} \right) S_k^2 \right\} \\ V(\bar{y}_{st}) &= \sum_{k=1}^K \left\{ \left( \frac{N_k}{N} \right)^2 \left[ E \left( \frac{1}{n_k} \right) - \frac{1}{N_k} \right] S_k^2 \right\} \\ E \left( \frac{1}{n_k} \right) &\approx \frac{1}{n \frac{N_k}{N}} + \frac{1 - \frac{N_k}{N}}{n^2 \frac{N_k}{N}} \end{aligned}$$

$E(1/n_k)$ , Stephan (1945) yakınsamasına dayalı yapılmıştır ( $\approx$  asimptotik olarak denk demektir) ve  $n_k \neq 0$  tüm  $k = 1, \dots, K$  için kabul edilmiştir [Holt ve Smith, 1979].

Buradan,  $\bar{y}_{st}$  varyansı,

$$V(\bar{y}_{st}) \approx \left(\frac{1}{n} - \frac{1}{N}\right) \sum_{k=1}^K \left\{ \frac{N_k}{N} S_k^2 \right\} + \frac{1}{n^2} \sum_{k=1}^K \left\{ \left(1 - \frac{N_k}{N}\right) S_k^2 \right\} \quad (4.16)$$

olarak tanımlanır.

Koşullu olmayan varyansının iki bileşenden oluştuğunu görüyoruz. Birinci bileşen bilinen çapa orantılı paylaştırmadaki tabakalı örnekleme varyansıdır. İkinci bileşen sonradan tabakalamanın önceden tabakalamaya tercih edilmesinden dolayı oluşan varyanstaki artış miktarıdır ve ikinci bileşen çoğunlukla küçüktür. Dolayısıyla, sonradan tabakalamayla elde edilen kazançların orantılı paylaştırılmış tabakalı örneklemedekilerle aynı olduğu ileri sürülebilir. Bununla birlikte, sonradan tabakalama yaygın bir şekilde kullanılmakta ve çok önemli kazançlar elde edilmektedir [Smith, 1991].

Büyük  $n$  çapları için ikinci bileşen ilk bileşene kıyasla daha küçüktür. Dolayısıyla büyük örnekler için, tahminlerde sonradan tabakalama tekniğinin sonuçları neredeyse orantılı paylaştırma kadar iyidir. Koşullu olmayan varyans formülünün kullanımının bir araştırmayı planlamada (örnek seçiminden önce) veya bir yöntemi başka birine göre değerlendirmede faydalı olduğu belirtilmiştir [Singh, 2003].

Holt ve Smith (1979), koşullu ve koşullu olmayan güven aralıklarının kapsama özellikleri incelendiğinde koşullu aralıkların  $\bar{y}_{st}$  ve  $V(\bar{y}_{st}|n)$  herhangi bir  $n$  düzenleri  $\bar{Y}$  parametresi için uygun olduğunu göstermiştir. Diğer taraftan,  $\bar{y}_{st}$  ve  $V(\bar{y}_{st})$  dayalı koşullu olmayan aralıkların büyük bir kısmının  $\bar{Y}$ 'yi içermedikleri belirlemiştir.

## 4.2. Tarama

Çok boyutlu bir tabloyla temsil edilmekte olan ayrıntılı bir bilginin nasıl iyi tahmin edilebileceği yaygın problemi söz konusudur. Bu durum genellikle demografik ve coğrafik çalışmalarda ortaya çıkar ve çoğu kez tarama yöntemi uygulanarak çözümlenmeye çalışılır [Norman, 1999].

Tarama, örnekleme araştırmalarında iki veya daha fazla sayıda yardımcı değişkene dayalı oluşturulan çapraz sınıflamalar yapılması durumunda yaygın kullanılan bir ağırlıklandırma yöntemidir. Çapraz sınıflamalardaki göze yığın çapları bilinmiyor ve yalnızca marjinal yığın çapları biliniyorken sonradan tabakalama yerine kullanılabilecek alternatif bir yöntem tarama olarak adlandırılır. Sonradan tabakalamada olduğu gibi amaç, örneği hedef yığın hakkında bilinene uydurmaya çalışmaktır [Little, 1993].

Tarama yöntemi, mevcut yardımcı bilgiyi kullanarak örnekleme ağırlıklarını düzenlemeyi içerir ve yardımcı değişkenlerin bilinen yığın toplamalarını kullanır. Bu yöntemle çapraz tablolarda marjinal oranlar başarılı bir şekilde kullanılarak tablodaki gözeler için oldukça güvenilmeyen bilgi çok daha güvenilir olan alt toplamalarla iyileştirilir. Göze ağırlıkları için istikrarlı tahminler elde edilmeye çalışılır [Holt ve Eliot, 1991]. Çok sayıda yardımcı değişken kullanıldığında ve göze sayısının çok olması durumlarında sonradan tabakalamadan daha iyi tahminler sağlayacağı söylenebilir.

Tarama, örnek verisini (dışarıdan) bilinen marjinal yığın çaplarına yinelemeli olarak orantılama işlemidir [Oh ve Scheuren, 1978]. Yineleme orantılı uyum, tarama işleminin basit yinelemesidir [Deming ve Stephan, 1940; Little, 1993]. Yineleme orantılı uyum (iterative proportional fitting) yöntemi ilk olarak Deming ve Stephan tarafından 1940 U.S. Nüfus Sayımından yığın ile örnek verisi arasındaki tutarlılığı sağlamanın bir yolu olarak önerilmiştir [Little, 1993]. Bu iki yöntemin aynı sezgisel açıklamaları olmasına karşın, standart hataların tahminleri için farklı istatistiksel özellikleri vardır [Lu ve Gelman, 2003].

Yineleme orantılı uyum, çeşitli disiplinlerce yeni baştan ele alınmış ve farklı adlar altında incelenmiştir. Bu ağırlıklandırma yöntemi literatürde tarama (raking) veya tarama oran (raking ratio) yöntemi olarak bilinmektedir. Çeşitli alanlarda, kenar ağırlıklandırma (rim weighting), marjinal ayarlama (marjinal calibration), marjinal ağırlıklandırma (marginal weighting), örnek dengeleme (sample balancing) adları altında incelenmiştir.

Tarama yönteminin uygulanmasında ilk olarak çapraz tablolarda örnek satır toplamaları yığın satır toplamalarına uymaya zorlanır. Sonra örnek düzeltilmiş sütun toplamaları yığın sütun toplamalarına uymaya zorlanır. Daha sonra örnek satır toplamaları uydurulabilmek için yeniden düzeltilir ve yakınsama sağlanana kadar devam edilir. Uygun yakınsamaya çoğunlukla hızlı bir şekilde ulaşılır. Buna karşın bazı durumlarda yakınsama yavaş olabilir ve garanti edilmez [Kalton ve Flores-Cervantes, 2003].

Çalışmada BTÖ tasarımı tarama ağırlıklandırma yöntemi, yalnızca iki kategorik yardımcı değişken durumu için incelenecektir. Yöntem daha fazla sayıda yardımcı değişken olması durumu için de benzer şekilde uygulanır.

Araştırma örnek verisi  $x_1$  ve  $x_2$  gibi iki yardımcı değişkene göre çapraz sınıflandırılıyor olsun.  $x_1$  değişkeninin  $K$  sayıda kategorisi ve  $x_2$  değişkeninin  $S$  sayıda kategorisi varken,  $x_1$  ve  $x_2$ 'ye göre çapraz sınıflama  $K \times S$  sayıda gözeden oluşur. Çapraz sınıflamaya ilişkin yığın ve örnek çaplarının gösterimleri aşağıda Çizelge 4.2'de verilmiştir.

Örnek verisinde  $n_{ks}$  göze çapları tüm  $k = 1, \dots, K$  ve  $s = 1, \dots, S$  gözeleri için biliniyordur. Yığın marjinal toplamalarından  $N_{k.}$ , satır marjinal toplamaları tüm  $k = 1, \dots, K$  için ve  $N_{.s}$ , sütun marjinal toplamaları tüm  $s = 1, \dots, S$  için bilinmektedir.

Çizelge 4.2.  $x_1$  ve  $x_2$  değişkenlerine dayalı çapraz tablo göze yığın ve örnek çaplarının gösterimi

| YIĞIN ÇAPLARI |   |          |          |                |          |          |          |
|---------------|---|----------|----------|----------------|----------|----------|----------|
|               |   | $x_2$    |          |                |          |          | Toplam   |
|               |   | 1        | 2        | s              |          | S        |          |
| $x_1$         | 1 | $N_{11}$ | $N_{12}$ | . . . $N_{1s}$ | . . .    | $N_{1S}$ | $N_{1.}$ |
|               | 2 | $N_{21}$ | $N_{22}$ | . . . $N_{2s}$ | . . .    | $N_{2S}$ | $N_{2.}$ |
|               | . | .        | .        | .              | .        | .        | .        |
|               | . | .        | .        | .              | .        | .        | .        |
|               | . | .        | .        | .              | .        | .        | .        |
|               | k | $N_{k1}$ | $N_{k2}$ | .              | $N_{ks}$ | $N_{kS}$ | $N_{k.}$ |
|               | . | .        | .        | .              | .        | .        | .        |
|               | . | .        | .        | .              | .        | .        | .        |
| K             |   | $N_{K1}$ | $N_{K2}$ | . . . $N_{Ks}$ | . . .    | $N_{KS}$ | $N_{K.}$ |
| Toplam        |   | $N_{.1}$ | $N_{.2}$ | . . . $N_{.s}$ | . . .    | $N_{.S}$ | $N$      |

| ÖRNEK ÇAPLARI |   |          |          |                |          |          |          |
|---------------|---|----------|----------|----------------|----------|----------|----------|
|               |   | $x_2$    |          |                |          |          | Toplam   |
|               |   | 1        | 2        | s              |          | S        |          |
| $x_1$         | 1 | $n_{11}$ | $n_{12}$ | . . . $n_{1s}$ | . . .    | $n_{1S}$ | $n_{1.}$ |
|               | 2 | $n_{21}$ | $n_{22}$ | . . . $n_{2s}$ | . . .    | $n_{2S}$ | $n_{2.}$ |
|               | . | .        | .        | .              | .        | .        | .        |
|               | . | .        | .        | .              | .        | .        | .        |
|               | . | .        | .        | .              | .        | .        | .        |
|               | k | $n_{k1}$ | $n_{k2}$ | .              | $n_{ks}$ | $n_{kS}$ | $n_{k.}$ |
|               | . | .        | .        | .              | .        | .        | .        |
|               | . | .        | .        | .              | .        | .        | .        |
| K             |   | $n_{K1}$ | $n_{K2}$ | . . . $n_{Ks}$ | . . .    | $n_{KS}$ | $n_{K.}$ |
| Toplam        |   | $n_{.1}$ | $n_{.2}$ | . . . $n_{.s}$ | . . .    | $n_{.S}$ | $n$      |

Tarama yöntemi  $N_{ks}$  yığın göze toplamlarının örnek üzerinden tahminini,  $N_{k.}$  satır ve  $N_{.s}$  sütun marjinal yığın dağılımlarını birbirini izleyen adımlarla ardı ardına sağlamak için  $n_{ks}$  örnek çaplarının yineleme orantılı uyumudur. Tarama yöntemiyle belirlenen nihai son  $k$  ıncı satır ve  $s$  inci sütun göze toplamı tahmini değerleri  $\tilde{N}_{ks}$  ile gösterilebilir.

Yöntemin uygulanmasına öncelikli olarak, tüm gözeler için,

$$N_{ks}^{(0)} = \hat{N}_{ks} = n_{ks}$$

tanımlaması yapılmasıyla başlanır. Daha sonra, birbiri ardınca satırlar ve sütunlar  $t = 1, 2, 3, \dots$  adımlarında,  $k = 1, \dots, K$  ve  $s = 1, \dots, S$  gözeleri için,

$$N_{ks}^{(2t-1)} = N_{ks}^{(2t-2)} \frac{N_{k.}}{\sum_{s=1}^S N_{ks}^{(2t-2)}}$$

$$N_{ks}^{(2t)} = N_{ks}^{(2t-1)} \frac{N_{.s}}{\sum_{k=1}^K N_{ks}^{(2t-1)}}$$

düzeltilir. Düzeltilmiş satır ve sütun marjinal toplamaları yığın satır ve sütun marjinal toplamalarına yakınsayana kadar yineleme sürdürülür [Ardilly ve Tillé, 2006].

Her bir adım bir önceki adımın sonuçlarıyla başlar, tüm denklemler eş zamanlı olarak istenilene yakın olduğunda yani,  $\tilde{N}_{ks}$  'ler için,

$$\sum_{s=1}^S \tilde{N}_{ks} = N_{k.}$$

$$\sum_{k=1}^K \tilde{N}_{ks} = N_{.s}$$

sağlandığında işlem sonlandırılır.  $\tilde{N}_{ks}$  'ler,  $k$  ıncı satır  $s$  inci sütun için son adımda elde edilen göze toplamı değeridir.  $n_{ks}$  'ler sıfır değerlerinden oluşmamışsa algoritma hızla yakınsama gösterecektir.

Tarama ile elde edilen  $\tilde{N}_{ks}$  tahmin edicisinin asimptotik olarak yansız, normal dağılımlı ve minimum varyanslı olduğu belirlenmiştir.  $\tilde{N}_{ks}$  , en iyi asimptotik yansız normal (Best Unbiased Normal) tahmin edicilerdir [Oh ve Scheuren, 1983].

Yığın toplamı,

$$Y = \sum_{k=1}^K \sum_{s=1}^S \sum_{i=1}^{N_{ks}} y_{ksi}$$

için tarama yöntemiyle örnek üzerinden tahmin edicisi,

$$\hat{Y}_t = \sum_{k=1}^K \sum_{s=1}^S \frac{\tilde{N}_{ks}}{n_{ks}} \left( \sum_{i=1}^{n_{ks}} y_{ksi} \right) \quad (4.17)$$

$$\hat{Y}_t = \sum_{k=1}^K \sum_{s=1}^S \frac{\tilde{N}_{ks}}{n_{ks}} y_{ks} \quad (4.18)$$

eşitliğiyle elde edilir [Oh ve Scheuren, 1987].  $y_{ks} = \sum_{i=1}^{n_{ks}} y_{ksi}$  ,  $k$  ıncı satır  $s$  inci sütundaki gözlemlerin toplam değeridir. Buradan, yığın toplamının tarama tahmin edicisi,

$$\hat{Y}_t = \sum_{k=1}^K \sum_{s=1}^S {}_tW_{ks}^{**} \left( \sum_{i=1}^{n_{ks}} y_{ksi} \right) \quad {}_tW_{ks}^{**} = \frac{\tilde{N}_{ks}}{n_{ks}} \quad (4.19)$$

olarak tanımlanabilir.  ${}_tW_{ks}^{**} = \frac{\tilde{N}_{ks}}{n_{ks}}$ ,  $k$  ıncı satır  $s$  inci sütun tarama ağırlığıdır,  $k = 1, \dots, K$ ,  $s = 1, \dots, S$ .

ve yığının ortalamasının tahmin edicisi,

$$\bar{y}_t = \frac{1}{N} \sum_{k=1}^K \sum_{s=1}^S \sum_{i=1}^{n_{ks}} {}_tW_{ks}^{**} y_{ksi} \quad {}_tW_{ks}^{**} = \frac{\tilde{N}_{ks}}{n_{ks}} \quad (4.20)$$

şeklinde belirlenir.  $\hat{Y}_t$  ve  $\bar{y}_t$  birer Horvitz-Thompson tahmin edicileridir. BTÖ de yığının ortalamasının Hájek tahmin edicisi,

$$\bar{y}_{t,w} = \frac{\sum_{k=1}^K \sum_{s=1}^S \sum_{i=1}^{n_{ks}} {}_tW_{ks}^{**} y_{ksi}}{\sum_{k=1}^K \sum_{s=1}^S \sum_{i=1}^{n_{ks}} {}_tW_{ks}^{**}} \quad {}_tW_{ks}^{**} = \frac{\tilde{N}_{ks}}{n_{ks}} \quad (4.21)$$

$$\bar{y}_{t,w} = \frac{\sum_{k=1}^K \sum_{s=1}^S {}_tW_{ks}^{**} y_{ks}}{\sum_{k=1}^K \sum_{s=1}^S {}_tW_{ks}^{**} n_{ks}}$$

olarak elde edilir. Tarama algoritmasının bir özelliği olarak  $n_{ks} = 0$  ise,  $\tilde{N}_{ks} = 0$  olacağını belirtmek önemlidir. Bu tür durumlarda  ${}_tW_{ks}^{**} = 0$  dır.

Oh ve Scheuren (1987), tarama ile elde edilen tahmin edicilerin koşullu yaklaşım özellikleri üzerinde durmuştur. Holt ve Smith (1979) ve Rao (1985) tarafından da savunulan koşullu yaklaşıma başvurulmuştur.

#### 4.2.1. Koşullu yan

$\hat{Y}_t$  tahmin edicisinin Oh ve Scheuren (1983) kapsamında incelenen  $\underline{n} = (n_{11}, n_{12}, \dots, n_{ks}, \dots, n_{KS})$  altındaki koşullu özellikleri incelenecektir.  $\bar{Y}_{ks}$ , k ıncı satır s inci sütun göze yığın ortalamasını gösteriyor olsun.  $\hat{Y}_t$  yığın toplamının tarama tahmin edicisi koşullu beklenen değeri,

$$E(\hat{Y}_t | \underline{n}) = \sum_{k=1}^K \sum_{s=1}^S \tilde{N}_{ks} \bar{Y}_{ks} = Y + \sum_{k=1}^K \sum_{s=1}^S (\bar{Y}_{ks} - \bar{Y})(\tilde{N}_{ks} - N_{ks})$$

şeklinde tanımlanır.  $\hat{Y}_t$ , yığının yapısına dayalı olan yanın önemine ve taramanın yakınsayıp yakınsamamış olması durumuna göre koşullu yanlı bir tahmin edicidir. (Yakınsama için tarama uygulandığında koşullu olmaksızın asimptotik olarak  $E(\tilde{N}_{ks}) = N_{ks}$  olur.)

Genel varyans kuralları uygulanarak (örn., Scheffe, 1959)

$$(\bar{Y}_{ks} - \bar{Y}) = (\bar{Y}_{k.} - \bar{Y}) + (\bar{Y}_{.s} - \bar{Y}) + (\bar{Y}_{ks} - \bar{Y}_{k.} - \bar{Y}_{.s} + \bar{Y})$$

yazılabilir [Oh ve Scheuren, 1987]. Böylece koşullu varyans, verilen  $\underline{n}$  altında,

$$\begin{aligned} \text{Yan}(\hat{Y}_t | \underline{n}) &= \sum_{k=1}^K (\bar{Y}_{k.} - \bar{Y})(\tilde{N}_{k.} - N_{k.}) + \sum_{s=1}^S (\bar{Y}_{.s} - \bar{Y})(\tilde{N}_{.s} - N_{.s}) \\ &\quad + \sum_{k=1}^K \sum_{s=1}^S (\bar{Y}_{ks} - \bar{Y}_{k.} - \bar{Y}_{.s} + \bar{Y})(\tilde{N}_{ks} - N_{ks}) \end{aligned}$$

ile tanımlanabilir. Eğer tarama yakınsama için ise bu durumda koşullu yanın ilk iki terimi sıfır olur. Taramanın her iki biçimi için de koşullu yanın üçüncü terimi sıfır olacaktır.  $y_{ks}$  'ler için herhangi bir etkileşim olmadığı kabul edilir, bu da yeterlidir.

Oh ve Scheuren (1987), çok sayıda değişkene dayalı büyük ölçekli araştırmalarda, bu varsayımı yapmak gerçekçi olmayabileceğini, ancak her şeye rağmen uygulamada etkileşimin çoğu kez  $y_{ks}$  'nin küçük bir parçası olduğunu belirtmiştir. Bundan dolayı, tarama oran tahmin edicisi, pek çok durumda, düşük örnek çaplarında dahi küçük yanlara sahip olduğu belirtilmiştir.

#### 4.2.2. Koşullu varyans

Tarama oran tahmin edicisinin varyansı için koşullu ve koşullu olmayan yaklaşımlar incelenmiştir. Oh ve Scheuren (1987) yineleme tekniklerini uygulamış ve koşullu yaklaşımda verilen  $\underline{n}$  altında  $\hat{Y}_t$  'nin koşullu varyansı,

$$V(\hat{Y}_t | \underline{n}) = \sum_{k=1}^K \sum_{s=1}^S n_{ks} \left( \frac{\tilde{N}_{ks}}{n_{ks}} \right)^2 \left( 1 - \frac{n_{ks}}{N_{ks}} \right) V_{ks} \quad (4.22)$$

şeklinde tanımlanmıştır. Burada  $V_{ks}$ ,  $ks$ -inci alt grubun yığın varyansıdır ve eğer  $N_{ks} = 0$  veya 1 ise  $V_{ks} = 0$  olur. (Ayrıca, veri olmaması durumunu göstermesinden dolayı,  $0/0=0$  kuralı uygulanmıştır.)

Koşullu varyans, taramanın yakınsamaya doğru gidip gitmediğini yansıtır. Buna rağmen çok az çalışılmıştır, çünkü koşullu varyansın tahminine kolayca uyarlanamaz. En önemli zorluk,  $n_{ks}$  'ler küçük olduğunda  $V_{ks}$  'nin sabit tahmin edicisini hesaplamada yetersiz olunmasından kaynaklanmaktadır [Oh ve Scheuren, 1987].

### 4.3. Ayarlama Tahmin Edicisi

Ayarlama, örnekleme araştırmalarında tahmin üzerine yazılmış pek çok yeni makalede temel teşkil eden bir konudur ve “ayarlama yaklaşımı” veya “ayarlama tahmin edicileri” başlıkları altında da incelenmektedir [Särndal, 2007].

Yardımcı bilgiyi tahmin aşamasında kullanılan en yaygın yöntemler regresyon ve oran tahmini yöntemleridir. Regresyon tahmini yöntemine benzeyen ve daha genel olan bir yöntem de ayarlama (calibration) tahmin edicisidir. Ayarlama tahmin edicisi Deville ve Särndal (1992) tarafından anket verisinde yardımcı bilgi kullanımı kapsamında verilmiştir. Deville ve Särndal ayarlama ağırlıklarının pek çok alternatif yapıları olduğunu ve örnekleme kapsamında asimptotik olarak aynı olduklarını ispatlamıştır.

Örnekleme araştırmalarında parametre tahmini, yaygın bir biçimde yardımcı bilgiye dayalı model-tabanlı yaklaşımlarla yapılmaktadır. Literatürde model-tabanlı yaklaşımlar olarak sınıflandırılabilir üç temel yöntem önerilmiştir: genelleştirilmiş regresyon tahmin edicileri (GREG) (Cassel, Särndal ve Wretman, 1976; Särndal, 1980); ayarlama tahmin edicileri (Deville ve Särndal, 1992); ve son zamanlarda önerilen ampirik olabilirlik yöntemleridir [Kott, 2003].

Model-tabanlı ile tahmin edicilerin yaklaşık (asimptotik) yansız olduğu ifade edilmektedir. Bu, çalışılan modelin doğru veya yanlış olmasına bakılmaksızın etkin olması ve eğer model doğru ise tahmin edicinin önemli düzeyde etkin olması demektir [Wu ve Sitter, 2001]. Ayarlama, örnekleme ağırlıklarının araştırma verisine uygulanması durumunda ilgilenilen farklı araştırma değişkenleri için model-yansız tahmin ediciler sağlar [Kott, 2003].

Ayarlama tahmin edicisi yardımcı değişken (değişkenler) bilgisini tahminlerin etkinliğini arttırmak için kullanır. Tahminde kullanılacak yardımcı değişkenler için yığın toplamalarının bilinmesini gerektirir. Tahmin edicinin etkinliği yardımcı değişkenlerin ilgilenilen değişken  $y$ ’yi ne kadar iyi açıklayabildiğine bağlıdır.

Ayarlama, orijinal olarak örnekleme hatalarını azaltmak amacıyla geliştirilmiştir. Uygulamada ise  $W_i$  örnekleme (tasarım) ağırlıklarında genel olarak iki durum için ayarlama düzeltmesi yapılır [Särndal, 2007].

- (i) Altında/üstünde kapsam olmasından dolayı oluşan yanı gidermek için düzeltme. Örnekleme ağırlıklarında kapsanmama ve örnek ile yığın arasındaki diğer ayrılıkları gidermek amacıyla ayarlama yapmaktır.
- (ii) Yanıtlanmama olmasından dolayı oluşan yanı gidermek için düzeltme. Ağırlıklar yanıt eğilimi faktörünün tersiyle düzeltilir. Düzeltme, hem yanıtlayan hem de yanıt alınamayanlar için bilinen yardımcı değişkenlere dayalı olarak yanıtlayanların ağırlıklarının ayarlanmasıyla gerçekleştirilir.

$U$ , yığını  $N$  sayıda  $u_1, u_2, \dots, u_N$  birimlerinden oluşuyor olsun.  $y$  ilgilenilen değişken ve  $x$  yardımcı değişkeni gösteriyor olsun.  $N$  çaplı bir  $U$  sonlu yığımından belirlenen bir örnekleme yöntemiyle yerine koymaksızın  $n$  çaplı bir  $s$  örneği alınıyor olsun, ( $s \subseteq U$ ).  $\pi_i = P(u_i)$ ,  $i$  inci birimin  $n$  çaplı örneğe seçilmesi olasılığı olsun ( $i = 1, \dots, N$ ).  $i \in s$  örnekleme birimleri için,  $(y_i, x_i)$  gözlemlendiği ve  $x$ 'in  $X = \sum_{i=1}^N x_i$  yığın toplamının bilindiği varsayalım.

Yığın toplamı

$$Y = \sum_{i=1}^N y_i$$

için örnek üzerinden HT tahmin edicisi,

$$\hat{Y} = \sum_{i=1}^n W_i y_i = \hat{Y}_{HT}$$

ile verilebilir.  $W_i = 1/\pi_i$ , örnekleme ağırlıkları yardımcı değişken bilgisine dayalı olmaksızın belirlenmiştir ve  $\sum_{i=1}^n W_i = N$  dir.

Ayarlama tahmin edicisiyle öncelikli olarak ilgilenilen değişkenle ilişkili olduğu belirlenen yardımcı değişkenin (veya değişkenlerin) gözlenen değerlerine örnekleme ağırlıkları atanır. Yardımcı değişken  $x$  'in yığın toplamı

$$X = \sum_{i=1}^N x_i$$

için bir tahmini örnekleme ağırlıklarına dayalı olarak,

$$\hat{X} = \sum_{i=1}^n W_i x_i = \hat{X}_{HT} \quad (4.23)$$

şeklinde belirlenir.

Bu yöntemle,  $y$  ilgilenilen değişkenin parametre tahminlerinin etkinliğinde artış sağlanmak amacıyla,  $x$  yardımcı değişkeni ve yardımcı değişkenin yığın toplamı bilgisi kullanılır. Bunun için,  $W_i$  örnekleme ağırlıklarında düzeltme yapılarak yeni  ${}_cW_i$  ayarlama ağırlıkları elde edilmeye çalışılır.

${}_cW_i$  ayarlama ağırlıkları, verilen bir uzaklık ölçümüne dayalı olarak  $W_i$  örnekleme ağırlıklarına olabildiğince yakın olacak şekilde belirlenmeye çalışılır. Bu durum, yardımcı değişken değerlerinin ağırlıklandırılmış örnek toplamının  $X$  yığın toplamına eşit yapılması gerektiğini ifade eder.

$W_i$ , örnekleme ağırlıklarında düzeltme yapma fikri, “ağırlıklandırma kontrol işlemlerini” inceleyen Zieschang (1986, 1990) tarafından BM Tüketici Harcaması Anketi kapsamında genelleştirilmiş en küçük kareler ağırlıklandırma algoritmasıyla

verilmiştir. Deville (1988), Lemel (1976)'in fikrini genişleterek bilinen  $x$ -toplamları üzerinde ayarlama yaparak  $W_i = 1/\pi_i$  örnekleme ağırlıklarında değişiklik yapmayı amaçlamıştır [Deville ve Särndal, 1992]

Yardımcı değişkeninin bilinen yığın toplamı  $X$ , ayarlama ağırlıklarına bağlı olarak aşağıdaki gibi tanımlanır.

$$X = \sum_{i=1}^N x_i = \sum_{i=1}^n {}^c W_i x_i \quad (4.24)$$

Bu ayarlama için bir kısıttır ve ayarlama kısıtı bazen “ayarlama denklemi” olarak da adlandırılmaktadır [Kott, 2003; Särndal, 2007].

Ayarlamadaki asıl fikir örnek birimleri için bir veya daha çok ayarlama kısıtlarının sağlandığı  ${}^c W_i$  ağırlıklarını hesaplamaktır. Dolayısıyla, kısıtın sağlanmasıyla tutarlılık da sağlanmış olur. Çünkü yardımcı değişken ile ilgilenilen değişken arasında güçlü bir ilişkinin olması yardımcı değişken için önemli rol oynayan ağırlıkların ilgilenilen değişken için de önemli olacağı anlamına gelir [Deville ve Särndal, 1992].  ${}^c W_i$  ayarlama ağırlıkları, Eş. 4.24'ü yani ayarlama denklemini sağlayacak şekilde belirlenir ve daha sonra  $Y$  yığın toplamının ayarlama tahmin edicisini belirlemede kullanılır.

Örnek birimleri için  $\pi_i$  seçim olasılıkları biliniyorsa  $Y$  yığın toplamı tahmini verilen bir uzaklık ölçümü ve ayarlama tahmin edicisi yöntemiyle ağırlıklandırma yapılarak elde edilebilir.  $Y$ 'nin ayarlama tahmin edicisi;

$$\hat{Y}_C = \sum_{i=1}^n {}^c W_i y_i \quad (4.25)$$

şeklinde tanımlanır.  $cW_i$ , bilinmeyen ayarlama ağırlıklarıdır ve yardımcı değişkene dayalı yapılan ağırlıklandırmayı ifade eder,  $i = 1, \dots, n$ . Burada  $\sum_{i=1}^n cW_i = N$  dir.

$\bar{Y}$  yığının ortalamasının ayarlama tahmin edicisi,

$$\bar{y}_c = \frac{1}{N} \sum_{i=1}^n cW_i y_i \quad (4.26)$$

ile elde edilir.

Ayarlama ağırlıklarını belirlemede literatürde iki yaklaşım vardır: minimum uzaklık yaklaşımı ve fonksiyonel yapı yaklaşımıdır. Fonksiyonel yapı yaklaşımı daha genel olmakla birlikte iki yaklaşımla aynı sonuçlara ulaşılmaktadır.

Deville ve Särndal (1992), minimum uzaklık yaklaşımı ile ayarlama ağırlıklarını incelemiştir. Estevao ve Särndal (2000), fonksiyonel yapı yaklaşımını önermiştir. Tez kapsamında minimum uzaklık yöntemiyle ağırlıkların belirlenmesi incelenecektir.

#### 4.3.1. Ayarlama ağırlıklarının belirlenmesi

Ayarlama ağırlıkları, yansız tahminler sağlayan  $W_i$  örnekleme ağırlıklarında mümkün olduğunca az değişiklik yapılarak belirlenmeye çalışılır. Araştırma istatistikçileri örnekleme ağırlıklarının yakınında durmayı isterler. Ağırlıkların uygulanmasıyla yansız tahminler elde edileceği denilemez ama yansızlığa çok yakın olunacağı söylenebilir [Deville ve Särndal, 1992; Wu, 2003].

$E(\cdot)$ , örnekleme tasarımıyla ilgili beklenen değeri ifade ediyorken, ortalama uzaklığın bir ölçümü,  $E_p \left\{ \sum_{i=1}^n (cW_i - W_i)^2 / W_i \right\}$  ki-kare istatistiğini anımsatır. Bu

terim daha da genelleştirilirse,  $1/q_i$ ,  $i$  inci birimin  $W_i$  ile bağlantılı olmayan belirlenen pozitif ağırlıkları iken ortalama uzaklık

$$E_p \left\{ \sum_{i=1}^n ({}_cW_i - W_i)^2 / (q_i W_i) \right\} \quad (4.27)$$

şeklinde tanımlanır [Deville ve Särndal, 1992].

Dolayısıyla, tüm örnek birimleri için  $X = \sum_{i=1}^n {}_cW_i x_i$  ayarlama kısıtına bağlı olarak Eş. 4.26 ortalama uzaklık denkleminin minimumu belirlenir. Bunun için,  $\hat{X}_{HT} = \sum_{i=1}^n W_i x_i$  ile  $X = \sum_{i=1}^n {}_cW_i x_i$  ağırlıklandırma kümeleri arasındaki yakınlığı saptamaya yarayacak bir kriter belirlemek gereklidir.

Bu kriter,  $W_i$  örnekleme ağırlıkları ile  ${}_cW_i$  ayarlama ağırlıkları arasındaki uzaklıktır ve uzaklığın ölçümü genel olarak  $G = G({}_cW_i, W_i) = G({}_cW_i / W_i)$  ile gösterilir.

Tüm s örneği için  $W_i$  ile  ${}_cW_i$  arasındaki toplam uzaklık;

$$\begin{aligned} \sum_i^n W_i G({}_cW_i / W_i) &= \sum_{i=1}^n W_i ({}_cW_i / W_i - 1)^2 / q_i \\ &= \sum_{i=1}^n ({}_cW_i - W_i)^2 / (q_i W_i) \end{aligned}$$

olarak ifade edilebilir.

Deville ve Särndal (1992),  $\hat{Y}_c$  'nin istatistiksel özelliklerini incelemiş ve yedi farklı uzaklık ölçümünü detaylı bir şekilde ele almıştır. İncelenen uzaklık ölçümleriyle elde edilen sonuçların hemen hemen aynı sonuçları verdiği ve asimptotik olarak

birbirilerine denk oldukları belirlenmiştir. Araştırmalarda  $G = G({}_cW_i, W_i)$  uzaklık ölçümlerinden ki-kare uzaklık ölçümü en yaygın kullanılanıdır.

Ki-kare uzaklık ölçümü;

$$\sum_i^n G({}_cW_i, W_i) = \frac{1}{2} \sum_{i=1}^n ({}_cW_i - W_i)^2 / q_i W_i = \frac{1}{2} \sum_{i=1}^n W_i ({}_cW_i / W_i - 1)^2 / q_i \quad (4.28)$$

şeklindedir.

$\hat{X}_{HT}$  ile  $X$  arasındaki yakınlığı saptamada, toplam uzaklığın minimum yapılmasında Lagrange çarpanı yöntemi kullanılabilir.  $L$  Lagrange fonksiyonu,

$$\begin{aligned} L &= \sum_i^n G({}_cW_i, W_i) + \lambda \left[ X - \sum_{i=1}^n {}_cW_i x_i \right] \\ L &= \frac{1}{2} \sum_{i=1}^n W_i ({}_cW_i / W_i - 1)^2 / q_i + \lambda \left[ X - \sum_{i=1}^n {}_cW_i x_i \right] \end{aligned} \quad (4.29)$$

biçiminde yazılabilir. Burada  $\lambda = (\lambda_1, \dots, \lambda_K)'$ ;  $K$  vektör Lagrange çarpanıdır,  $K$  yardımcı değişken sayısıdır.

Lagrange fonksiyonundan yararlanılarak  ${}_cW_i$ 'nin en iyi değeri,  $L$ 'nin  ${}_cW_i$ 'ye göre türevinin alınıp sıfıra eşitlenmesiyle aşağıdaki gibi hesaplanır.

$$\frac{\partial L}{\partial {}_cW_i} = 2 \left[ \frac{{}_cW_i}{W_i q_i} - 1 \right] - \lambda x_i' = 0$$

$x_i$  yerine  $x_i'$  gösterimi kullanılmıştır ki matris yoluyla çözüm yapılabilsin.

$$\frac{\partial L}{\partial {}_cW_i} = 2({}_cW_i - 1) - \lambda W_i q_i x_i' = 0$$

Bu minizasyon ayarlama ağırlıklarını verir. Diferansiyel denklemde  ${}_cW_i$  yalnız bırakıldığında,

$${}_cW_i = W_i(1 + q_i x_i' \lambda), \quad i=1, \dots, n \quad (4.30)$$

sonucuna ulaşılmış olur. Yeni ağırlıklar yansız tahminler verecektir denilemez ancak gerçek bir beklenti olarak yansızlığın çok yakınında olunacağı söylenebilir.

Burada  $\lambda$  değerine ihtiyaç vardır. Eş. 4.30'un her iki tarafı sağdan  $x_i$  ile çarpıldığında,

$${}_cW_i x_i - W_i x_i = W_i q_i x_i x_i' \lambda \quad (4.31)$$

eşitliğine ulaşılır. Eş. 4.31'in tüm  $n$ 'ler üzerinden toplam alındığında,

$$\sum_{i=1}^n {}_cW_i x_i - \sum_{i=1}^n W_i x_i = \sum_{i=1}^n W_i q_i x_i x_i' \lambda$$

ve Eş. 4.23 ile Eş. 4.24'den aşağıdaki eşitlik elde edilir.

$$X - \hat{X} = \left[ \sum_{i=1}^n W_i q_i x_i x_i' \right] \lambda \quad (4.32)$$

$\lambda$ 'nın hesaplanabilmesi için  $\lambda$ 'nın yalnız bırakılması gereklidir. Bunun için Eş. 4.32'nin her iki tarafı soldan  $\left[ \sum_{i=1}^n W_i q_i x_i x_i' \right]^{-1}$  ters matrisi ile çarpıldığında  $\lambda$  lagrange çarpanları vektörüne,

$$\lambda = \left[ \sum_{i=1}^n W_i q_i x_i x_i' \right]^{-1} (X - \hat{X}) \quad (4.33)$$

ulaşılmış olur.  $T = \sum_{i=1}^n W_i q_i x_i x_i'$  simetrik matrisi tanımlanırsa

$$\lambda = T^{-1}(X - \hat{X}) \quad (4.34)$$

olarak da yazılabilir.  $\lambda$  değeri Eş. 4.30'ta yerine yazılacak olursa ayarlama ağırlıkları hesaplanmış olacaktır.  $Y$  yığın toplamının ayarlama tahmin edicisi,

$$\hat{Y}_C = \sum_{i=1}^n {}_c W_i y_i \quad {}_c W_i = W_i (1 + q_i x_i' \lambda) \quad (4.35)$$

şeklinde hesaplanır.  $q_i$ 'ye verilecek değerlere bağlı olarak ayarlama tahmin edicisi farklı yapılarda hesaplanır. Ayarlama tahmin edicisi  $q_i = 1/x_i$  alınmasıyla bilinen oran tahmin edicisine dönüşürken,  $q_i = 1$  alınması durumunda GREG tahmin edicisine ulaşılır. Aşağıda oran ve GREG tahmin edicisinin ayarlama ile elde edilmesi verilecektir.

#### 4.3.2. Yığın toplamının oran tahmin edicisinin ayarlamayla belirlenmesi

Ayarlama ağırlıklarında bazen eşit olmayan  $1/q_i$  de kullanılmaktadır. Ayarlama denkleminde  $q_i = 1/x_i$  eşit olmayan ağırlığı alınmasıyla iyi bilinen oran tahmin edicisine ulaşılır.  $q_i = 1/x_i$  alındığında ( $x_i$  pozitif sabit sayı alınarak),  $x_i' \lambda = x_i \lambda$  olur.

$$T = \sum_{i=1}^n W_i q_i x_i x_i' = \sum_{i=1}^n W_i \frac{1}{x_i} x_i x_i' = \sum_{i=1}^n W_i x_i'$$

$$\lambda = \left( \sum_{i=1}^n W_i x_i \right)^{-1} (X - \hat{X}),$$

$$\lambda = \hat{X}^{-1} (X - \hat{X})$$

$$\lambda = \frac{X}{\hat{X}} - 1$$

$${}_cW_i = W_i(1 + q_i x'_i \lambda) = W_i(1 + \frac{1}{x_i} x'_i \lambda)$$

$${}_cW_i = W_i(1 + \lambda)$$

$${}_cW_i = W_i(1 + \frac{X}{\hat{X}} - 1)$$

$${}_cW_i = W_i(\frac{X}{\hat{X}}) \quad (4.36)$$

$\hat{Y}_C = \sum_{i=1}^n {}_cW_i y_i$  eşitliğinde yerine yazılacak olduğunda,

$$\hat{Y}_C = \sum_{i=1}^n W_i(\frac{X}{\hat{X}}) y_i = \frac{X}{\hat{X}} \sum_{i=1}^n W_i y_i$$

$$\hat{Y}_C = \frac{\hat{Y}}{\hat{X}} X \quad (4.37)$$

bilinen oran tahmin edicisine ulaşılmış olur.

### 4.3.3. Genelleştirilmiş regresyon tahmin edicisi

Araştırma örneklemede regresyon kullanımıyla ilgili ilk referanslar Jessen (1942) ve Cochran (1942) tarafından verilmiştir. Fuller ve Isaki (1981) ile Isaki ve Fuller (1982), regresyon modeli belirli değişkenleri içeriyorsa model altında oluşturulan yığın ortalamasının regresyon tahmin edicisinin tasarım yansız olduğunu göstermişlerdir. Cassel, Särndal ve Wretman (1976), regresyon tahmin edicilerini belirlenmesinde hem model hem de tasarım ilkelerini dikkate almış ve “genelleştirilmiş regresyon tahmin edicileri” (GREG) terimini önermiştir [Fuller, 2002].

Cassel, Särndal ve Wretman (1976), Deville ve Särndal (1992), Deville, Särndal ve Sautory (1993), Kott (2003), Fuller (2002), Ardilly ve Tillé (2006) ve daha başka araştırmacılar GREG tahmin edicisi terimini kullanmıştır.

Yardımcı veri kullanılması ilgilenilen yığın hakkında çok daha iyi bilgiler sağlanmasına yardımcı olur ve çoğunlukla tahmin edilen yığın parametrelerinin örnekleme varyanslarını minimize eder. Regresyon tahmini iyi etkinliğe sahip tahmin ediciler kurmak amacıyla, yığın bilgisini veya büyük çaplı bir örnekten alınan bilgiyi tahminde kullanan önemli bir yöntemdir [Fuller, 2002].

Regresyon tahmin edicisi ile ilgili temel bilgi doğrusal modele dayalıdır. Klasik doğrusal model regresyon tahmininin temelini oluşturur. Doğrusal tahmin ediciler araştırma örneklemede önemli bir avantaj sağlar çünkü ağırlıklar bir kez hesaplandıktan sonra herhangi bir analiz değişkeni için de kullanılabilir [Fuller, 2002].

GREG ağırlıklandırma düzeltmesi, örneklemede standart regresyon tahmin edicisinden çıkarsanır [örneğin, Cochran, 1977, Bölüm 7'ye bakınız]. Yöntem regresyon tahmin edicisinde ağırlıkların bir değişikliği olarak kullanılan yardımcı değişkenler için düzeltmeyi içerir [Bethlehem ve Keller, 1987; Fuller, 2002; Kalton ve Flores-Cervantes, 2003].

$n$  çaplı bir  $s$  örneği  $N$  çaplı sonlu bir  $U$  yığından seçilmiş ve  $i$  inci birimin örneği seçilmiş olması olasılığı  $\pi_i$  ile gösteriliyor olsun. İlgilenilen değişken  $y$  ve  $i$  inci birimin değeri  $y_i$  ve  $y_i$  tüm  $i \in s$  için gözleniyor olsun.  $K$  sayıda  $x_1, x_2, \dots, x_K$  yardımcı değişken olsun ve  $i \in s$  için değerleri gözlemlensin.  $X = (X_1, X_2, \dots, X_K)'$ ,  $K$ -boyutlu  $x_1, x_2, \dots, x_K$  değişkenlerinin bilinen yığın toplamaları vektörünü gösteriyorken doğrusal regresyon modeli,

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_K x_{Ki} + \varepsilon_i \quad (4.38)$$

$i = 1, \dots, N$  için verilir. Burada  $\beta_1, \dots, \beta_K$  bilinmeyen model parametreleridir (regresyon katsayılarıdır).  $\varepsilon_i$ ,  $E(\varepsilon_i) = 0$  ve  $V(\varepsilon_i) = \sigma_i^2$  ile rasgele hatalardır,  $\varepsilon_i \sim NI(0, \sigma_i^2)$ ,  $i = 1, \dots, N$ .

$$\varepsilon_i = Y_i - \sum_{k=1}^K \beta_k x_{ki} = Y_i - \underline{X}_i' \underline{\beta}$$

Genel ağırlıklandırma teknikleri altında GREG tahmin edicisini vermeye çalışalım. En küçük kareler yönteminin uygulanmasıyla  $\underline{\beta}$  için optimum değer,

$$\underline{\beta} = \left( \sum_{i=1}^N \underline{X}_i \underline{X}_i' \right)^{-1} \left( \sum_{i=1}^N \underline{X}_i Y_i \right) \quad (4.39)$$

ve matris gösterimi ile,

$$\underline{\beta} = (X'X)^{-1} X'y$$

olarak belirlenir. Basit tesadüfi örnekleme altında  $\underline{\beta}$ 'nin tahmin edicisi,

$$\hat{\underline{\beta}} = \left( \sum_{i=1}^n \underline{x}_i \underline{x}_i' \right)^{-1} \left( \sum_{i=1}^n \underline{x}_i y_i \right) \quad (4.40)$$

şeklindedir.  $\bar{Y}$  yığın ortalamasının genel regresyon tahmin edicisi,

$$\bar{y}_{Reg} = \hat{\underline{\beta}}' \bar{\underline{X}} \quad (4.41)$$

belirlenir. Bethlehem ve Kersten (1985), genel regresyon tahmin edicisinin ancak ve ancak tüm  $i$ 'ler için  $\underline{c}\underline{X}_i = 1$  olmasını sağlayan  $\underline{c}$  gibi bir vektör mevcutsa yaklaşık yansız olduğunu göstermiştir. Buradan, Eş. 4.41 aşağıdaki gibi,

$$\bar{y}_{Reg} = \bar{y} - \hat{\underline{\beta}}' (\bar{\underline{x}} - \bar{\underline{X}}) = \bar{y}_{GREG} \quad (4.42)$$

şeklinde yeniden yazılabilir.

Y yığın toplamının regresyon tahmin edicisi matris yoluyla gösterimi,

$$\hat{Y}_{Reg} = \hat{Y}_{HT} + (X - \hat{X}_{HT})' \hat{\beta} = \hat{Y}_{GREG} \quad (4.43)$$

yapısındadır.  $\hat{Y}_{HT} = \hat{Y} = \sum_i^n W_i y_i$ ,  $y$ 'nin yığın toplamının HT tahmin edicisidir,  $W_i = 1/\pi_i$ . Ayrıca,  $\hat{X}_{HT} = \hat{X} = (\hat{X}_1, \hat{X}_2, \dots, \hat{X}_K)'$   $x_1, x_2, \dots, x_K$  değişkenlerinin yığın toplamı Horvitz-Thompson tahmin edicileri vektörüdür.

Yığın toplamı veya ortalamasının regresyon tahmin edicisini formüle etmenin pek çok yolu vardır. Literatürde farklı gösterimleri ve hesaplamaları yer almıştır. Bu tahmin edicinin, örnek birimlerine ağırlıklar atanmış olduğu gösterimi kullanılabilir.  $\bar{y}_{Reg}$  regresyon tahmin edicisi  $w_i$  düzeltilmiş ağırlıklarına dayalı tanımlanacak olursa

$$\bar{y}_{Reg} = \sum_{i=1}^n {}^r W_i y_i \quad (4.44)$$

şeklinde verilir. Bu durumda,  $i$  inci örnek birimine atanan  $w_i$  ağırlığı,

$${}^r W_i = \underline{x}_i' \left( \sum_{i=1}^n \underline{x}_i \underline{x}_i' \right) \bar{X} \quad (4.45)$$

formülü ile hesaplanır [Bethlehem ve Kersten, 1985].

Regresyon tahmin edicisinde Eş. 4.45'le tanımlanmış regresyon ağırlıkları çok büyük veya çok küçük hatta negatif değerli olabilmektedir. Deville ve Särndal (1992)'nin ayarlama tahmin edicileri kısıt koyarak regresyon ağırlıklarını sınırlandırmıştır. En küçük kareler regresyon tahmininin sonucunda oluşacak istenilmeyen negatif ağırlıkların olmasını önlemek amacıyla, ayarlama tahmin edicileri uygulanabilir.

Ayarlama tahmini işlemi ile ayarlama kısıtı altında hesaplanacak ağırlıkların örnekleme ağırlıklarının yakınında olması sağlanır. Bu özellikle pozitif ağırlıklar

elde edilmesini sağlamış olur. Regresyon tahmin edicisi ayarlama tahmin edicisinin özel bir durumu olarak verilebilir. Ayarlama tahmin edicileri, yardımcı değişken seçiminde regresyon tahminin özelliklerini kullanarak geliştirilebilir [Jayasuriya ve Valliant, 1995 ].

#### 4.3.4. Regresyon tahmin edicisinin ayarlamayla belirlenmesi

Ayarlama tahmin edicileri, örnekleme ağırlıkları ile ayarlama ağırlıkları arasındaki verilen uzaklığının minimize edilmesiyle belirlenmeye çalışılır. Uzaklık fonksiyonu için genelleştirilmiş en küçük kareler uzaklık fonksiyonu kullanıldığında, regresyon tahmin edicisi ayarlama tahmin edicisinin özel bir durumu olur [Jayasuriya ve Valliant, 1995 ].

Ayarlama yaklaşımı çerçevesinde genelleştirilmiş regresyon tahmini, yardımcı değişken bilgisini sistematik yolla tahmin aşamasına dahil etmiştir. Bu nedenle GREG tahminin ayarlama yöntemine benzemediğini belirtmek önemlidir. İki yaklaşım birbirinden farklıdır. Ayarlama tekniği, GREG tahmini için bir alternatif çıkarsamadır [Särndal, 2007].

$\hat{Y}_C$ , ayarlama tahmin edicisinde  ${}^cW_i$ , ayarlama ağırlıkları yerine yazıldığında,

$$\begin{aligned}\hat{Y}_C &= \sum_{i=1}^n W_i (1 + q_i x_i' \lambda) y_i \\ \hat{Y}_C &= \sum_{i=1}^n W_i y_i + (X - \hat{X})' \hat{B} \quad , \quad \hat{B} = T^{-1} \sum_{i=1}^n W_i q_i x_i y_i\end{aligned}\tag{4.46}$$

sonucuna ulaşılır.  $\hat{B}$ , çoklu regresyon katsayısının ağırlıklandırılmış tahmin edicisidir.

Uygulamalarda çoğunlukla  $1/q_i = 1$  alınması durumu hakimdir,  $q_i = 1$  durumu için, yığın toplamı için ayarlama ağırlıkları,

$${}^cW_i = W_i (1 + x_i' \lambda), \quad i=1, \dots, n\tag{4.47}$$

$$\lambda = T^{-1}(X - \hat{X})$$

$$T = \sum_{i=1}^n W_i x_i x_i'$$

Eş. 4.47'yle belirlenmiş olur. Yığın toplamının ayarlama tahmin edicisi,

$$\hat{Y}_C = \sum_{i=1}^n W_i y_i + (X - \hat{X})' \hat{B} = \hat{Y}_{GREG} \quad (4.48)$$

$$\hat{B} = T^{-1} \sum_{i=1}^n W_i x_i y_i$$

eşitliğiyle belirlenmiş olur. Ki-kare uzaklık fonksiyonu ve  $q_i = 1$  durumunda alınmasıyla ayarlama tahmin edicisi genelleştirilmiş regresyon tahmin edicisine dönüşmüştür. Ayrıca,

$$\hat{Y}_{GREG} = \hat{Y}_C = \sum_{i=1}^n {}_c W_i y_i = \hat{Y}_{HT} + (X - \hat{X}_{HT})' \hat{B} \quad (4.49)$$

şeklinde tanımlamak da mümkün.  $\hat{Y}_{HT}$  ve  $\hat{X}_{HT} = \hat{X}$  sırasıyla y ve x değişkenlerinin Horvitz-Thompson tahmin edicileridir. Böylece, ayarlama tahmin edicisi ile genelleştirilmiş regresyon (GREG) tahmin edicisi arasında bir bağlantı kurulmuş olur [Särndal, 2007].

#### 4.3.5. Yanlılık

Ayarlama tahmin edicisi yansız Horvitz-Thompson tahmin edicisine dayalı,

$$\hat{Y}_C = \hat{Y}_{HT} + \sum_{i=1}^n ({}_c W_i - W_i) y_i \quad (4.50)$$

yazılabilir. Buradan  $\hat{Y}_C$  ayarlama tahmin edicisinin yanı,

$$E(\hat{Y}_C) - Y = E\left(\sum_{i=1}^n ({}_C W_i - W_i) y_i\right)$$

biçimindedir. Tasarım yansızlığına yakın olmayı amaçlamak,

$$E\left(\sum_{i=1}^n ({}_C W_i - W_i) y_i\right) \approx 0$$

dağılımda yakınsama özelliğine haiz olunması istenir (y değişkeni ne olursa olsun). Yansızlığın sağlanması için ayarlama,  ${}_C W_i$  ile  $W_i$  arasında küçük ayrılıklar olması için çabalamalıdır.  ${}_C W_i$  ayarlama ağırlıkları genellikle tasarım yansız tahminler verir ve Horvitz-Thompson tahmin edicisinden daha küçük varyanslı oldukları belirlenmiştir [Särndal, 2007].

Ayarlama ağırlıklandırmasının bir özelliği olarak  $\{{}_C W_i \mid i \in s\}$  ile  $\{W_i \mid i \in s\}$  birbirine çok yakın olmalı ki farklılıkları minimize edilebilsin. Bunun sonucu olarak geniş çaplı bir örnek alınmasıyla  $\hat{Y}_C$ , y-değişkeninin ne olduğuna bağlı olmaksızın rasgeleleştirilmiş yansızlığa çok yakın olacaktır.

#### 4.3.6. $\hat{Y}_C$ 'nin varyans ve varyans tahmini

$\hat{Y}_{GREG}$ ,  $\hat{Y}_C$  'nin özel bir durumudur.  $\hat{Y}_C$  ayarlama tahmin edicisinin  $\hat{Y}_{GREG}$  regresyon tahmin edicisine asimptotik olarak denk olduğu görülmüştür. Dolayısıyla  $\hat{Y}_C$  'nin asimptotik varyansı (AV) GREG tahmin edicisinin varyansı ile aynı olup şöyledir [Wu, 2003]:

$$AV(\hat{Y}_C) = \sum_{j=1}^N \sum_{i=1}^N \{(\pi_{ij} - \pi_i \pi_j)({}_C W_i \varepsilon_i)({}_C W_j \varepsilon_j)\} = AV(\hat{Y}_{GREG}) \quad (4.51)$$

$W_i = 1/\pi_i$ ,  $\varepsilon_i = y_i - x_i' B$  dir. B, Eş. 4.39'u sağlar ve

$$\left(\sum_{i=1}^N q_i x_i x_i'\right)B = \sum_{i=1}^N q_i x_i y_i \quad (4.52)$$

$B$ , ağırlıklandırılmış en küçük kareler denklemini minimize eder,

$$\sum_{i=1}^N q_i (y_i - x_i' B)^2 = \sum_{i=1}^N q_i \varepsilon_i^2. \quad (4.53)$$

Eş. 4.51'de yer alan varyansı tahmin etmek için,  $E_i$  artıkları kullanılamaz, çünkü  $B$  bilinmemektedir.  $\hat{B}_{ws}$  tahmin edicisi, Eş. 4.53'le verilen  $SS_U$  bilinmeyen  $q_i \varepsilon_i^2$  değerinin yığın toplamıdır. Bu toplam değerinin ayarlanmış ağırlıklar tahmin edicisi,

$$SS_{sw} = \sum_{i=1}^n {}_c W_i q_i E_i^2$$

$$SS_{sw} \left( \sum_{i=1}^n {}_c W_i q_i x_i x_i' \right) \hat{B}_{ws} = \sum_{i=1}^n {}_c W_i q_i x_i y_i \quad (4.54)$$

şeklinde ve örnek-tabanlı normal eşitliğini sağlayan  $\hat{B}_{ws}$  tahmin edicisi ile minimize edilir. Örnek-tabanlı artıklar,  $e_i = y_i - x_i' \hat{B}_{ws}$  olarak hesaplanabilir.  $\hat{Y}_{GREG}$ 'nin varyans tahmin edicisi aşağıdaki gibi yazılır.

$$\hat{V}(\hat{Y}_{GREG}) = \sum_{j=1}^n \sum_{i=1}^n \left\{ \left( (\pi_{ij} - \pi_i \pi_j) / \pi_{ij} \right) ({}_c W_i e_i) ({}_c W_j e_j) \right\} \quad (4.55)$$

Ayarlama ağırlıkları  $\hat{V}(\hat{Y}_{GREG})$ 'de  $e_i$  artıklarını ağırlıklandırmak için kullanılmıştır. Tasarım-tabanlı bakış açısıyla ele alınacak olunursa  $e_i$  artıkları,  $W_i$  standart örnekleme ağırlıklarıyla basit bir şekilde ağırlıklandırılabilir ve tasarım-tutarlı bir varyans tahmin edicisi elde edilir.

Eş. 4.55 ile verilen varyans tahmini, yalnızca tasarım-tutarlı bir varyans tahmin edicisi değil aynı zamanda  $E_{\xi}(\hat{Y}_c - Y)^2$  model hata kareler ortalaması için yaklaşık model-yansız bir tahmin edicidir.

$\hat{Y}_{REG}$  tahmin edicisi ağırlıklandırma sınıfları içinde kolay uygulanabilir olması ve ayarlama dayalı belirlenen diğer tahmin edicilere asimptotik olarak denk olması nedeniyle önemlidir. Farklı  $F$  fonksiyonlarıyla belirlenen ayarlama tahmin edicileri aynı geniş-örnek varyansını gösterir. Dolayısıyla, Eş. 4.55 varyans tahmini, herhangi bir  $\hat{Y}_c$  tahmin edicisinin varyansının tahmini için kullanılabilir [Deville ve ark., 1993].

#### 4.3.7. Neden ayarlanmış ağırlıklar?

Ayarlama tahmin edicisi yönteminin aşağıda yer alan özellikleri taşıdığı savunulmaktadır:

1. Tutarlılığı sağlamak için ayarlama. Ayarlama çoğu kez tutarlı tahminler elde etmenin bir yolu olarak tanımlanmıştır. Burada tutarlılık, rasgelelik tutarlılığı anlamında olmayıp bilinen toplamlarla tutarlılık anlamındadır. Bu yöntemde,  $X = \sum_{i=1}^n x_i$  ayarlama kısıtını sağlayan bir ağırlıklandırma sistemi uygulanır [Estevao ve Särndal, 2000].

Ayarlama denklemleri, tutarlılığı ağırlık sistemi üzerine yükler, böylece yardımcı değişkenlere uygulandığında aynı yardımcı değişkenlerin bilinen toplamlarına uyumu (tutarlılığı) sağlanır. Bilinen veya tahmin edilen toplam artan doğruluğa ilave faydalar getirir. Düşük varyans ve/veya azaltılmış yanıtlamama yanı gibi [Särndal, 2007].

2. Temel ağırlıklara çok yakın olma.  $W_i = 1/\pi_i$ , temel örnekleme ağırlıklarının en önemli özelliği model yansız tahminler vermeleridir. Dolayısıyla, bu ağırlıklardan

olabilecek herhangi bir sapma az olmalıdır ki model yansızlık en azından yaklaşık veya asimptotik olarak söz konusu olsun [Estevao ve Särndal, 2000].

3. Yardımcı değişken toplamları üzerinde kontrol sağlama. Ayarlama ilgililenilen değişken ile ilişkili olduğu bilinen ne kadar çok yardımcı değişken bilgisi kullanılmışsa o kadar “iyi” düzeyde ağırlıklandırma sistemi yapılmış demektir. Ayarlama tahmin edicilerin duyarlılığı  $y$  değişkeni ile ilişkili olan çok sayıda (otokorelasyon sorunu olmayan ve yığın toplamı bilinen) yardımcı değişken kullanımıyla artış gösterecektir [Estevao ve Särndal, 2000].

4. Uygunluk ve şeffaflık için ayarlama. Harms ve Duchesne [2006]’ya göre, “Ayarlama yaklaşımı gerçek uygulamalardaki kullanımının popülaritesi artmaktadır, çünkü örnekleme ağırlıklarına ve doğal ayarlama kısıtlarına dayalı hesaplanan tahminleri yapmak ve yorumlamak kolaydır.” Bilinen toplamlar üzerine ayarlama açık ve doğal yolla belirlenir [Estevao ve Särndal, 2000].

Ayarlamanın basit biçimleri “doğal kısıtlar” dışında herhangi bir varsayıma başvurmaz. Önemli bir başka avantajı ise pek çok uygulamada ayarlamanın tüm çalışma değişkenlerine uygulanabilecek tek bir ağırlıklandırma sistemi sağlamasıdır. [Särndal, 2007].

Sonradan tabakalama ve tarama yöntemlerinde yardımcı değişkenlerin kategorik ya da kategorik hale dönüştürülmüş olması istenir. Regresyon tahmini yapılabilmesi için değişkenlerin nicel yapıda olması gereklidir. GREG ve ayarlama tahmin edicisinde ise tahmin modeline nitel veya nicel değişkenler dahil edilebilir. Bu nedenle ayarlama tahmin edicisi daha genel bir tahmin yöntemidir.

#### **4.5. Ayarlama Tahmin Edicisinde $\bar{Y}$ ve $Y$ ’nin Hájek Tahmin Edicisi**

Ayarlama tahmin edicisi Horvitz-Thompson yapısında verilen bir tahmin edicidir. Hájek tahmin edicisi yapısında ayarlama tanımlanabilir. Yığın ortalamasının ayarlama tahmin edicisi,

$$\bar{y}_{C,w} = \frac{\sum_{i=1}^n {}_cW_i y_i}{\sum_{i=1}^n {}_cW_i} \quad (4.56)$$

şeklinde Hájek yapısında belirlenebilir.  ${}_cW$ , bilinmeyen ayarlama ağırlıklarıdır ve yardımcı değişkene dayalı yapılan ağırlıklandırmayı ifade eder,  $i = 1, \dots, n$ . Burada  $\sum_{i=1}^n {}_cW = N$  dir.  $Y$  yığın toplamının Hájek yapısında ayarlama tahmin edicisi,

$$\hat{Y}_{C,w} = N \bar{y}_{C,w} \quad (4.57)$$

olur.

Verilen iki ağırlıklandırma kümesi  $\hat{X}_{HT} = \sum_{i=1}^n W_i x_i$  ile  $X$  arasındaki yakınlığı saptamaya yarayacak bir kriter  $G = G({}_cW_i, W_i)$  uzaklık kriteri için ki-kare uzaklık ölçümü kullanılabilir.  $\hat{X}_{HT}$  ile  $X$  arasındaki yakınlığı saptamada, toplam uzaklığı minimum yapılmasıyla ayarlama ağırlıkları ve  $\lambda$ : Lagrange çarpanı,

$${}_cW_i = W_i(1 + q_i x_i' \lambda), \quad i=1, \dots, n$$

$$\lambda = \left[ \sum_{i=1}^n W_i q_i x_i x_i' \right]^{-1} (X - \hat{X})$$

olarak belirlenir. Burada,  $q_i = 1$  alınmasıyla ayarlama denklemi GREG tahmin edicisine dönüşeceği belirtilmişti.

Deville ve Särndal (1992),  $\hat{X}$  yardımcı değişkenler örnek toplamaları vektörünü,  $j$  inci yardımcı değişken için  $\hat{X}_j = \sum_{i=1}^n W_i x_{ji}$  formülünden yararlanılarak belirlenmesini önermiştir.

$\hat{X}_j$ , bir Horvitz-Thompson tahmin edicisi yapısındadır. Ancak uygulamalı araştırmalarda Hájek yapısındaki tahmin edicilerin kullanımı önerilmekte ve bazı araştırmalarda yalnızca Hájek tahmin edicileri kullanılabilmektedir. Bu nedenle  $\hat{X}_j$  için Hájek tahmin edicisi tanımlanarak istenen tahminleri,

$${}_w\hat{X}_j = N \frac{\sum_{i=1}^n W_i x_{ji}}{\sum_{i=1}^n W_i} \quad (4.58)$$

ağırlıklandırılmış denklemlerle elde edilmesi önerilebilir.  ${}_w\hat{X} = ({}_w\hat{X}_1, {}_w\hat{X}_2, \dots, {}_w\hat{X}_K)'$ ,  $K$  sayıda yardımcı değişken için yığın toplamları Hájek tahmin edicileri vektörüdür.

Lagrange çarpanları vektörü ve  ${}_cW_i$  ayarlanmış ağırlıkları

$$\lambda = \left[ \sum_{i=1}^n W_i' x_i x_i' \right]^{-1} (X - {}_w\hat{X}) \quad (4.59)$$

$${}_cW_i = W_i' (1 + x_i' \lambda)$$

eşitlikleriyle elde edilir. Buradan,  $\bar{y}_{c,w}$  ve  $\hat{Y}_{c,w}$  tahmin edicileri Eş. 4.58 ve Eş. 4.59'a dayalı yukarıda tanımlanan sırasıyla Eş. 4.56 ve Eş. 4.57'den belirlenir.

#### 4.6. Aşamalı Ayarlama

Ayarlama tahmin edicisinde örnekleme tasarımı ve yardımcı değişkenlere dayalı bir çalışma modeli altında etkin tahminler yapılması amaçlanır.  $y$  ve  $x$ 'ler arasındaki ilişkiye dayalı modelle, yardımcı bilginin tahmin aşamasında en iyi nasıl kullanılabileceği saptanmaya çalışılır [Wu ve Sitter, 2001]. Ayarlama ağırlıklarının  $y_i$  gözlemlerinin doğrusal ağırlıkları olarak görebilmemize olanak sağlar [Deville ve Särndal, 1992].

Örnekleme arařtırmalarında genellikle  $y$  ilgilenilen deęiřkeniyle iliřkili çok sayıda yardımcı deęiřkene ait bilgiler derlenir. Regresyon tahmin edicisinin amacı yardımcı deęiřkenler yardımıyla ilgilenilen deęiřkenin parametreleri hakkında çıkarsamalar yapmaktır. Yardımcı deęiřkenlerin bağımsız olması ve  $y$ 'yi en iyi açıklayanların regresyon modeline alınarak parametre tahminlerinin yapılması amaçlanır.

Doęrusal regresyon modeli,  $y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_K x_{Ki} + \varepsilon_i$  biçiminde ele alınan modelde parametre tahmini yapılır.

$y$ 'yi en iyi açıklayan deęiřkenleri seçmek amacıyla başvuru olan yöntem aşamalı regresyon (stepwise) yöntemidir. Aşamalı regresyon yardımcı deęiřkenler arasından  $y$ 'yi açıklama niteliğine sahip olanların seçimini saęlayan bir yöntemdir. Aşamalı regresyonda veri kümesindeki deęiřkenler ile tüm olası modeller denenerek modeller oluřturulması yaklařımı kullanılabilir. Bu modeller arasında  $R^2$  deęeri en yüksek olan model uygun regresyon modeli olarak kabul edilir [Özdamar, 1999].

Ancak ayarlama tahmin edicisinde yardımcı deęiřkenler için bağımsız olma kořulu aranmamaktadır. Örneęin, Kalton ve Flores-Cervantes (2003) ve Creedy (2003) çalışmasında ayarlama aęırlıklandırması için kullanılan yardımcı deęiřkenlerin bağımsız olmadığı, güçlü iliřkili olduęu görölmektedir.

Uygulamalı arařtırmalarda ilgilenilen deęiřkeni açıklama niteliğine sahip çok sayıda deęiřken varken, deęiřkenlerin bağımsız olması kořulu nedeniyle modele çok az sayıda deęiřken alınmaktadır. Modele alınacak deęiřken seçimi kořullu olarak yapılabilir. Kořullu stratejileri tahminin doęruluęunu arttırır ve uç deęerler olmasını engeller [Wywiał, 2003].

Pek çok arařtırmada genellikle bir deęiřkene göre daęılımı incelenen örnek verisinin, başka deęiřkenlere dayalı incelemeleri de yapılmak istenilir. Böyle durumlarda kořullu daęılım kullanımı ve deęiřkenlerin tahmin modeline kořullu olarak aşamalı eklenmesi istenilen tahminlere daha iyi ulařmamızda katkı saęlayacaktır.

Ayarlama tahmin edicisinde koşullu dağılım özelliklerinden yararlanılabilir. Ağırlıklandırma yöntemlerinden ayarlama tahmin edicisi, yardımcı değişkenlerin modele aşamalı (koşullu) eklenmesine olanak sağlayarak yığın parametre tahminlerinin daha iyi tahminlerini sağlayabilmektedir. Diğer ağırlıklandırma yöntemlerinin yapısı ve formülasyonu gereği koşullu tahminler yapılması olanaklı değildir.

$y$  ilgilenilen değişkeniyle ( $y = f\{x_1, \dots, x_L\}$ ) ilişkili olduğu belirlenen  $L$  sayıda  $x_1, x_2, x_3, x_4, \dots, x_L$  yardımcı değişkenleri varken, bu yardımcı değişkenlere göre ayarlama ağırlıklandırması yapılmak istenilebilir. Aşamalı ayarlama yapılarak yardımcı değişkenlere dayalı ilgilenilen değişkene ilişkin yığın ortalaması veya toplamı yapılmak istenilebilir. Bu durum da değişkenler ayarlama denklemine aşamalı (koşullu) olarak eklenebilir.

Ayarlama tahmin edicisinde  $X_1, X_2, \dots, X_L$  yığın toplamlarının bilindiği varsayılır.

Birinci aşamada,  $x_1$  yardımcı değişkenine göre örnek dağılımının yığın dağılıma uyumu ayarlama tahmin edicisiyle incelenmek istenebilir. Buna göre  $x_1$  değişkenin  $X_1$  yığın toplamı bilindiği varsayılır.  ${}_cW_i$  ayarlama ağırlıkları,  $\hat{X}_1$  ile  $X_1$  arasındaki yakınlığı saptamada toplam uzaklığın minimum yapılmasıyla belirlenir. Burada,

$$X_1 = \sum_{i=1}^N x_{1i}$$

$$\hat{X}_1 = \sum_{i=1}^n {}_cW_i x_{1i}$$

dir ve ayarlama ağırlıkları

$${}_cW_i = W_i(1 + x'_i\lambda), \quad i = 1, \dots, n$$

$$\lambda = \left[ \sum_{i=1}^n W_i x_i x_i' \right]^{-1} (X_1 - \hat{X}_1)$$

$\bar{Y}$  yığının ortalamasının ayarlama tahmin edicisi,

$$\bar{y}_C = \sum_{i=1}^n {}_C W_i y_i = \bar{y}_{GREG}$$

olarak belirlenir.  $\bar{y}_C$ , tahmin edicisinde  ${}_C W_i$ , ayarlama ağırlıkları  $W_i$  örnekleme ağırlıklarına,  $x_1$  yardımcı değişkenine ve  $X_1$  yığın toplamı bilgisine dayalı belirlenmiştir,  $i = 1, \dots, n$ .

İkinci aşamada,  $x_1$ , yardımcı değişkenlerine göre ayarlama ağırlıklandırması yapıldığı bilinen veride,  $x_2$  yardımcı değişkenine göre ayarlama ağırlıklandırması yapılması isteniliyor olabilir. Bu durumda,  $f(x_2 \setminus x_1)$  koşullu dağılımı söz konusu demektir. Buradan,  $x_1$  değişkenine dayalı ayarlama ağırlıklandırması yapıldığı bilinen veride,  $x_2$  değişkenine dayalı koşullu ayarlama ağırlıklandırması yapılması önerilebilir.

Bu durumda, ikinci aşama ağırlıklandırması bir koşullu ağırlıklandırma olacaktır.  $x_2$  değişkenine dayalı ayarlama ağırlıklandırmadır ancak temel ağırlıklar olarak birinci aşamada  $x_1$ 'e dayalı belirlenen  ${}_C W_i$  ayarlama ağırlıklarını kullanılır. Buradan, ikinci aşama ayarlama ağırlıkları,  $\hat{X}_1$  ile  $X_1$  arasındaki yakınlığı saptamada toplam uzaklığın minimum yapılmasıyla belirlenir. Burada,

$$X_2 = \sum_{i=1}^N x_{2i}$$

$$\hat{X}_2 = \sum_{i=1}^n W_i x_{2i}$$

şeklindedir ve

$${}_c\tilde{W}_i = {}_cW_i(1 + x_i'\lambda) \quad i = 1, \dots, n \quad (4.60)$$

$$\lambda = \left[ \sum_{i=1}^n {}_cW_i x_i x_i' \right]^{-1} (X_2 - \hat{X}_2)$$

$x_1$  modeldeyken  $x_2$ 'ye göre koşullu ayarlama ile  $\bar{Y}$  yığın ortalamasının ayarlama tahmin edicisi,

$$\bar{y}_C = \sum_{i=1}^n {}_c\tilde{W}_i y_i = \bar{y}_{GREG} \quad (4.61)$$

eşitliğiyle belirlenir.  ${}_c\tilde{W}_i$ , ayarlama ağırlıkları  ${}_cW_i$  ağırlıklarına dayalı belirlenmiştir.  $x_2$  yardımcı değişkenine ve  $X_2$  yığın toplamı bilgisine dayalı belirlenir,  $i = 1, \dots, n$ .

Aşama sayısı daha da arttırılabilir veya her aşamada birden çok sayıda değişkene dayalı ayarlama yapılabilir. Yardımcı değişkenlerin ilgilenilen değişkenle ilişkili olması durumu mutlaka göz önünde bulundurulmalıdır.

#### 4.7. Ağırlıklandırma Yöntemlerini Kıyaslamada Varyans Şişme Faktörü

Ağırlıklandırma yöntemlerinin amacı, yanıtlanmama, yığın ile örnek arasındaki ayrılıkların ve kapsamamanın neden olacağı yanı azaltmaktır. Ağırlıklandırma yöntemleriyle veride düzeltmeler yapılırken ağırlıkların değişkenliğinde artışlar olmakta, dolayısıyla da araştırma tahminlerinin doğruluğunda azalma olmaktadır.

Doğrulukta bu kaybın yararlı bir ölçümü  $VIF = 1 + CV^2(W_j)$ 'dir. Ağırlıklandırma yöntemlerini kıyaslamada hesaplanan ağırlıkların değişkenlik ölçüsü olan  $VIF$ , varyans şişme faktörü (variance inflation factor) kullanılabilir.  $CV^2(W_j)$ , ağırlıkların değişim katsayısı olarak tanımlanmıştır [Kish, 1992; Kalton ve Flores-Cervantes, 2003].

Araştırma verisi analizlerinde ağırlıkların kullanılması tahminlerdeki yanlılığı gideriyor olmasına karşın, tahminlerin varyanslarında artışa yol açıyor olabilir. Bir yığın parametresinin tahmininde ağırlıkların kullanılmasının etkisi varyans şişme faktörü ( $VIF$ ),

$$VIF = n \frac{\sum_{j=1}^n W_j^2}{\left( \sum_{j=1}^n W_j \right)^2}$$

ile tanımlanır. Burada,  $n$  örnek çapı ve  $W_j$  tahminde kullanılan  $j$  inci örnekleme birimi ağırlığıdır (toplam atanan ağırlıktır).  $VIF$ , ağırlıkların değişim katsayısı olarak da aşağıdaki gibi yazılabilir [DHS, 2005]:

$$VIF = n \frac{\sum_{j=1}^n W_j^2}{\left( \sum_{j=1}^n W_j \right)^2} = 1 + CV^2(W_j) \quad (4.62)$$

$$CV^2(W_j) = \frac{n}{\sum_{j=1}^n W_j^2} \left\{ \sum_{j=1}^n W_j^2 - \frac{1}{n} \left( \sum_{j=1}^n W_j \right)^2 \right\} = \frac{\text{Ağırlıkların varyansı}}{(\text{Ağırlıkların ortalaması})^2}$$

Literatürde örnekleme, yanıtlanmama ve ayarlama ağırlıklandırmasının araştırma tahminlerinin yanında azalma sağlarken varyansta artışa yol açtığı belirtilmiştir. Kish (1992), varyanstaki orantısal artış için basit bir  $CV^2(W_j)$  formülünü vermiştir. Veride ağırlıklandırma yapılması  $\sigma^2$  birim varyanslarını  $\{1 + CV^2(W_j)\}\sigma^2$ 'ye arttırmakta veya duyarlı birimlerin sayısı  $n$ 'den  $n/(1 + cv^2)$ 'ye azaltmaktadır. Burada  $CV^2(W_j)$  görel kaybı göstermektedir.

$CV^2(W_j)$ , varyanstaki artış için iyi bir yaklaşımken hata kareler ortalamasındaki yan azalmasından oluşacak potansiyel azalmayı yansıtmamaktadır. Little ve Vartivarian (2005), çalışmalarında yanıtlanmama ağırlıklandırmasının varyansta artışla sonuçlanmadığını göstermişlerdir. Yanlılığı azaltmada yanıtlanmama ağırlıklandırması yapmanın varyansta artma değil azalma eğilimi gösterdiğini belirlemişlerdir.

## 5. UYGULAMA

Çalışmada Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü tarafından gerçekleştirilen 2003 Türkiye Nüfus ve Sağlık Araştırmasının (TNSA–2003) doğurgan yaşlardaki 15-49 yaşında evlenmiş kadınlar verisi kullanılmıştır.

TNSA–2003 ile doğurganlık, bebek ve çocuk ölümlülüğü ve gebeliği önleyici yöntem kullanımı gibi önemli demografik karakteristikler ile çeşitli sağlık göstergeleri için kabul edilebilir duyarlılık derecelerine sahip tahminlerin elde edilmesi amaçlanmıştır.

Evlenmiş Kadın Sorukağı; Temel özellikler, Doğurganlık, Evlilik, Gebeliği önleyici yöntemler, Anne sağlığı ve emzirme, Aşılanma ve sağlık, Doğurganlık tercihleri, Kocanın temel özellikleri, Kadın çalışması ve statüsü, Cinsel yoldan bulaşan hastalıklar ve AIDS, Annelerin ve çocukların antropometrik ölçümleri konularında bilgi toplamak üzere tasarlanmıştır.

TNSA–2003 ülkemizi kır-kent yerleşim yeri ve coğrafi bölgelere göre temsil eden 81 ilde, 700 yerleşim yerinde, 10836 hanede gerçekleştirilmiş olup bu hanelerde 8075 evlenmiş kadınla görüşülmüştür. Araştırmada 2000 Genel Nüfus Sayımından belirlenen  $N=12630510$  sayıdaki 15–49 yaşında evlenmiş kadınlar hedef yığınının,  $n=8447$  sayıda seçilen örnek birimlerinden  $n_R=8075$  sayıdaki kadından yanıt alınabilmiştir.

Bu bölümde 2003 Türkiye Nüfus ve Sağlık Araştırmasının (TNSA-2003) doğurgan yaşlardaki 15-49 yaşında evlenmiş kadınlar verisi üzerinden, 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısı ( $\bar{Y}$  'nın) tahmini yapılacaktır.  $\bar{Y}$  'nın tahmini için (i) örnekleme ağırlıklandırması, (ii) yanıtlanmama ağırlıklandırması, (iii) ayarlama ağırlıklandırması aşamalarında incelenecek ve sonuçları kıyaslanacaktır. Ayrıca,  $\bar{Y}$  'nın tahmini için aşamalı ayarlama tahmin edicisi incelenecek ve sonuçları yorumlanacaktır.

Bununla birlikte, çalışmada doğurgan yaştaki 15–49 yaşında evlenmiş kadınların yaşayan çocuk sayısının canlı doğum sayısına oranıyla ilgilenilmektedir.  $R = \text{Yaşayan çocuk sayısı} / \text{Canlı doğum sayısının tahmininde}$ , bilinen  $r_b$  bileşik oran tahmin edicisi ve Ayhan (2003)'ün önerdiği  ${}_A r_b$  bileşik oran tahmin edicisi belirlenecek elde edilen sonuçlar kıyaslanacaktır.

Çalışmada TNSA–2003 verisine uygulanacak ağırlıklandırma aşamalarından ve ağırlıklandırma yöntemlerinden elde edilen tahminlerin  $\bar{Y}$  ve  $R$  gerçek değerlerini ne kadar iyi tahmin edebildikleri araştırılacaktır.

2000 Genel Nüfus Sayımı verilerinden 15-49 yaşında evlenmiş kadınlar için,

- $\bar{Y} = 2,59$  (*kadın başına düşen ortalama canlı doğan çocuk sayısı*)
- $R = 0,929253$  (*Yaşayan çocuk sayısı/Canlı doğum sayısının*)

belirlenmiştir.

TNSA–2003'ün yerleşim yerleri hakkındaki ilk bilgiler 2000 Genel Nüfus Sayımı sonuçlarından elde edilmiştir. 2000 Genel Nüfus Sayımı sonuçlarından tüm yerleşim yerlerinin (il ve ilçe merkezleri, bucak ve köyler) nüfus büyüklüklerinin ve hanehalkı sayılarının bilgisayar ortamında kullanılmasına olanak veren bir liste sağlamıştır.

Araştırmanın örnekleme tasarımı ve büyüklüğü, Türkiye geneli, kentsel ve kırsal alanlar ve beş demografik bölge (Batı, Güney, Orta, Kuzey ve Doğu) için analiz yapmaya imkan vermektedir. Araştırmanın örneği, 2002 yılının ikinci yarısı içinde benimsenen 12 coğrafi bölge (NUTS 1) için yapılmasına da olanak vermektedir. Bu bölgeler arasında yer alan iki bölge; İstanbul ve Güney Doğu Anadolu Projesi (GAP) bölgeleri de, sahip oldukları özel konumları nedeniyle örnekleme tasarımında büyük gözlem sayılarıyla temsil edilmişlerdir. Böylece yerleşim birimleri 40 tabakaya gruplanmıştır.

Çizelge 5.1. Bölge, NUTS 1 Bölgesi, Kır-kent, yerleşim yeri tipi ve illere göre tabaka listesi, Türkiye 2003

| Tabaka (h) | Bölge | NUTS 1 Bölge      | Tip                    | İl                                                                               |
|------------|-------|-------------------|------------------------|----------------------------------------------------------------------------------|
| 1          | Batı  | İstanbul          | Kent/Metropol/Düzensiz | İstanbul                                                                         |
| 2          | Batı  | İstanbul          | Kent/Metropol/Düzenli  | İstanbul                                                                         |
| 3          | Batı  | İstanbul          | Kent                   | İstanbul                                                                         |
| 4          | Batı  | İstanbul          | Kır                    | İstanbul                                                                         |
| 5          | Batı  | Batı Marmara      | Kent                   | Edirne, Kırklareli, Tekirdağ, Balıkesir,                                         |
| 6          | Batı  | Batı Marmara      | Kır                    | Edirne, Kırklareli, Tekirdağ, Balıkesir,                                         |
| 7          | Batı  | Ege               | Kent/Metropol          | İzmir                                                                            |
| 8          | Batı  | Ege               | Kent                   | İzmir, Aydın, Denizli, Muğla, Manisa                                             |
| 9          | Batı  | Ege               | Kır                    | İzmir, Aydın, Denizli, Muğla, Manisa                                             |
| 10         | Orta  | Ege               | Kent                   | Afyon, Kütahya, Uşak                                                             |
| 11         | Orta  | Ege               | Kır                    | Afyon, Kütahya, Uşak                                                             |
| 12         | Batı  | Doğu Marmara      | Kent/Metropol          | Bursa                                                                            |
| 13         | Batı  | Doğu Marmara      | Kent                   | Bursa                                                                            |
| 14         | Batı  | Doğu Marmara      | Kır                    | Bursa                                                                            |
| 15         | Batı  | Doğu Marmara      | Kent/Deprem            | Kocaeli, Sakarya, Yalova                                                         |
| 16         | Batı  | Doğu Marmara      | Kır/Deprem             | Kocaeli, Sakarya, Yalova                                                         |
| 17         | Orta  | Doğu Marmara      | Kent                   | Bilecik, Eskişehir                                                               |
| 18         | Orta  | Doğu Marmara      | Kır                    | Bilecik, Eskişehir                                                               |
| 19         | Orta  | Doğu Marmara      | Kent/Deprem            | Bolu, Düzce                                                                      |
| 20         | Orta  | Doğu Marmara      | Kır/Deprem             | Bolu, Düzce                                                                      |
| 21         | Orta  | Batı Anadolu      | Kent/Metropol          | Ankara                                                                           |
| 22         | Orta  | Batı Anadolu      | Kent                   | Ankara, Konya, Karaman                                                           |
| 23         | Orta  | Batı Anadolu      | Kır                    | Ankara, Konya, Karaman                                                           |
| 24         | Güney | Akdeniz           | Kent/Metropol          | Adana                                                                            |
| 25         | Güney | Akdeniz           | Kent                   | Antalya, Burdur, Isparta, Adana, İçel, Hatay, K. Maraş, Osmaniye                 |
| 26         | Güney | Akdeniz           | Kır                    | Antalya, Burdur, Isparta, Adana, İçel, Hatay, K. Maraş, Osmaniye                 |
| 27         | Orta  | Orta Anadolu      | Kent                   | Kırşehir, Nevşehir, Niğde, Aksaray, Kırıkkale, Kayseri, Sivas, Yozgat            |
| 28         | Orta  | Orta Anadolu      | Kır                    | Kırşehir, Nevşehir, Niğde, Aksaray, Kırıkkale, Kayseri, Sivas, Yozgat            |
| 29         | Kuzey | Batı Karadeniz    | Kent                   | Zonguldak, Bartın, Karabük, Kastamonu, Sinop, Samsun                             |
| 30         | Kuzey | Batı Karadeniz    | Kır                    | Zonguldak, Bartın, Karabük, Kastamonu, Sinop, Samsun                             |
| 31         | Orta  | Batı Karadeniz    | Kent                   | Çankırı, Amasya, Çorum, Tokat                                                    |
| 32         | Orta  | Batı Karadeniz    | Kır                    | Çankırı, Amasya, Çorum, Tokat                                                    |
| 33         | Kuzey | Doğu Karadeniz    | Kent                   | Artvin, Giresun, Gümüşhane, Ordu, Rize, Trabzon                                  |
| 34         | Kuzey | Doğu Karadeniz    | Kır                    | Artvin, Giresun, Gümüşhane, Ordu, Rize, Trabzon                                  |
| 35         | Doğu  | Kuzeydoğu Anadolu | Kent                   | Erzincan, Erzurum, Bayburt, Ağrı, Kars, Ardahan, Iğdır                           |
| 36         | Doğu  | Kuzeydoğu Anadolu | Kır                    | Erzincan, Erzurum, Bayburt, Ağrı, Kars, Ardahan, Iğdır                           |
| 37         | Doğu  | Ortadoğu Anadolu  | Kent                   | Bingöl, Elazığ, Malatya, Tunceli, Bitlis, Hakkâri, Muş, Van                      |
| 38         | Doğu  | Ortadoğu Anadolu  | Kır                    | Bingöl, Elazığ, Malatya, Tunceli, Bitlis, Hakkâri, Muş, Van                      |
| 39         | Doğu  | Güneydoğu Anadolu | Kent                   | Adıyaman, Gaziantep, Kilis, Diyarbakır, Şanlıurfa, Mardin, Siirt, Batman, Şırnak |
| 40         | Doğu  | Güneydoğu Anadolu | Kır                    | Adıyaman, Gaziantep, Kilis, Diyarbakır, Şanlıurfa, Mardin, Siirt, Batman, Şırnak |

(Kaynak: TNSA-2003 araştırması)

### 5.1. Örneklemeye Ağırlıklandırması

TNSA–2003’ün örnek seçiminde ağırlıklı, çok aşamalı, tabakalı küme örnekleme yaklaşımı kullanılmıştır. Araştırmanın hedef örneğinin dağılımı, 2000 Genel Nüfus Sayımı sonuçlarına dayanmaktadır. TNSA-2003’ün örnek seçimi üç aşamada gerçekleştirilmiştir.

Örnek seçiminin ilk aşamasında yerleşim birimleri her bir 40 tabaka için gruplanmış ve 2000 Genel Nüfus Sayımı esas alınarak oluşturulan bu yerleşim birimleri listelerinden nüfus büyüklüklerine olasılıklı orantılı olarak basit tesadüfi seçim yapılmıştır. Seçimin bu aşaması sonucunda TNSA–2003 örneğini oluşturan yerleşim birimlerini ve buralardan seçilecek küme sayılarını gösteren bir liste oluşmuştur.

Örnek seçiminin ikinci aşaması için, 2000 yılında Devlet İstatistik Enstitüsü (DİE) tarafından, belediyesi olan yerleşim yerlerinde yapılan ve 2002 yılında güncellenen bina sayımı verileri kullanılmıştır.

Güncellenmiş hanehalkı listeleri kullanılarak sistematik tesadüfi örnekleme yoluyla her küme içinden sabit sayıda hanehalkı seçilmiştir. Seçilen hanelerde genellikle yaşayan ve/veya görüşme tarihinden bir gece önce bu hanelerde kalmış olan 15–49 yaşlarında ve en az bir kez evlenmiş tüm kadınlar araştırma kapsamında Evlenmiş Kadın SoruKağıdı uygulaması için kabul edilmişlerdir.

$W_h$ , örnekleme ağırlığı seçilen örnekleme yöntemine göre örnekleme birimi seçim olasılığının tersine eşittir. TNSA–2003 araştırmasında herhangi bir tabakaya verilen ağırlık, o tabakadaki örnek sayısını belirlemek üzere kullanılan örnek oranının tersidir.

$W_h = 1 / f_h$  h inci tabaka örnekleme ağırlığı ve  $f_h$ , h inci tabaka örnek oranıdır.  $f_h = p_{1h} * p_{2h} * p_{3h}$  tabakadaki bütün seçim aşamalarındaki seçilme olasılıklarının çarpımıdır.  $p_{ih}$ , h tabakası için i inci örnekleme aşamasındaki örnekleme biriminin

seçilme olasılığıdır,  $h = 1, \dots, H$ ,  $H$ : tabaka sayısıdır. TNSA–2003 araştırmasının  $W_h$  örnekleme ağırlıkları Çizelge 5.2’deki gibi belirlenmiştir.

Çizelge 5.2. Tabakalara göre kadınlar için örnekleme ağırlıkları, yanıt oranları, yanıtlayıcı örnek çapları, düzeltilmiş örnekleme ağırlıkları ve nihai ağırlıklar, TNSA- 2003

| Tabaka | Örnekleme oranının tersi<br>$W_h$ | Hanehalkı<br>$1/r_h^{hane}$ | Kadınlar<br>$1/r_h^{kadin}$ | $n_{hR}$ | $W'_h$  | $(W'_h)^s$ | Nihai kadın ağırlığı<br>$(W'_h)^s \times 1000000$ |
|--------|-----------------------------------|-----------------------------|-----------------------------|----------|---------|------------|---------------------------------------------------|
| 1      | 1160555 / 960                     | 891 / 779                   | 672 / 630                   | 630      | 1708,8  | 1,076474   | 1076474                                           |
| 2      | 1587651 / 960                     | 870 / 682                   | 478 / 449                   | 449      | 2602,12 | 1,659981   | 1659981                                           |
| 3      | 24989 / 100                       | 68 / 63                     | 52 / 50                     | 50       | 324,996 | 0,272980   | 272980                                            |
| 4      | 76858 / 60                        | 46 / 46                     | 35 / 34                     | 34       | 1209,74 | 0,962433   | 962433                                            |
| 5      | 469931 / 500                      | 410 / 391                   | 285 / 269                   | 269      | 2455,64 | 0,802196   | 802196                                            |
| 6      | 362247 / 240                      | 220 / 218                   | 119 / 115                   | 115      | 5688,92 | 1,150401   | 1150401                                           |
| 7      | 685892 / 400                      | 348 / 300                   | 195 / 183                   | 183      | 1680,23 | 1,546953   | 1546953                                           |
| 8      | 686133 / 150                      | 144 / 137                   | 96 / 94                     | 94       | 1333,83 | 3,583791   | 3583791                                           |
| 9      | 667273 / 240                      | 211 / 204                   | 139 / 135                   | 135      | 1655,51 | 2,305124   | 2305124                                           |
| 10     | 202772 / 150                      | 129 / 127                   | 94 / 89                     | 89       | 5173,43 | 1,058475   | 1058475                                           |
| 11     | 211704 / 60                       | 48 / 47                     | 50 / 48                     | 48       | 2235,38 | 2,739621   | 2739621                                           |
| 12     | 352876 / 400                      | 348 / 300                   | 225 / 200                   | 200      | 598,79  | 0,840259   | 840259                                            |
| 13     | 129118 / 100                      | 83 / 75                     | 46 / 46                     | 46       | 2734,71 | 1,042909   | 1042909                                           |
| 14     | 109307 / 60                       | 33 / 33                     | 27 / 26                     | 26       | 3442,6  | 1,841054   | 1841054                                           |
| 15     | 377921 / 100                      | 90 / 86                     | 70 / 62                     | 62       | 846,61  | 3,259059   | 3259059                                           |
| 16     | 148605 / 60                       | 56 / 56                     | 39 / 38                     | 38       | 1587,51 | 1,855263   | 1855263                                           |
| 17     | 182284 / 100                      | 86 / 85                     | 68 / 65                     | 65       | 1305,67 | 1,408203   | 1408203                                           |
| 18     | 65446 / 60                        | 45 / 45                     | 21 / 21                     | 21       | 867,53  | 0,796109   | 796109                                            |
| 19     | 47999 / 100                       | 80 / 77                     | 57 / 55                     | 55       | 1349,47 | 0,377212   | 377212                                            |
| 20     | 83237 / 60                        | 55 / 55                     | 44 / 43                     | 43       | 641,78  | 1,036076   | 1036076                                           |
| 21     | 915073 / 500                      | 451 / 386                   | 287 / 260                   | 260      | 513,44  | 1,722755   | 1722755                                           |
| 22     | 431779 / 150                      | 128 / 124                   | 99 / 99                     | 99       | 945,73  | 2,168697   | 2168697                                           |
| 23     | 298404 / 240                      | 173 / 172                   | 116 / 107                   | 107      | 946,82  | 1,130884   | 1130884                                           |
| 24     | 276431 / 400                      | 361 / 349                   | 276 / 270                   | 270      | 1527,76 | 0,533328   | 533328                                            |
| 25     | 1052242 / 900                     | 808 / 734                   | 593 / 557                   | 557      | 1826,15 | 1,028638   | 1028638                                           |
| 26     | 681896 / 540                      | 470 / 446                   | 302 / 286                   | 286      | 3430,47 | 1,085906   | 1085906                                           |
| 27     | 523267 / 500                      | 457 / 438                   | 354 / 343                   | 343      | 4348,88 | 0,822517   | 822517                                            |
| 28     | 373756 / 240                      | 210 / 205                   | 162 / 159                   | 159      | 2191,87 | 1,186317   | 1186317                                           |
| 29     | 336258 / 500                      | 427 / 395                   | 275 / 267                   | 267      | 2945,04 | 0,546506   | 546506                                            |
| 30     | 318422 / 240                      | 207 / 204                   | 156 / 153                   | 153      | 1263,75 | 1,001856   | 1001856                                           |
| 31     | 224473 / 200                      | 180 / 176                   | 138 / 136                   | 136      | 1644,66 | 0,850111   | 850111                                            |
| 32     | 201222 / 90                       | 82 / 82                     | 60 / 59                     | 59       | 1570,77 | 1,659488   | 1659488                                           |
| 33     | 310851 / 600                      | 497 / 474                   | 362 / 355                   | 355      | 1628,01 | 0,404297   | 404297                                            |
| 34     | 349165 / 240                      | 203 / 199                   | 136 / 126                   | 126      | 1883,16 | 1,169152   | 1169152                                           |
| 35     | 212359 / 500                      | 462 / 452                   | 392 / 384                   | 384      | 1590,35 | 0,323444   | 323444                                            |
| 36     | 218260 / 240                      | 200 / 199                   | 158 / 151                   | 151      | 2634,27 | 0,797725   | 797725                                            |
| 37     | 371366 / 500                      | 478 / 449                   | 383 / 371                   | 371      | 1855,92 | 0,595771   | 595771                                            |
| 38     | 257644 / 240                      | 227 / 220                   | 208 / 195                   | 195      | 1108,02 | 0,862345   | 862345                                            |
| 39     | 756933 / 1000                     | 922 / 877                   | 762 / 742                   | 742      | 1368,89 | 0,596458   | 596458                                            |
| 40     | 356146 / 480                      | 455 / 449                   | 416 / 403                   | 403      | 899,22  | 0,566475   | 566475                                            |

(Kaynak: TNSA-2003 araştırması)

Seçilen örnekte yanıt alınamayan birimlerin olması durumunda, yanıtlayanların yanıtlamayanları temsil edebilmesi için  $W_h$  örnekleme ağırlıklarında yanıtlamama düzeltmesi yapılmalıdır.

TNSA–2003 araştırmasında 2000 Genel Nüfus Sayımında belirlenen  $N= 12630510$  sayıdaki 15–49 yaşında evlenmiş kadınlar hedef yığınının ağırlıklı çok aşamalı tabakalı küme örnekleme yöntemiyle  $n=8447$  sayıda seçilen örnek birimlerinden  $n_R=8075$  sayıda kadından yanıt alınmıştır.

$W_h$ , örnekleme ağırlıkları için tabakalarda yanıtlamama düzeltmesi yapılır,  $h = 1, \dots, H$ . Yanıtlanamama için düzeltilmiş ağırlıklar,  $W_h$  örnekleme ağırlığının hane halkı ve kadın düzeyi yanıt oranlarının tersiyle çarpılarak  $(W_h (1/r_h^{hane})(1/r_h^{kadın}))$  belirlenir.

Ancak düzeltilmiş örnekleme ağırlıkları toplamının, yığın toplamına eşit olmasını sağlamak için düzeltilmiş örnekleme ağırlıkları  $(W_h (1/r_h^{hane})(1/r_h^{kadın}))$ 'ler ile yani,

$$N / \left\{ \sum_{h=1}^H \sum_{i=1}^{nhR} W_h (1/r_h^{hane})(1/r_h^{kadın}) \right\} = 12630510 / 10901679 = 1,158584$$

ile çarpılır. Böylece  $W'_h$  düzeltilmiş örnekleme ağırlıkları yukarıda Çizelge 5.2'deki gibi belirlenmiş olur. Mesela, birinci tabakanın ağırlığı,

$$W'_1 = (1160555 / 960)(891 / 779)(672 / 630) 1,158584 = 1708,799$$

olarak elde edilir.

Araştırmalarda önerilen ve çeşitli avantajları olduğu düşünülen bir işlemde ağırlıkların standartlaştırılmasıdır. TNSA–2003 araştırmasında  $(W'_h)^s$  standartlaştırılmış ağırlıkları ve nihai kadın (örnekleme) ağırlıkları Çizelge 5.2'deki gibi verilmiştir.  $(W'_h)^s$  standartlaştırılmış ağırlıkları,

$$(W'_h)^s = W'_h \frac{\sum_{h=1}^H n_{hR}}{\sum_{h=1}^H W'_h n_{hR}}$$

$W'_h$  ağırlıkların  $T = \sum_{h=1}^H n_{hR} / \sum_{h=1}^H W'_h n_{hR}$  faktörüyle çarpılmasıyla bulunur. Tüm örnek birimleri üzerinden standartlaştırılmış ağırlıklar toplamı yanıtlayanlar örneğinin çapına eşittir [ORC Macro, 1996].

Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü nihai kadın ağırlıklarını,  $(W'_h)^s$  standartlaştırılmış ağırlıklarını 1000000 ile çarparak belirlemiştir. Veride tam sayılı ağırlıklar olması amaçlanmıştır.

Çizelge 5.3'te TNSA–2003 ham verisi özet olarak sunulmuştur. TNSA–2003 verisinde nihai kadın ağırlıkları verilmiştir.  $W'_h$  örnekleme ağırlıkları tez çalışması kapsamında çizelgeye aktarılmıştır. Çalışma kapsamında yığın ortalaması ve oranı tahminlerinin  $W'_h$  ağırlıklarına dayalı hesaplamaları yapılacaktır.

Aşağıda sırasıyla  $\bar{Y}$  yığın ortalaması (15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısı) için örnekleme, yanıtlanmama ve ayarlama ağırlıklandırmalarında incelenecek, aşamalı ayarlama ağırlıklandırması uygulanarak tahmin sonuçları kıyaslanacaktır.

Daha sonra  $R$  yığın oranı (15–49 yaşında evlenmiş kadınların yaşayan çocuk sayısının canlı doğum sayısına oranı) tahmini incelenecektir.

Çizelge 5.3. TNSA–2003 doğurgan yaşlardaki 15-49 yaşında evlenmiş kadınlar verisi

|          | 1    | 2    | 3            | 4                    | 5        | 6                                      | 7                                    | 8                                | 9                                   | 10                                   | 11                                  | ... | 4900         |
|----------|------|------|--------------|----------------------|----------|----------------------------------------|--------------------------------------|----------------------------------|-------------------------------------|--------------------------------------|-------------------------------------|-----|--------------|
| Birim no | küme | hane | Tanım kümesi | Nihai kadın ağırlığı | $W'_h$   | ( $\mathcal{Y}$ )<br>Canlı doğan çocuk | ( $\mathcal{X}_1$ )<br>Yerleşim yeri | ( $\mathcal{X}_2$ )<br>Yaş grubu | ( $\mathcal{X}_3$ )<br>Eğitim grubu | ( $\mathcal{X}_4$ )<br>Yaşayan çocuk | ( $\mathcal{X}_5$ )<br>Eşin eğitimi | ... | Pikniğe git. |
| 1        | 101  | 4    | 1            | 1076474              | 1708,8   | 2                                      | 1                                    | 3                                | 1                                   | 2                                    | 1                                   | ... | 1            |
| 2        | 101  | 6    | 1            | 1076474              | 1708,8   | 1                                      | 1                                    | 6                                | 2                                   | 1                                    | 2                                   | ... | 1            |
| 3        | 101  | 7    | 1            | 1076474              | 1708,8   | 3                                      | 1                                    | 4                                | 1                                   | 3                                    | 2                                   | ... | 1            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| 2010     | 278  | 5    | 2            | 1659981              | 2602,12  | 1                                      | 1                                    | 6                                | 3                                   | 1                                    | 5                                   | ... | 1            |
| 2011     | 278  | 6    | 2            | 1659981              | 2602,12  | 2                                      | 1                                    | 5                                | 3                                   | 2                                    | 2                                   | ... | 0            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| 4232     | 2701 | 1    | 6            | 1130884              | 1570,77  | 4                                      | 1                                    | 2                                | 1                                   | 3                                    | 4                                   | ... | 0            |
| 4233     | 2701 | 2    | 6            | 1130884              | 1570,77  | 3                                      | 1                                    | 4                                | 1                                   | 3                                    | 3                                   | ... | 0            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| 5001     | 3001 | 1    | 15           | 822517               | 1305,67  | 3                                      | 2                                    | 2                                | 2                                   | 3                                    | 2                                   | ... | 1            |
| 5002     | 3001 | 2    | 15           | 822517               | 1305,67  | 3                                      | 2                                    | 3                                | 3                                   | 3                                    | 4                                   | ... | 1            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| 6554     | 301  | 1    | 33           | 323444               | 513,436  | 2                                      | 1                                    | 2                                | 1                                   | 1                                    | 1                                   | ... | 1            |
| 6555     | 301  | 2    | 33           | 323444               | 513,436  | 1                                      | 1                                    | 2                                | 0                                   | 0                                    | 1                                   | ... | 1            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| .        | .    | .    | .            |                      | .        | .                                      | .                                    | .                                | .                                   | .                                    | .                                   | .   | .            |
| 8074     | 4032 | 14   | 40           | 776,1402             | 765,9054 | 5                                      | 2                                    | 3                                | 0                                   | 5                                    | .                                   | ... | 0            |
| 8075     | 4032 | 15   | 40           | 776,1402             | 765,9054 | 8                                      | 2                                    | 5                                | 0                                   | 6                                    | 1                                   | ... | 1            |

(Kaynak: Hacettepe Üniversitesi Nüfus Etütleri Enstitüsü)

### 5.1.1. Yanıtlayıcı ortalaması

Yığından tesadüfi olarak alınan  $n$  çaplı örnekte  $n_R$  sayıda yanıtlayan birim olduğunda sadece yanıtlayan kadın örneğine dayalı yığın ortalamasının ağırlıksız tahmin edicisi,

$$\bar{y}_R = \frac{\sum_{i=1}^{8075} y_i}{8075} = 2,78$$

şeklinde belirlenir. Buradan, TNSA–2003 araştırmasından Türkiye’de *15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısının* bir tahmini 2,78 olarak belirlenmiştir.  $\bar{y}_R$  yığın ortalamasının ağırlıksız tahmin edicisinin varyansının tahmini BTÖ tasarımı varsayımı altında,

$$\hat{V}(\bar{y}_R) = \frac{N-n}{N} \frac{s^2}{n} = \frac{12630510-8075}{12630510} \frac{4,741}{8075} = 0,000587$$

elde edilir.

### 5.1.2. Yığın ortalamasının Horvitz-Thompson tahmin edicisi

TNSA–2003 araştırmasının ham verisinde  $W_h'$  düzeltilmiş örnekleme ağırlıklarına dayalı olarak  $\bar{Y}$  yığın ortalamasının Horvitz-Thompson tahmin edicisi,

$$\bar{y}_{HT} = \frac{1}{N} \sum_{h=1}^H \sum_{i=1}^{n_{hR}} W_h' y_{hi} = \frac{1}{12630510} 33214885,414 = 2,6297$$

elde edilmiştir.

$\bar{y}_{HT}$  ’nin bir varyans tahmin edicisi, Lu ve Gelman (2003)’ün önermiş olduğu

$$\hat{V}_{HT}(\hat{\theta}) = \sum_{i=1}^n W_i^2 (y_i - \hat{\theta})^2$$

eşitliğinden hesaplanabilir.

Ancak  $\sum_{i=1}^n W_i = 1$  olması şartına dikkat edilmelidir. Ağırlıklar toplamı 1'e eşit değil ise  $W_i$  'ler  $W_i / \sum_{i=1}^n W_i$  ile yeniden belirlenerek  $\hat{V}_{HT}(\hat{\theta})$  formülünde kullanılmalıdır. Aşağıda verilen ağırlıklandırma aşamalarının tümünde ağırlıklar toplamı 1 olması sağlanarak varyans tahminleri yapılmıştır.

Buradan,  $\bar{y}_{HT}$  'nin varyansının bir tahmini

$$\hat{V}_{HT}(\bar{y}_{HT}) = \sum_{h=1}^H \sum_{i=1}^{n_{hR}} W_h'^2 (y_{hi} - \bar{y}_{HT})^2 = 0,000547$$

olarak elde edilmiştir. Bu varyans eşitliğinde  $W_i'$  ağırlıkları için  $\sum_{i=1}^n W_i' = 1$  olması sağlanarak sonuca ulaşılmıştır.

### 5.1.3. Yığın ortalamasının Hájek tahmin edicisi

Yığın ortalaması Hájek ağırlıklandırılmış tahmin edicisi,

$$\bar{y}_w = \frac{\sum_{h=1}^H W_h' y_{hR}}{\sum_{h=1}^H W_h' n_{hR}} = \frac{1}{12630510} 33214885,414 = 2,6297$$

olarak belirlenir.  $W_h'$ , düzeltilmiş örnekleme ağırlıkları kullanılarak yığın ortalamasının HA tahmin edicisi belirlenmiştir.  $\bar{y}_{HT}$  ve  $\bar{y}_w$  tahmin edicilerinin aynı sonucu verdiği görülmektedir. Her iki tahmin edicinin paydasında  $N = \sum_{h=1}^H W_h' n_{hR}$  olduğu için sonuçları da aynı olmuştur.

$\bar{y}_w$  tahmin edicisinin varyansının bir tahmini,

$$\hat{V}_{HT}(\bar{y}_w) = \sum_{h=1}^H \sum_{i=1}^{n_{hR}} W_h'^2 (y_{hi} - \bar{y}_{HT})^2 = 0,000547 = \hat{V}_{HT}(\bar{y}_{HT})$$

ile elde edilir.  $\bar{y}_{HT}$  'nin varyans tahmin edicisine eşittir.

Horvitz-Thompson tahmin edicisinde ağırlıklar toplamının N yığın çapına eşit olması gereklidir. Diğer bir durumda Horvitz-Thompson tahmin edicisiyle yığın ortalaması veya toplamının tahminleri yapılamayacaktır.

TNSA–2003'ün ham verisinde  $W_h'$  ağırlıkları değil, nihai örnek ağırlıkları verilmiştir. Araştırmacılar TNSA–2003'ün ham verisiyle tahminler yapmak istediklerinde nihai örnek ağırlıklarını kullanacaklardır. Ancak araştırmanın nihai örnek ağırlıkları toplamı yığın çapına eşit olmadığı için HT tahmin edicisi kullanılamayacaktır. Araştırmacılar Hájek tahmin edicisini kullanarak istenen yığın ortalaması ve toplamı tahminlerini yapmaları önerilebilir.

Bu araştırmada  $W_h'$  düzeltilmiş örnekleme ağırlıkları,  $(W_h')^s$  standartlaştırılmış nihai ağırlıklar veya nihai ağırlıklardan herhangi birinin kullanılması durumunda Hájek tahmin edicisinin yığın ortalaması ve toplamının iyi tahminlerini sağladığı belirlenmiştir. Mesela,  $(W_h')^s$  standartlaştırılmış nihai ağırlıklar veya  $(W_h')^s \times 1000000$  nihai ağırlıklarıyla yığın ortalamasının HA tahmin edicisi,

$$\bar{y}_w = \frac{\sum_{h=1}^H (W_h')^s y_{hR}}{\sum_{h=1}^H (W_h')^s n_{hR}} = \frac{\sum_{h=1}^H 1000000(W_h')^s y_{hR}}{\sum_{h=1}^H 1000000(W_h')^s n_{hR}} = 2,632026827$$

şeklinde belirlenmiştir. Buradan, TNSA–2003 araştırmasından Türkiye'de 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısının bir tahmini 2,63 olarak belirlenmiştir.

Literatürde başvuru kaynaklarında, uygulamalı araştırmalarda Hájek tahmin edicisi kullanımı önerilebilir. Çalışmanın bundan sonraki bölümlerinde yığın ortalaması ve oranının tahmin edicilerinin Hájek yapısındaki eşitliklerine başvurulacaktır.

## 5.2. Yanıtlanmama Ağırlıklandırması

Yanıtlanmama ağırlıklandırması yöntemlerinden ağırlıklandırma sınıf düzeltmesi yönteminin kullanımına karar verilmiştir. Yanıtlanmama için ağırlıklandırma, örneği belirli uygun ağırlıklandırma sınıflarına bölmeyi gerektirir. Bunun için yardımcı değişkenlerle ağırlıklandırma sınıfları oluşturulmasına ihtiyaç vardır.

TNSA–2003 de yanıtlanmama düzeltmesinde ağırlıklandırma sınıflarını oluşturmada kullanılabilecek iki değişken olarak bölgeler ve yerleşim yeri bilgisi vardır. Burada yerleşim yeri değişkenine göre yanıtlanmama ağırlıklandırması yapılacaktır. Yanıtlanmama ağırlıklandırması için ağırlıklandırma sınıf düzeltmesi yöntemi kullanılacaktır.

Araştırmada bölgeler açısından da ağırlıklandırma sınıf düzeltmesi yöntemi incelenmiştir. Bölgelere dayalı belirlenen tahmin sonuçlarının yerleşim yerine göre belirlenen tahmin sonuçlarının aynı olduğu görülmüştür. Dolayısıyla, ağırlıklandırma sınıflarını oluşturmada yerleşim yeri veya bölgeler değişkenlerinden herhangi biri kullanılabilir.

Çalışmada, yanıtlanmama düzeltmesi için ağırlıklandırma sınıflarını oluşturmada kullanılacak yardımcı değişkenin yerleşim yeri ( $x_1$ ) yardımcı değişkeni olmasına karar verilmiştir.  $x_1$ , yerleşim yeri değişkeni kentsel ve kırsal bölge şeklinde  $C=2$  düzeye sahiptir,  $c = 1, 2$ .

Çizelge 5.4. Örnek uygulaması: yerleşim yeri ve bölge

| Bölgelere ve kentsel/kırsal yerleşim yerlerine göre TNSA-2003 örneklemindeki hanehalkları ve uygun kadınların, hanehalkı ve kadın görüşmelerinin sonuçlarına göre yüzde dağılımı ve yanıtlama oranları, Türkiye 2003                                                                                                                                                                                                                                                                                                                                                                                    |               |             |       |       |       |       |       |        |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|-------------|-------|-------|-------|-------|-------|--------|
| Sonuç                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Yerleşim yeri |             |       |       | Bölge |       |       | Toplam |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | Kent          | Kır         | Batı  | Güney | Orta  | Kuzey | Doğu  |        |
| <b>Seçilen hanehalkları</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |               |             |       |       |       |       |       |        |
| Sorukağıdı dolduruldu (C)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 81.6          | 87.4        | 79.0  | 85.1  | 83.0  | 80.2  | 89.2  | 83.0   |
| Ziyaret sırasında evde hanehalkı yok veya görüşülebilecek nitelikte kimse yok (HP)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 2.9           | 1.2         | 3.4   | 3.1   | 2.1   | 1.8   | 1.3   | 2.5    |
| Sonraya bırakıldı (P)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 0.1           | 0.0         | 0.1   | 0.0   | 0.1   | 0.0   | 0.0   | 0.0    |
| Reddetti (R)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 4.4           | 0.4         | 6.3   | 2.4   | 2.0   | 1.8   | 1.7   | 3.4    |
| Konut bulunamadı (DNF)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 0.3           | 0.2         | 0.4   | 0.6   | 0.0   | 0.3   | 0.2   | 0.3    |
| Hanehalkının tümü araştırma tarihlerinde evde yok (HA)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 5.0           | 7.4         | 4.7   | 4.2   | 8.0   | 9.9   | 3.6   | 5.6    |
| Konutta adreste yaşayan yok / Adres konut değil (DV)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 5.4           | 3.1         | 5.5   | 4.6   | 4.6   | 6.0   | 3.6   | 4.8    |
| Konut yıkılmış (DD)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 0.0           | 0.0         | 0.0   | 0.0   | 0.0   | 0.0   | 0.0   | 0.0    |
| Kısmen dolduruldu (PC)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 0.2           | 0.1         | 0.3   | 0.1   | 0.1   | 0.1   | 0.1   | 0.1    |
| Diğer (O)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 0.1           | 0.2         | 0.3   | 0.1   | 0.1   | 0.1   | 0.2   | 0.2    |
| Toplam                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 100.0         | 100.0       | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100    |
| Örneklemdaki hanehalkı sayısı                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 9,754         | 3,295       | 4,267 | 1,797 | 2,433 | 1,587 | 2,965 | 13,049 |
| Hanehalkı cevaplama oranı (HRR) <sup>1</sup>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 91.3          | 97.9        | 88.3  | 93.3  | 95.1  | 95.4  | 96.4  | 92.9   |
| <b>Uygun kadınlar</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |               |             |       |       |       |       |       |        |
| Sorukağıdı dolduruldu (EWC)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 95.5          | 95.9        | 94.1  | 95.0  | 95.7  | 97.0  | 96.9  | 95.6   |
| Evde değil (EWNH)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 2.4           | 3.1         | 3.1   | 2.7   | 2.7   | 2.5   | 2.0   | 2.6    |
| Sonraya bırakıldı (EWP)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 0.1           | 0.0         | 0.1   | 0.0   | 0.2   | 0.0   | 0.0   | 0.1    |
| Reddetti (EWR)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 1.3           | 0.0         | 1.7   | 1.0   | 0.7   | 0.1   | 0.6   | 1.0    |
| Kısmen dolduruldu (EWPC)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 0.4           | 0.3         | 0.6   | 0.9   | 0.3   | 0.0   | 0.2   | 0.4    |
| Diğer (EWO)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 0.3           | 0.6         | 0.4   | 0.3   | 0.3   | 0.4   | 0.3   | 0.4    |
| Toplam                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 100.0         | 100.0       | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100    |
| Kadın sayısı                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 6,259         | 2,188       | 2,478 | 1,171 | 1,550 | 929   | 2,319 | 8,447  |
| Uygun kadın cevaplama oranı (EWRR) <sup>2</sup>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 95.5          | 95.9        | 94.1  | 95.0  | 95.7  | 97.0  | 96.9  | 95.6   |
| Genel cevaplama oranı (ORR) <sup>3</sup>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | <b>87.1</b>   | <b>93.9</b> | 83.0  | 88.7  | 91.0  | 92.5  | 93.4  | 88.8   |
| <b>Not:</b> Hanehalkı cevaplama oranı, görüşmesi tamamlanan hanehalkı sayısının, görüşmesi tamamlanan, ziyaret sırasında evde hanehalkının bulunmadığı, uygun cevaplayıcı olmayan, ertelenen, reddedilen ve görüşmesi tamamlanmayan hanehalkları sayılarının toplamına bölünmesiyle hesaplanmıştır. Uygun kadın cevaplama oranı ise tamamlanmış görüşme sayısının, tamamlanmış, evde yok, ertelenmiş, reddedilmiş, kısmen doldurulmuş ve “diğer” kodu almış görüşme sayılarının toplamına bölünmesiyle hesaplanmıştır. Genel cevaplama oranı hanehalkı ve uygun kadın cevaplama oranlarının çarpımıdır. |               |             |       |       |       |       |       |        |
| <sup>1</sup> Hanehalkı cevaplama oranı (HRR), çeşitli cevaplama kategorilerinde yer alan hanehalkı sayıları kullanılarak, şu şekilde hesaplanmıştır:                                                                                                                                                                                                                                                                                                                                                                                                                                                    |               |             |       |       |       |       |       |        |
| $\frac{C}{C + HP + P + HR + DNF + PC}$                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |               |             |       |       |       |       |       |        |
| <sup>2</sup> Uygun kadın cevaplama oranı (EWRR), çeşitli cevaplama kategorilerinde yer alan kadın sayıları kullanılarak, şu şekilde hesaplanmıştır:                                                                                                                                                                                                                                                                                                                                                                                                                                                     |               |             |       |       |       |       |       |        |
| $\frac{EWC}{EWC + EWNH + EWP + EWR + EWPC + EWO}$                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |               |             |       |       |       |       |       |        |
| <sup>3</sup> Genel cevaplama oranı (ORR) şu şekilde hesaplanmıştır:                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |               |             |       |       |       |       |       |        |
| $ORR = HRR * EWRR$                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |               |             |       |       |       |       |       |        |

(Kaynak: TNSA-2003 araştırması)

TNSA-2003'te kadınlar için genel cevaplama oranı (ORR) yüzde 89 olarak hesaplanmıştır. Bu oran yüzde 83 (Batı bölgesi) ile yüzde 93 (Doğu Anadolu) arasında değişmektedir. yerleşim yerleri açısından değerlendirildiğinde ise, bu oranın yüzde 87 (Kent) ile yüzde 94 (Kır) arasında değiştiği görülmektedir.

$x_1$ , yerleşim yeri yardımcı değişkenine göre ağırlıklandırma sınıflarına ayrılan veriye ağırlıklandırma sınıf düzeltmesi yöntemi uygulanacaktır. Ağırlıklandırma sınıflarındaki yanıtlayanların düzeltilmiş örnekleme ağırlıkları, bu yöntemle belirlenecek yanıtlamama ağırlıklarıyla artırılarak benzer yanıtlamayanlar temsili sağlanmaya çalışılacaktır.

### 5.2.1. Ağırlıklandırma sınıf düzeltmesi

Ağırlıklandırma sınıf düzeltmesi örneği gruplara (ağırlıklandırma sınıflarına) ayıran ve her bir sınıfa yanıt oranının tersini ağırlık olarak atayan bir yanıtlanmama düzeltmesidir. Ağırlıklandırma sınıfları oluşturulduktan sonra, her sınıftaki yanıtlayıcıların örnekleme ağırlıkları o sınıfın cevap oranı tersi ile çarpılır, böylece o sınıftaki yanıtlamayanların oranı telafi edilmiş ve yanıtlamama düzeltmesi yapılmış olur.

Örnek,  $x_1$  yerleşim yeri değişkenine göre kentsel ve kırsal olmak üzere iki ağırlıklandırma sınıfına ayrılır.  $n_c$  çaplı ağırlıklandırma sınıflarında  $n_{cR}$  sayıda yanıtlayan birimleri olmuştur,  $c = 1, 2$ . Ağırlıklandırma sınıf düzeltmesi yöntemiyle,  $c$  inci sınıf yanıtlamama ağırlığı,

$$W_c^* = \frac{n_c}{n_{cR}}$$

olur.

Çizelge 5.5. TNSA–2003 örneğindeki uygun kadınların, kentsel/kırsal yerleşim yerlerine göre yanıt oranı ve yanıtlamama ağırlıkları

| Yerleşim yeri | $n_{cR}$ | Yanıt oranı (ORR)<br>$n_{cR} / n_c$ | Yanıtlamama ağırlığı<br>$W_c^* = n_c / n_{cR}$ |
|---------------|----------|-------------------------------------|------------------------------------------------|
| Kent          | 5976     | 0,871                               | 1,148                                          |
| Kır           | 2099     | 0,939                               | 1,065                                          |
| Toplam        | 8075     | -                                   | -                                              |

Çizelge 5.6.  $\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W'_h y_{chi}$  ve  $\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W'_h$  toplam değerlerinin hesaplamaları

| Birim no | $W'_h$  | (y)<br>Canlı doğan çocuk | ( $x_1$ )<br>Yerleşim yeri | $W_c^*$ | $W_c^* W'_h$    | $W_c^* W'_h y_{chi}$ |
|----------|---------|--------------------------|----------------------------|---------|-----------------|----------------------|
| 1        | 1708,8  | 2                        | 1                          | 1,148   | 1961,701        | 3923,402             |
| 2        | 1708,8  | 1                        | 1                          | 1,148   | 1961,701        | 1961,701             |
| .        | .       | .                        |                            | .       | .               | .                    |
| .        | .       | .                        |                            | .       | .               | .                    |
| .        | .       | .                        |                            | .       | .               | .                    |
| 633      | 2602,12 | 3                        | 1                          | 1,148   | 2987,237        | 8961,711             |
| 634      | 2602,12 | 4                        | 1                          | 1,148   | 2987,237        | 11948,95             |
| 635      | 2602,12 | 7                        | 1                          | 1,148   | 2987,237        | 20910,66             |
| .        | .       | .                        |                            | .       | .               | .                    |
| .        | .       | .                        |                            | .       | .               | .                    |
| .        | .       | .                        |                            | .       | .               | .                    |
| 6932     | 2235,4  | 1                        | 1                          | 1,148   | 2566,215        | 2566,215             |
| 6933     | 2235,4  | 1                        | 1                          | 1,148   | 2566,215        | 2566,215             |
| .        | .       | .                        |                            | .       | .               | .                    |
| .        | .       | .                        |                            | .       | .               | .                    |
| .        | .       | .                        |                            | .       | .               | .                    |
| 8074     | 899,22  | 5                        | 2                          | 1,065   | 957,6731        | 4788,365             |
| 8075     | 899,22  | 8                        | 2                          | 1,065   | 957,6731        | 7661,385             |
|          |         |                          |                            | TOPLAM  | <b>14204155</b> | <b>37298172</b>      |

Ağırlıklandırma sınıf düzeltmesi yöntemiyle yığın ortalamasının Hájek yapısındaki tahmini,

$$\bar{y}_{w,a} = \frac{\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W'_h y_{chi}}{\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W'_h} \quad W'_h = \frac{N_h}{n_{hR}} \quad W_c^* = \frac{n_c}{n_{cR}}$$

ile belirlenecektir.  $\bar{y}_{w,a}$  formülünün pay ve paydası yukarıdaki gibi Çizelge 5.6'dan belirlenir.

Buradan, ağırlıklandırma sınıf düzeltmesi yöntemiyle  $\bar{Y}$  yığın ortalamasının HA yapısındaki tahmin edicisi,

$$\bar{y}_{w,a} = \frac{\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W_h' y_{chi}}{\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{chR}} W_c^* W_h'} = \frac{37298172}{14204155} = 2,617$$

olarak belirlenir.

$\bar{y}_{w,A}$ 'nın varyansının tahmini için,

$$\hat{V}_{HT}(\bar{y}_{w,A}) = \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{hR}} [(W_c^* W_h')^2 (y_{chi} - \bar{y}_{w,A})^2] = 0,000540$$

eşitliğinden elde edilmiştir. Varyans formülünde, ağırlıklar toplamının 1 olması

sağlanacak şekilde  $(W_c^* W_h')$  ağırlıkları  $\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{hR}} (W_c^* W_h')$  toplamlarına bölünmüştür.

Bu ağırlık değerleri  $\hat{V}_{HT}(\bar{y}_{w,A})$ 'de kullanılmıştır.

Ağırlıklandırmanın son aşaması olarak, veride yardımcı değişkenlere dayalı ilave ağırlıklandırma yapılmak istenilebilir. Böylece yığının kapsanmaması, örnek ile yığın arasındaki ayrılıkları gidermek ve parametre tahminlerinin duyarlılığını arttırmak amacıyla ayarlama ağırlıklandırması yapılmasına karar verilsin.

### 5.3. Ayarlama Ağırlıklandırması

TNSA–2003 araştırmasında ayarlama ağırlıklandırması yöntemlerinden,

- a) Sonradan tabakalama
- b) Tarama
- c) Ayarlama tahmin edicisi
- d) Aşamalı ayarlama tahmin edicisi

yöntemleri incelenecektir.  $\bar{Y}$ , 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısı tahminleri yapılacaktır. *(Uygulamada ayarlama ağırlıklandırmasının genel bir gösterimi olarak ağırlıklarda \*\* üst indis gösterimi kullanımı tercih edilmiştir.)*

Ayarlama ağırlıklandırma yöntemlerinde öncelikli olarak  $y$  ilgilenilen değişkeniyle ilişkili olan ve yığın toplamı bilinen yardımcı değişkenler tespit edilir. TNSA–2003 araştırmasında  $y$ , 15–49 yaşında evlenmiş kadınların canlı doğan çocuk sayısı değişkeniyle güçlü ilişkili  $x_2$ : Yaş grubu ve  $x_3$ : Eğitim durumu kategorik yapıdaki yardımcı değişkenleriyle ayarlama ağırlıklandırması yapılacağına karar verilmiştir. Burada,  $x_2$ ,  $K=7$  düzeyli;  $x_3$ ,  $L=4$  düzeylidir,  $k = 1, \dots, 7$   $l = 1, \dots, 4$ .

$\bar{Y}$  yığın ortalamasının tahmin edicisi,  $y$  ilgilenilen değişkeninin gözlem değerlerine  $W'_h$  örnekleme ağırlıklarının atanmasıyla,  $W_c^*$  yanıtlanama ağırlıklarının atanmasıyla ve  ${}_{st}W_{kl}^{**}$  sonradan tabakalama (veya  ${}_{tr}W_{kl}^{**}$  tarama veya  ${}_cW_i$  ayarlama) ağırlıklarının atanmasıyla yapılacaktır.

Ayarlama ağırlıklandırması yöntemlerinden sonradan tabakalama ve tarama yönteminde öncelikli olarak  $x_2$ : Yaş grubu ve  $x_3$ : Eğitim durumu yardımcı değişkenlerine göre örnek verisinin çapraz sınıflaması yapılmalı ve yığındaki dağılımlarının belirlenmesi gereklidir.

Çizelge 5.7’de, TNSA–2003’de kadınlar yanıtlayıcı örneğinde  $x_2$  ve  $x_3$  çapraz sınıflaması verilmektedir. 2000 Genel Nüfus Sayımı’ndan 15–49 yaşında evlenmiş kadınların  $x_2$  ve  $x_3$ ’e göre yığındaki dağılımları belirlenmiş ve Çizelge 5.8’de verilmiştir.

Çizelge 5.7. Yaş grupları ve eğitim durumu yanıt alınan örnek dağılımı ( $n_{klR}$ ’ler)

|                   |           | $x_3$ : Eğitim durumu                                     |                    |                       |                         |        |
|-------------------|-----------|-----------------------------------------------------------|--------------------|-----------------------|-------------------------|--------|
|                   |           | Okuma yazma bilmeyen- bilen fakat bir okul bitirmemiş (0) | İlkokul Mezunu (1) | İlköğretim Mezunu (2) | Lise ve üzeri Mezun (3) | Toplam |
| $x_2$ : Yaş grubu | (1) 15-19 | 78                                                        | 114                | 32                    | 16                      | 240    |
|                   | (2) 20-24 | 193                                                       | 569                | 128                   | 190                     | 1080   |
|                   | (3) 25-29 | 256                                                       | 785                | 117                   | 358                     | 1516   |
|                   | (4) 30-34 | 316                                                       | 822                | 110                   | 258                     | 1506   |
|                   | (5) 35-39 | 387                                                       | 727                | 87                    | 209                     | 1410   |
|                   | (6) 40-44 | 424                                                       | 626                | 66                    | 181                     | 1297   |
|                   | (7) 45-49 | 378                                                       | 477                | 45                    | 126                     | 1026   |
| Toplam            |           | 2032                                                      | 4120               | 585                   | 1338                    | 8075   |

Çizelge 5.8. Yaş grupları ve eğitim durumuna göre yığın dağılımı ( $N_{kl}$ ’ler)

|                   |           | $x_3$ : Eğitim durumu                                     |                    |                       |                         |          |
|-------------------|-----------|-----------------------------------------------------------|--------------------|-----------------------|-------------------------|----------|
|                   |           | Okuma yazma bilmeyen- bilen fakat bir okul bitirmemiş (0) | İlkokul Mezunu (1) | İlköğretim Mezunu (2) | Lise ve üzeri Mezun (3) | Toplam   |
| $x_2$ : Yaş grubu | (1) 15-19 | 69129                                                     | 286270             | 58176                 | 39936                   | 453511   |
|                   | (2) 20-24 | 201565                                                    | 1042410            | 141496                | 341894                  | 1727365  |
|                   | (3) 25-29 | 308860                                                    | 1403440            | 182969                | 483396                  | 2378665  |
|                   | (4) 30-34 | 368663                                                    | 1285089            | 171013                | 419626                  | 2244391  |
|                   | (5) 35-39 | 474445                                                    | 1290072            | 145111                | 373329                  | 2282957  |
|                   | (6) 40-44 | 522265                                                    | 1022539            | 100492                | 277055                  | 1922351  |
|                   | (7) 45-49 | 593857                                                    | 775233             | 73826                 | 178354                  | 1621270  |
| Toplam            |           | 2538784                                                   | 7105053            | 873083                | 2113590                 | 12630510 |

(Çizelgede gösterilen ( ) içindeki kodlamalar TNSA–2003 araştırmasının ham verisinde kullanılan tanımlamalardır.)

Aşağıda sırasıyla ayarlama ağırlıklandırması yöntemlerinden,

- sonradan tabakalama,
- tarama,
- ayarlama tahmin edicisi
- aşamalı ayarlama tahmin edicisi

yöntemleriyle 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısının tahminleri yapılacaktır.

Bu yöntemlerle elde edilen tahmin edicilerin varyanslarının tahminleri Lu ve Gelman (2003)'ün önermiş olduğu varyans formülüyle belirlenecektir. Yöntemlerle belirlenen ağırlıkların değişkenlik ölçüsü olan varyans şişme faktörleri hesaplanacak, sonuçlar kıyaslanacaktır.

### 5.3.1. Sonradan tabakalama

Sonradan tabakalama yöntemiyle yığın yardımcı değişkenlere göre çeşitli tabakalara ayrılarak her bir kategorideki yanıtların dağılımının tahmin edilmesi, yığın ile örnek arasında bilinen farklılıkların düzeltilmesi amaçlanır.

Çalışmada,  $x_2$ : Yaş grubu ve  $x_3$ : Eğitim durumu yardımcı değişkenlerine göre yanıt alınanlar örnek verisi ve yığın yukarıda verilen Çizelge 5.7 ve Çizelge 5.8'deki gibi sonradan tabakalara ayrılmıştır.

Sonradan tabaka ağırlıkları,

$${}_{st}W_{kl}^{**} = \frac{N_{kl}}{n_{klR}}$$

ile belirlenir.  $\bar{Y}$  yığın ortalamasının sonradan tabakalama tahmin edicisi,

$$\bar{y}_{w,st} = \frac{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} {}_{st}W_{kl}^{**} W_c^* W_h' y_{klchi}}{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} {}_{st}W_{kl}^{**} W_c^* W_h'}$$

$${}_{st}W_{kl}^{**} = \frac{N_{kl}}{n_{klR}}$$

$$\bar{y}_{w,st} = \frac{\sum_{k=1}^K \sum_{l=1}^L {}_{st}W_{kl}^{**} \sum_{c=1}^C W_c^* \sum_{h=1}^H W_h' y_{klch}}{\sum_{k=1}^K \sum_{l=1}^L {}_{st}W_{kl}^{**} \sum_{c=1}^C W_c^* \sum_{h=1}^H W_h' n_{klch}}$$

için ağırlıklandırılmış formülünün kullanımı önerilebilir.  ${}_{st}W_{kl}^{**}$ ,  $k$  ıncı satır  $l$  inci sütun sonradan tabaka ağırlığıdır,  $k=1,...,7$ ,  $l=1,...,4$ .  $W_h'$ ,  $h$  ıncı tabaka düzeltilmiş örnekleme ağırlığı ve  $W_c^*$ ,  $c$  inci ağırlıklandırma sınıfı yanıtlamama ağırlığıdır,  $h=1,...,40$ ,  $c=1, 2$ .

${}_{st}W_{kl}^{**}$  sonradan tabaka ağırlıkları Çizelge 5.7 ve Çizelge 5.8'de verilen çapraz sınıflamalara göre aşağıdaki gibi belirlenir.

Çizelge 5.9. Sonradan tabaka ağırlıkları ( ${}_{st}W_{kl}^{**} = N_{kl} / n_{klR}$ )

|                   |       | $X_3$ : Eğitim durumu                                |                             |                             |                             |
|-------------------|-------|------------------------------------------------------|-----------------------------|-----------------------------|-----------------------------|
|                   |       | Okuma yazma bilmeyen-bilen fakat bir okul bitirmemiş | İlkokul mezunu              | İlköğretim mezunu           | Lise ve üzeri mezun         |
| $x_2$ : Yaş grubu | 15-19 | ${}_{st}W_{11}^{**}=886,27$                          | ${}_{st}W_{12}^{**}=2511,1$ | ${}_{st}W_{13}^{**}=1818$   | ${}_{st}W_{14}^{**}=2496$   |
|                   | 20-24 | ${}_{st}W_{21}^{**}=1044,4$                          | ${}_{st}W_{22}^{**}=1832$   | ${}_{st}W_{23}^{**}=1105,4$ | ${}_{st}W_{24}^{**}=1799,4$ |
|                   | 25-29 | ${}_{st}W_{31}^{**}=1206,5$                          | ${}_{st}W_{32}^{**}=1787,8$ | ${}_{st}W_{33}^{**}=1563,8$ | ${}_{st}W_{34}^{**}=1350,3$ |
|                   | 30-34 | ${}_{st}W_{41}^{**}=1166,7$                          | ${}_{st}W_{42}^{**}=1563,4$ | ${}_{st}W_{43}^{**}=1554,7$ | ${}_{st}W_{44}^{**}=1626,5$ |
|                   | 35-39 | ${}_{st}W_{51}^{**}=1226$                            | ${}_{st}W_{52}^{**}=1774,5$ | ${}_{st}W_{53}^{**}=1667,9$ | ${}_{st}W_{54}^{**}=1786,3$ |
|                   | 40-44 | ${}_{st}W_{61}^{**}=1231,8$                          | ${}_{st}W_{62}^{**}=1633,4$ | ${}_{st}W_{63}^{**}=1522,6$ | ${}_{st}W_{64}^{**}=1530,7$ |
|                   | 45-49 | ${}_{st}W_{71}^{**}=1571,1$                          | ${}_{st}W_{72}^{**}=1625,2$ | ${}_{st}W_{73}^{**}=1640,6$ | ${}_{st}W_{74}^{**}=1415,5$ |

$VIF$ , varyans şişme faktörü ağırlıklardaki değişkenliğin bir ölçüsüdür.  $VIF$ 'in değerinin küçük olması istenir. Sonradan tabakalama ağırlıkları için varyans şişme faktörü,

$$VIF = 1 + CV^2({}_{st}W_{kl}^{**}) = \left[ n_R \left( \sum_{k=1}^K \sum_{l=1}^L \sum_{i=1}^{n_{klR}} {}_{st}W_{kl}^{**2} \right) / \left( \sum_{k=1}^K \sum_{l=1}^L \sum_{i=1}^{n_{klR}} {}_{st}W_{kl}^{**} \right)^2 \right]$$

$$= \left[ n_R \left( \sum_{k=1}^K \sum_{l=1}^L n_{klR} {}_{st}W_{kl}^{**2} \right) / \left( \sum_{k=1}^K \sum_{l=1}^L n_{klR} {}_{st}W_{kl}^{**} \right)^2 \right]$$

$$VIF({}_{st}W_{kl}^{**}) = 1,067742728$$

şeklinde tüm yanıt alınan örnek birimleri üzerinden belirlenir.

$\bar{y}_{w,st}$ 'nin hesaplamalarında öncelikli olarak gözlem değerlerine  $W'_h$  düzeltilmiş örnekleme ağırlıkları atanır. Daha sonra  $W'_h$  ağırlıkları atanmış veride,  $W_c^*$  yanıtlamama ağırlıklarının atanması yapılır, bunun için veri  $x_1$ : Yerleşim yeri değişkenine göre kent ve kır olarak ayrılır ve  $W_c^*$  yanıtlamama ağırlıkları da veriye atanır.

Son olarak  $W'_h$  ve  $W_c^*$ 'lere göre ağırlıklandırılmış veri,  $x_2$ : Yaş grubu ve  $x_3$ : Eğitim durumu yardımcı değişkenlerine göre sonradan tabakalanır ve gözlem değerlerine  ${}_{st}W_{kl}^{**}$  sonradan tabaka ağırlıkları atanarak istenen tahminler elde edilmeye çalışılır. .

Sonradan tabakalama için ağırlıklandırma aşaması ve  $\bar{y}_{w,st}$  için hesaplamaların özeti aşağıda Çizelge 5.10 ile verilmiştir.

Çizelge 5.10. Sonradan tabaka ağırlıklandırması hesaplamaları

| Birim no | $W'_h$  | $y$ | $x_1$ | $W_c^*$ | $x_2$ | $x_3$ | ${}_{st}W_{kl}^{**}$ | ${}_{st}W_{kl}^{**} W_c^* W'_h$ | ${}_{st}W_{kl}^{**} W_c^* W'_h y_{klchi}$ |
|----------|---------|-----|-------|---------|-------|-------|----------------------|---------------------------------|-------------------------------------------|
| 1        | 1708,8  | 2   | 1     | 1,148   | 2     | 1     | 1832                 | 3593836,441                     | 7187672,881                               |
| 2        | 1708,8  | 1   | 1     | 1,148   | 2     | 1     | 1832                 | 3593836,441                     | 3593836,441                               |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| 633      | 2602,12 | 3   | 1     | 1,148   | 4     | 1     | 1563,4               | 4670246,077                     | 14010738,23                               |
| 634      | 2602,12 | 4   | 1     | 1,148   | 7     | 2     | 1640,6               | 4900860,761                     | 19603443,04                               |
| 635      | 2602,12 | 7   | 1     | 1,148   | 6     | 1     | 231,8                | 692441,4997                     | 4847090,498                               |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| 6932     | 2235,4  | 1   | 1     | 1,148   | 7     | 1     | 1571,1               | 6209143,493                     | 6209143,493                               |
| 6933     | 2235,4  | 1   | 1     | 1,148   | 2     | 2     | 1832                 | 7240246,247                     | 7240246,247                               |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| 8074     | 899,22  | 5   | 2     | 1,065   | 3     | 1     | 1206,5               | 1155432,566                     | 5777162,828                               |
| 8075     | 899,22  | 8   | 2     | 1,065   | 6     | 2     | 1633,4               | 1564263,201                     | 12514105,61                               |
|          |         |     |       |         |       |       | <b>Toplam</b>        | <b>21758233235</b>              | <b>57138232563</b>                        |

$\bar{Y}$  'nın sonradan tabakalama tahmin edicisi,

$$\bar{y}_{w,st} = \frac{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} {}_{st}W_{kl}^{**} W_c^* W'_h y_{klchi}}{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} {}_{st}W_{kl}^{**} W_c^* W'_h} = \frac{57138232563}{21758233235} = 2,626$$

belirlenir. Eğitim durumu ve yaş gruplarına göre sonradan tabakalanmış TNSA–2003 verisinden, Türkiye’de 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısının 2,6 olduğu belirlenir.

$\bar{y}_{w,st}$  'nin varyansının tahmini,

$$\hat{V}_{HT}(\bar{y}_{w,st}) = \sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{hR}} [(_{st}W_{kl}^{**}W_c^*W_h')^2 (y_{klchi} - \bar{y}_{w,st})^2] = 0,000604$$

eşitliğinden elde edilmiştir. Bu varyans formülünde,  $_{st}W_{kl}^{**}W_c^*W_h'$  ağırlıklarının tüm yanıtlayan birimleri için  $\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{hR}} _{st}W_{kl}^{**}W_c^*W_h'$  toplamına bölünmüş değerleri kullanılmıştır. Böylece varyans tahmin formülü sonucu hesaplaması yapılmıştır.

Sonradan tabakalama yapılan veriye hem örnekleme, hem yanıtlamama ve hem de sonradan tabaka ağırlıklarının atanmasıyla istenen yığın ortalaması tahmini belirlenmiştir. Bu ağırlıkların birlikte değişimi için  $VIF$ , varyans şişme faktörü değeri belirlenebilir:

$$VIF = \left[ n_R \left( \sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} (_{st}W_{kl}^{**}W_c^*W_h')^2 \right) / \left( \sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} _{st}W_{kl}^{**}W_c^*W_h' \right)^2 \right]$$

$$VIF (_{st}W_{kl}^{**}W_c^*W_h') = 1,438373205.$$

### 5.3.2. Tarama

Tarama, örnek araştırmalarında iki veya daha fazla sayıda yardımcı değişkene dayalı oluşturulan çapraz sınıflar olması durumunda yaygın kullanılan bir ağırlıklandırma yöntemidir. Tarama yöntemi, mevcut yardımcı bilgiyi kullanarak, örnekleme ağırlıklarını düzenlemeyi içerir ve yardımcı değişkenlerin bilinen marjinal yığın toplamalarını kullanır.

Yöntemin uygulamasında öncelikli olarak belirlenen yardımcı değişkenlere göre örnek dağılımının çapraz sınıflaması yapılır. Çalışmada,  $x_2$ : Yaş grubu ve  $x_3$ : Eğitim

durumu yardımcı değişkenlerine göre örneğin dağılımının yığının dağılımına uygunluğunun sağlanmasına karar verilmiştir.

$x_2$  ve  $x_3$  değişkenlerinin yığın dağılımı ve yanıtlayıcıların örnek dağılımı yukarıda Çizelge 5.7 ve Çizelge 5.8’de verilmişti. Tarama yöntemiyle,  $n_{klR}$  yanıtlayıcı örnek çaplarına ve  $N_{k.}$  satır ve  $N_{.l}$  sütun marjinal yığın çaplarına dayalı olarak  $N_{kl}$  yığın göze toplamı için tahminler yapılmaya çalışılacaktır,  $k = 1, \dots, 7$   $l = 1, \dots, 4$ .

Tarama ağırlıklarının hesaplanmasında ilk olarak örnek satır toplamı yığın satır toplamına uymaya zorlanır. Sonra, örnek düzeltilmiş sütun toplamı yığın sütun toplamına uymaya zorlanır. Daha sonra, örnek satır toplamı yığın satır toplamına uydurulabilmek için yeniden düzeltilir. Yığın satır ve sütun toplamına yakınsama sağlanana kadar satır-sütun düzeltme işlemlerine devam edilir. Uygun yakınsamaya çoğunlukla çabuk ulaşılır.

Tarama yöntemiyle yığın satır ve sütun toplamına yakınsama sağlandığında, belirlenen nihai son  $k$  ıncı satır  $l$  inci sütun göze toplamı tahmini değerleri  $\tilde{N}_{kl}$  ile gösterilir.

Her bir adım bir önceki adımın sonuçlarıyla başlar, tüm eşitlikler eş zamanlı olarak istenilene yakın olduğunda yani,  $\sum_{l=1}^L \tilde{N}_{kl} = N_{k.}$  ve  $\sum_{k=1}^K \tilde{N}_{kl} = N_{.l}$  sağlandığında işlem sonlandırılır.  $\tilde{N}_{kl}$ ’ler  $k$  ıncı satır  $l$  inci sütun için son adımda elde edilen göze toplamı değeridir.

Tarama ağırlıkları,

$${}_{tr}W_{kl}^{**} = \frac{\tilde{N}_{kl}}{n_{klR}}$$

ile belirlenir. Tarama yöntemiyle  $\bar{Y}$  'nın ağırlıklandırılmış (HA) tahmin edicisi,

$$\bar{y}_{w,tr} = \frac{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} tr W_{kl}^{**} W_c^* W_h' y_{klchi}}{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} tr W_{kl}^{**} W_c^* W_h'}$$

$$tr W_{kl}^{**} = \frac{\tilde{N}_{kl}}{n_{klR}}$$

$$\bar{y}_{w,tr} = \frac{\sum_{k=1}^K \sum_{l=1}^L tr W_{kl}^{**} \sum_{c=1}^C W_c^* \sum_{h=1}^H W_h' y_{klch}}{\sum_{k=1}^K \sum_{l=1}^L tr W_{kl}^{**} \sum_{c=1}^C W_c^* \sum_{h=1}^H W_h' n_{klch}}$$

kullanımı önerilebilir.  $tr W_{kl}^{**}$ ,  $k$  ıncı satır  $l$  inci sütun tarama ağırlığıdır,  $k = 1, \dots, 7$ ,  $l = 1, \dots, 4$ .  $W_h'$ ,  $h$  ıncı tabaka düzeltilmiş örnekleme ağırlığı ve  $W_c^*$ ,  $c$  inci ağırlıklandırma sınıfı yanıtlamama ağırlığıdır,  $h = 1, \dots, 40$ ,  $c = 1, 2$ .

Çizelge 5.7 ve Çizelge 5.8'deki  $x_2$  ve  $x_3$  değişkenleri çapraz sınıflama verisinin tarama yinelemeleri Çizelge 5.11'de aşağıdaki gibi verilecektir.

Düzeltilmenin 1. Turu satır düzeltmesiyle başlar. 1. Tur satır düzeltmesinde, her bir satırda  $n_{klR}$  'ler  $N_{k.}$  yığın satır toplamlarına uydurulur.  $x_2$  'nin birinci satırındaki  $n_{11R}$  'ler 453511/240, ikinci satırındaki  $n_{12R}$  'ler 1727365/1080 ile çarpılarak ve böyle devam ettirilerek satır düzeltmeleri yapılır.

Mesela,  $x_2$  'nin  $k=1$  ve  $x_3$  'ün  $l=1$  gözesi için 1. Tur satır düzeltmesi değeri  $78(453511/240)=147391,1$  olarak belirlenir.

2. Tur sütun düzeltmesinde,  $N_{.l}$  yığın sütun toplamlarına uyum sağlamak için 1. Tur satır düzeltmesi sonuçları düzeltilir. 1. Tur satır düzeltmesinden,  $x_3$  'ün birinci sütunundaki değerler 2538784/3181028, ikinci sütunundaki değerler

7105053/6440880 ile çarpılır, üçüncü ve dördüncü sütunlar da benzer şekilde belirlenir.

Mesela,  $x_2$ 'nin  $k=1$  ve  $x_3$ 'ün  $l=1$  gözesi için 2. Tur sütun düzeltmesi değeri  $147391,1(2538784/3181028)=117633,1$  olarak belirlenir.

Yineleme işlemi marjinal satır ve sütun toplamaları yığının marjinal satır ve sütun toplamalarına uyana kadar tekrarlanır. Uyumun 9. Turda sağlandığı Çizelge 5.11'den görülmektedir. 9. Turda marjinal toplamalara bakıldığında  $N_{k.}$  satır ve  $N_{.l}$  sütun marjinal yığın çaplarına eşit olduğu görülmektedir.

Çizelge 5.11. Çizelge 5.7 ve Çizelge 5.8'deki veri için tarama yinelemesi

| 1. TUR: Satır Düzeltmesi |   |          |          |          |          |          |
|--------------------------|---|----------|----------|----------|----------|----------|
|                          |   | $x_3$    |          |          |          |          |
|                          |   | 1        | 2        | 3        | 4        | TOPLAM   |
| $x_2$                    | 1 | 147391,1 | 215417,7 | 60468,13 | 30234,07 | 453511   |
|                          | 2 | 308686,5 | 910065,4 | 204724,7 | 303888,3 | 1727365  |
|                          | 3 | 401674,3 | 1231697  | 183577,7 | 561716,4 | 2378665  |
|                          | 4 | 470934,6 | 1225026  | 163932,9 | 384497,3 | 2244391  |
|                          | 5 | 626598,8 | 1177099  | 140863,3 | 338395,8 | 2282957  |
|                          | 6 | 628432,4 | 927827,1 | 97822,02 | 268269,5 | 1922351  |
|                          | 7 | 597310   | 753748,3 | 71108,33 | 199103,3 | 1621270  |
| TOPLAM                   |   | 3181028  | 6440880  | 922497,2 | 2086105  | 12630510 |

| 2. TUR: Sütun Düzeltmesi |   |          |          |          |          |          |
|--------------------------|---|----------|----------|----------|----------|----------|
|                          |   | $x_3$    |          |          |          |          |
|                          |   | 1        | 2        | 3        | 4        | TOPLAM   |
| $x_2$                    | 1 | 117633,1 | 237631,2 | 57229,12 | 30632,41 | 443125,8 |
|                          | 2 | 246363,3 | 1003910  | 193758,5 | 307892,2 | 1751924  |
|                          | 3 | 320577   | 1358707  | 173744,2 | 569117,3 | 2422146  |
|                          | 4 | 375853,8 | 1351349  | 155151,8 | 389563,2 | 2271917  |
|                          | 5 | 500089,7 | 1298480  | 133317,9 | 342854,3 | 2274741  |
|                          | 6 | 501553   | 1023503  | 92582,12 | 271804,1 | 1889442  |
|                          | 7 | 476714,2 | 831473,6 | 67299,37 | 201726,6 | 1577214  |
| TOPLAM                   |   | 2538784  | 7105053  | 873083   | 2113590  | 12630510 |

Çizelge 5.11 (Devam). Çizelge 5.7 ve Çizelge 5.8'deki veri için tarama yinelemesi

| 3. TUR: Satır Düzeltmesi |   |          |          |          |          |          |
|--------------------------|---|----------|----------|----------|----------|----------|
|                          |   | $x_3$    |          |          |          |          |
|                          |   | 1        | 2        | 3        | 4        | TOPLAM   |
| $x_2$                    | 1 | 120389,9 | 243200,4 | 58570,35 | 31350,32 | 453511   |
|                          | 2 | 242909,7 | 989836,8 | 191042,4 | 303576,1 | 1727365  |
|                          | 3 | 314822,2 | 1334317  | 170625,3 | 558900,9 | 2378665  |
|                          | 4 | 371299,9 | 1334976  | 153271,9 | 384843,3 | 2244391  |
|                          | 5 | 501895,8 | 1303169  | 133799,4 | 344092,6 | 2282957  |
|                          | 6 | 510288,7 | 1041330  | 94194,64 | 276538,1 | 1922351  |
|                          | 7 | 490030,2 | 854699,1 | 69179,24 | 207361,4 | 1621270  |
| TOPLAM                   |   | 2551637  | 7101528  | 870683,2 | 2106663  | 12630510 |

| 4. TUR: Sütun Düzeltmesi |   |          |          |          |          |          |
|--------------------------|---|----------|----------|----------|----------|----------|
|                          |   | $x_3$    |          |          |          |          |
|                          |   | 1        | 2        | 3        | 4        | TOPLAM   |
| $x_2$                    | 1 | 119783,5 | 243321,1 | 58731,78 | 31453,41 | 453289,9 |
|                          | 2 | 241686,2 | 990328,2 | 191568,9 | 304574,3 | 1728158  |
|                          | 3 | 313236,5 | 1334979  | 171095,6 | 560738,7 | 2380050  |
|                          | 4 | 369429,7 | 1335639  | 153694,4 | 386108,7 | 2244871  |
|                          | 5 | 499367,8 | 1303816  | 134168,1 | 345224   | 2282576  |
|                          | 6 | 507718,4 | 1041847  | 94454,26 | 277447,5 | 1921467  |
|                          | 7 | 487562   | 855123,4 | 69369,91 | 208043,3 | 1620099  |
| TOPLAM                   |   | 2538784  | 7105053  | 873083   | 2113590  | 12630510 |

| 5. TUR: Satır Düzeltmesi |   |          |          |          |          |          |
|--------------------------|---|----------|----------|----------|----------|----------|
|                          |   | $x_3$    |          |          |          |          |
|                          |   | 1        | 2        | 3        | 4        | TOPLAM   |
| $x_2$                    | 1 | 119842   | 243439,8 | 58760,43 | 31468,75 | 453511   |
|                          | 2 | 241575,3 | 989874   | 191481,1 | 304434,6 | 1727365  |
|                          | 3 | 313054,2 | 1334202  | 170996   | 560412,5 | 2378665  |
|                          | 4 | 369350,7 | 1335353  | 153661,5 | 386026,1 | 2244391  |
|                          | 5 | 499451,1 | 1304034  | 134190,5 | 345281,6 | 2282957  |
|                          | 6 | 507952   | 1042326  | 94497,73 | 277575,2 | 1922351  |
|                          | 7 | 487914,5 | 855741,7 | 69420,07 | 208193,7 | 1621270  |
| TOPLAM                   |   | 2539140  | 7104970  | 873007,4 | 2113392  | 12630510 |

Çizelge 5.11 (Devam). Çizelge 5.7 ve Çizelge 5.8'deki veri için tarama yinelemesi

| 6. TUR: Sütun Düzeltmesi |          |          |          |          |          |
|--------------------------|----------|----------|----------|----------|----------|
|                          | $x_3$    |          |          |          | TOPLAM   |
|                          | 1        | 2        | 3        | 4        |          |
| 1                        | 119825,2 | 243442,7 | 58765,52 | 31471,69 | 453505,1 |
| 2                        | 241541,5 | 989885,5 | 191497,6 | 304463,1 | 1727388  |
| 3                        | 313010,4 | 1334218  | 171010,9 | 560464,9 | 2378704  |
| $x_2$ 4                  | 369298,9 | 1335368  | 153674,8 | 386062,2 | 2244404  |
| 5                        | 499381,1 | 1304049  | 134202,2 | 345313,9 | 2282946  |
| 6                        | 507880,9 | 1042338  | 94505,92 | 277601,1 | 1922326  |
| 7                        | 487846,1 | 855751,7 | 69426,08 | 208213,2 | 1621237  |
| TOPLAM                   | 2538784  | 7105053  | 873083   | 2113590  | 12630510 |

| 7. TUR: Satır Düzeltmesi |          |          |          |          |          |
|--------------------------|----------|----------|----------|----------|----------|
|                          | $x_3$    |          |          |          | TOPLAM   |
|                          | 1        | 2        | 3        | 4        |          |
| 1                        | 119826,8 | 243445,8 | 58766,29 | 31472,11 | 453511   |
| 2                        | 241538,3 | 989872,5 | 191495,1 | 304459,1 | 1727365  |
| 3                        | 313005,2 | 1334196  | 171008,1 | 560455,7 | 2378665  |
| $x_2$ 4                  | 369296,7 | 1335360  | 153673,9 | 386059,9 | 2244391  |
| 5                        | 499383,5 | 1304055  | 134202,8 | 345315,6 | 2282957  |
| 6                        | 507887,4 | 1042352  | 94507,15 | 277604,7 | 1922351  |
| 7                        | 487856   | 855769,1 | 69427,49 | 208217,4 | 1621270  |
| TOPLAM                   | 2538794  | 7105051  | 873080,8 | 2113584  | 12630510 |

| 8. TUR: Sütun Düzeltmesi |          |          |          |          |          |
|--------------------------|----------|----------|----------|----------|----------|
|                          | $x_3$    |          |          |          | TOPLAM   |
|                          | 1        | 2        | 3        | 4        |          |
| 1                        | 119826,3 | 243445,9 | 58766,43 | 31472,19 | 453510,8 |
| 2                        | 241537,3 | 989872,8 | 191495,6 | 304459,9 | 1727366  |
| 3                        | 313004   | 1334196  | 171008,5 | 560457,2 | 2378666  |
| $x_2$ 4                  | 369295,3 | 1335361  | 153674,3 | 386060,9 | 2244391  |
| 5                        | 499381,5 | 1304056  | 134203,1 | 345316,5 | 2282957  |
| 6                        | 507885,4 | 1042352  | 94507,38 | 277605,4 | 1922350  |
| 7                        | 487854,1 | 855769,3 | 69427,66 | 208218   | 1621269  |
| TOPLAM                   | 2538784  | 7105053  | 873083   | 2113590  | 12630510 |

Çizelge 5.11 (Devam). Çizelge 5.7 ve Çizelge 5.8'deki veri için tarama yinelemesi

| 9. TUR: Satır Düzeltmesi |   |          |          |          |          |          |
|--------------------------|---|----------|----------|----------|----------|----------|
|                          |   | $x_3$    |          |          |          |          |
|                          |   | 1        | 2        | 3        | 4        | TOPLAM   |
| $x_2$                    | 1 | 119826,3 | 243446   | 58766,46 | 31472,2  | 453511   |
|                          | 2 | 241537,2 | 989872,5 | 191495,5 | 304459,7 | 1727365  |
|                          | 3 | 313003,9 | 1334196  | 171008,4 | 560456,9 | 2378665  |
|                          | 4 | 369295,2 | 1335361  | 153674,3 | 386060,9 | 2244391  |
|                          | 5 | 499381,6 | 1304056  | 134203,1 | 345316,5 | 2282957  |
|                          | 6 | 507885,6 | 1042352  | 94507,41 | 277605,5 | 1922351  |
|                          | 7 | 487854,4 | 855769,8 | 69427,7  | 208218,1 | 1621270  |
| TOPLAM                   |   | 2538784  | 7105053  | 873082,9 | 2113590  | 12630510 |

9. Tura bakıldığında marjinal satır ve sütun toplamlarının,  $N_{k.}$  yığın marjinal satır ve  $N_{.l}$  yığın marjinal sütun toplamlarına yakınsamanın sağlanmış olduğu görülmektedir. (9. Tur marjinal toplamları Çizelge 5.8'le kıyaslanabilir.) Dolayısıyla yineleme işlemine son verilir.

Tarama yöntemiyle belirlenen nihai son  $k$  ıncı satır  $l$  inci sütun göze toplamı tahmini değerleri  $\tilde{N}_{kl}$  ile gösterilebilir.

Çizelge 5.12.  $\tilde{N}_{kl}$ ,  $k$  ıncı satır  $l$  inci sütun için son adımdaki göze toplamı değeri

|                   |   | $x_3$ : Eğitim durumu       |                             |                             |                             |
|-------------------|---|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
|                   |   | 1                           | 2                           | 3                           | 4                           |
| $x_2$ : Yaş grubu | 1 | $\tilde{N}_{11} = 119826,3$ | $\tilde{N}_{12} = 243446$   | $\tilde{N}_{13} = 58766,46$ | $\tilde{N}_{14} = 31472,2$  |
|                   | 2 | $\tilde{N}_{21} = 241537,2$ | $\tilde{N}_{22} = 989872,5$ | $\tilde{N}_{23} = 191495,5$ | $\tilde{N}_{24} = 304459,7$ |
|                   | 3 | $\tilde{N}_{31} = 313003,9$ | $\tilde{N}_{32} = 1334196$  | $\tilde{N}_{33} = 171008,4$ | $\tilde{N}_{34} = 560456,9$ |
|                   | 4 | $\tilde{N}_{41} = 369295,2$ | $\tilde{N}_{42} = 1335361$  | $\tilde{N}_{43} = 153674,3$ | $\tilde{N}_{44} = 386060,9$ |
|                   | 5 | $\tilde{N}_{51} = 499381,6$ | $\tilde{N}_{52} = 1304056$  | $\tilde{N}_{53} = 134203,1$ | $\tilde{N}_{54} = 345316,5$ |
|                   | 6 | $\tilde{N}_{61} = 507885,6$ | $\tilde{N}_{62} = 1042352$  | $\tilde{N}_{63} = 94507,41$ | $\tilde{N}_{64} = 277605,5$ |
|                   | 7 | $\tilde{N}_{71} = 487854,4$ | $\tilde{N}_{72} = 855769,8$ | $\tilde{N}_{73} = 69427,7$  | $\tilde{N}_{74} = 208218,1$ |

Tarama yöntemiyle ağırlıklar,  ${}_{tr}W_{kl}^{**} = \tilde{N}_{kl} / n_{klR}$  formülüyle belirlenir.  ${}_{tr}W_{kl}^{**}$ ,  $k$  ıncı satır  $s$  inci sütun tarama ağırlıkları, Çizelge 5.7’de verilen  $n_{klR}$ ’ler ve Çizelge 5.12’teki  $\tilde{N}_{kl}$ ’lerden aşağıdaki gibi belirlenir,  $k = 1, \dots, 7$   $l = 1, \dots, 4$ .

Çizelge 5.13.  ${}_{tr}W_{kl}^{**} = \tilde{N}_{kl} / n_{klR}$  tarama ağırlıkları

|                      |       | $x_3$ : Eğitim durumu                                          |                                 |                                 |                                 |
|----------------------|-------|----------------------------------------------------------------|---------------------------------|---------------------------------|---------------------------------|
|                      |       | Okuma yazma<br>bilmeyen- bilen<br>fakat bir okul<br>bitirmemiş | İlkokul<br>mezunu               | İlköğretim<br>mezunu            | Lise ve üzeri<br>mezun          |
| $x_2$ : Yaş<br>grubu | 15-19 | ${}_{tr}W_{11}^{**} = 1536,235$                                | ${}_{tr}W_{12}^{**} = 2135,491$ | ${}_{tr}W_{13}^{**} = 1836,452$ | ${}_{tr}W_{14}^{**} = 1967,013$ |
|                      | 20-24 | ${}_{tr}W_{21}^{**} = 1251,488$                                | ${}_{tr}W_{22}^{**} = 1739,67$  | ${}_{tr}W_{23}^{**} = 1496,059$ | ${}_{tr}W_{24}^{**} = 1602,419$ |
|                      | 25-29 | ${}_{tr}W_{31}^{**} = 1222,671$                                | ${}_{tr}W_{32}^{**} = 1699,613$ | ${}_{tr}W_{33}^{**} = 1461,61$  | ${}_{tr}W_{34}^{**} = 1565,522$ |
|                      | 30-34 | ${}_{tr}W_{41}^{**} = 1168,656$                                | ${}_{tr}W_{42}^{**} = 1624,527$ | ${}_{tr}W_{43}^{**} = 1397,039$ | ${}_{tr}W_{44}^{**} = 1496,36$  |
|                      | 35-39 | ${}_{tr}W_{51}^{**} = 1290,392$                                | ${}_{tr}W_{52}^{**} = 1793,75$  | ${}_{tr}W_{53}^{**} = 1542,564$ | ${}_{tr}W_{54}^{**} = 1652,232$ |
|                      | 40-44 | ${}_{tr}W_{61}^{**} = 1197,843$                                | ${}_{tr}W_{62}^{**} = 1665,099$ | ${}_{tr}W_{63}^{**} = 1431,93$  | ${}_{tr}W_{64}^{**} = 1533,732$ |
|                      | 45-49 | ${}_{tr}W_{71}^{**} = 1290,62$                                 | ${}_{tr}W_{72}^{**} = 1794,067$ | ${}_{tr}W_{73}^{**} = 1542,838$ | ${}_{tr}W_{74}^{**} = 1652,525$ |

Tarama yöntemiyle belirlenen ağırlıkların değişkenliği için varyans şişme faktörü,

$$VIF = 1 + CV^2({}_{tr}W_{kl}^{**}) = \left[ n_R \left( \sum_{k=1}^K \sum_{l=1}^L \sum_{i=1}^{n_{klR}} {}_{tr}W_{kl}^{**2} \right) / \left( \sum_{k=1}^K \sum_{l=1}^L \sum_{i=1}^{n_{klR}} {}_{tr}W_{kl}^{**} \right)^2 \right]$$

$$= \left[ n_R \left( \sum_{k=1}^K \sum_{l=1}^L n_{klR} {}_{tr}W_{kl}^{**2} \right) / \left( \sum_{k=1}^K \sum_{l=1}^L n_{klR} {}_{tr}W_{kl}^{**} \right)^2 \right]$$

$$VIF({}_{tr}W_{kl}^{**}) = 1,01871687.$$

$\bar{y}_{w,tr}$  tahmin edicisinin hesaplanmasında sonradan tabakalamada olduğu gibi hesaplama yolu izlenir. Öncelikle gözlem değerlerine  $W'_h$  düzeltilmiş örnekleme ağırlıkları atanır. Daha sonra  $W'_h$  ağırlıkları atanmış veri  $x_1$ : Yerleşim yeri değişkenine göre kent ve kır olarak ikiye ayrılır ve  $W'_c$  yanıtlamama ağırlıkları da veriye atanır.

Son olarak  $W'_h$  düzeltilmiş örnekleme ağırlıklarına ve  $W_c^*$  yanıtlamama ağırlıklarına göre ağırlıklandırılmış veri,  $x_2$ : Yaş grubu ve  $x_3$ : Eğitim durumu yardımcı değişkenlerine göre sınıflanır. Böylece sınıflanan veriye  ${}_{tr}W_{kl}^{**}$  tarama ağırlıkları atanarak istenen yığın ortalaması tahmini yapılmaya çalışılır.

$\bar{y}_{w,tr}$  için hesaplamalarının özeti aşağıda Çizelge 5.14 ile verilmektedir. Yaş grubu ve eğitim durumuna göre çapraz sınıflama yapılmasıyla tarama yöntemi uygulanmış veriden, Türkiye’de 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısının tahmini,

$$\bar{y}_{w,tr} = \frac{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} {}_{tr}W_{kl}^{**} W_c^* W'_h y_{klchi}}{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} {}_{tr}W_{kl}^{**} W_c^* W'_h} = \frac{58890128564}{22440789028} = 2,624245$$

şeklinde elde edilir.

Çizelge 5.14. Tarama ağırlıklandırması hesaplamaları

| Birim no | $W'_h$  | $y$ | $x_1$ | $W_c^*$ | $x_2$ | $x_3$ | ${}_{tr}W_{kl}^{**}$ | ${}_{tr}W_{kl}^{**} W_c^* W'_h$ | ${}_{tr}W_{kl}^{**} W_c^* W'_h y_{klchi}$ |
|----------|---------|-----|-------|---------|-------|-------|----------------------|---------------------------------|-------------------------------------------|
| 1        | 1708,8  | 2   | 1     | 1,148   | 2     | 1     | 1251,49              | 2455045,404                     | 4910090,807                               |
| 2        | 1708,8  | 1   | 1     | 1,148   | 2     | 1     | 1251,49              | 2455045,404                     | 4910090,807                               |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| 633      | 2602,12 | 3   | 1     | 1,148   | 4     | 1     | 1168,66              | 3491052,258                     | 10473156,77                               |
| 634      | 2602,12 | 4   | 1     | 1,148   | 7     | 2     | 1794,07              | 5359303,037                     | 26796515,19                               |
| 635      | 2602,12 | 7   | 1     | 1,148   | 6     | 1     | 1197,84              | 3578240,739                     | 25047685,17                               |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| 6932     | 2235,4  | 1   | 1     | 1,148   | 7     | 1     | 1290,62              | 5100658,631                     | 5100658,631                               |
| 6933     | 2235,4  | 1   | 1     | 1,148   | 2     | 2     | 1739,67              | 6875348,902                     | 6875348,902                               |
| .        | .       | .   | .     | .       | .     | .     | .                    | .                               | .                                         |
| 8074     | 899,22  | 5   | 2     | 1,065   | 3     | 1     | 1222,67              | 1170919,097                     | 5854595,484                               |
| 8075     | 899,22  | 8   | 2     | 1,065   | 6     | 2     | 1665,1               | 1594620,48                      | 12756963,84                               |
|          |         |     |       |         |       |       | <b>Toplam</b>        | <b>22440789028</b>              | <b>58890128564</b>                        |

Yaş grubu ve eğitim durumuna göre çapraz sınıflama yapılmasıyla tarama yöntemi uygulanmış veriden, Türkiye’de 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısının tahmini,

$$\bar{y}_{w,tr} = \frac{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} tr W_{kl}^{**} W_c^* W_h' y_{klchi}}{\sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} tr W_{kl}^{**} W_c^* W_h'} = \frac{58890128564}{22440789028} = 2,624245$$

şeklinde elde edilir.

$\bar{y}_{w,tr}$  'nin varyansının tahmini için,

$$\hat{V}_{HT}(\bar{y}_{w,tr}) = \sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} [(tr W_{kl}^{**} W_c^* W_h')^2 (y_{chi} - \bar{y}_{w,A})^2] = 0,000575$$

eşitliğinden belirlenir.

Örnekleme ve yanıtlamama ağırlıklandırması yapılan veriye tarama yöntemi uygulanmasıyla yanıtlayanlar örneğindeki gözlem değerlerine  $_{st} W_{kl}^{**} W_c^* W_h'$  ağırlıkları atanmıştır. Bu  $_{st} W_{kl}^{**} W_c^* W_h'$  ağırlıklarının varyans şişme faktörü değeri,

$$VIF = \left[ n_R \left( \sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} (tr W_{kl}^{**} W_c^* W_h')^2 \right) / \left( \sum_{k=1}^K \sum_{l=1}^L \sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{n_{klchR}} tr W_{kl}^{**} W_c^* W_h' \right)^2 \right]$$

$$VIF (tr W_{kl}^{**} W_c^* W_h') = 1,378709224$$

olarak belirlenir.

### 5.3.3. Ayarlama tahmin edicisi

Araştırma örneklemede ayarlama ağırlıklarıyla, yığının bilinen çeşitli karakteristiklerine göre örnek verisi yığın parametrelerine daha uygun hale getirilmeye çalışılır. Ayarlama ağırlıklandırması ile yardımcı değişkenlerin sonlu yığın toplamı bilgileri kullanılarak tahminlerin etkinliğinde artış sağlanmak istenir.

TNSA–2003 çalışmasında  $\bar{Y}$ , 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısı ayarlama tahmin edicisiyle belirlenmeye çalışılacaktır.  $y$  (15–49 yaşında evlenmiş kadınların canlı doğan çocuk sayısı) değişkeniyle ilişkili olan ve yığın toplamı bilinen  $x_2$ : Yaş grubu ve  $x_3$ : Eğitim durumu yardımcı değişkenlerine dayalı bir ayarlama tahmini yapılacaktır.

Yığında  $N=12630510$  sayıda 15–49 yaşında evlenmiş kadın vardır. 2000 Genel Nüfus Sayımından  $x_2$  ve  $x_3$  değişkenlerinin yığın toplamı bilgisi,

$$X_2 = 54319581 \quad X_3 = 15191989$$

olduğu belirlenmiştir.

TNSA–2003 çalışmasında  $W'_i$  düzeltilmiş örnekleme ağırlıklarında ayarlama yapılarak yeni  ${}_cW_i$  ayarlama ağırlıklarının elde edilmesi amaçlanır.  ${}_cW_i$  ayarlama ağırlıkları mümkün olduğunca  $W'_i$  ağırlıklarına yakın olacak şekilde belirlenir. (Uygulamada ayarlama ağırlıklandırmasının genel bir gösterimi olarak  $**$  üst indis gösterimi kullanımı tercih edildiği için  ${}_cW_i$  ayarlama ağırlıkları için  ${}_cW_i^{**}$  gösterimi kullanılacaktır.)

$\hat{X}_{HT}$  ile  $X$  arasındaki yakınlığı saptamada, toplam uzaklığı minimum yapılmasıyla  ${}_cW_i$  ayarlama ağırlıkları belirlenmeye çalışılır. Ki-kare uzaklık ölçümüne ve

Lagrange çarpanları yöntemine dayalı olarak  ${}_cW_i$  ayarlama ağırlıkları ve  $\lambda$  : Lagrange çarpanı,

$${}_cW_i^{**} = W_i'(1 + x_i'\lambda), \quad i=1, \dots, n$$

$$\lambda = \left[ \sum_{i=1}^{n_R} W_i' x_i x_i' \right]^{-1} (X - \hat{X})$$

eşitliklerinden belirlenir. Burada,  $q_i = 1$  alınmıştır. (Böylece ayarlama denklemi bir GREG tahmin edicisine dönüşmüştür.)

$\bar{Y}$  yığının ortalamasının ayarlama tahmin edicisi,

$$\bar{y}_{C,w} = \frac{\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* {}_cW_i^{**} y_{ci}}{\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* {}_cW_i^{**}} \quad {}_cW_i^{**} = W_i'(1 + x_i'\lambda)$$

eşitliğiyle elde edilir.  $\bar{y}_{C,w}$ , Hájek ağırlıklandırılmış türü bir tahmin edicidir.  ${}_cW_i^{**}$ ,  $i$  inci birim ayarlama ağırlığı ve  $W_c^*$ ,  $c$  inci ağırlıklandırma sınıfı yanıtlanama ağırlığıdır.  $W_i'$ ,  $i$  inci örnekleme birimi için düzeltilmiş örnekleme ağırlığıdır ve  $i \in h$  iken  $W_i' = W_h'$  dir,  $i = 1, \dots, 8075$ ,  $c = 1, 2$ .

Deville ve Särndal (1992), yardımcı değişkenler için  $\hat{X}$  örnek toplamının bir HT tahmin edicisi yapısında,

$$\hat{X} = \sum_{i=1}^n W_i' x_i$$

ile belirlenmesini önermiştir.  $x_2$  ve  $x_3$  yardımcı değişkenlerinin örnek toplamları,

$$\hat{X}_2 = \sum_{i=1}^{8075} W_i' x_{2i} = 54319581$$

$$\hat{X}_3 = \sum_{i=1}^{8075} W_i' x_{3i} = 15191989$$

olarak elde edilir.

$\lambda$  ve  ${}_c W_i^{**}$  ağırlıkları sırasıyla belirlenecektir.  $\lambda$ , Lagrange çarpanı,

$$(X - \hat{X}) = \begin{bmatrix} -906404,81 \\ -767935,6653 \end{bmatrix}$$

$$\sum_{i=1}^{8075} W_i' X_i X_i' = \begin{bmatrix} 277091028 & 67149102 \\ 67149102 & 32767037 \end{bmatrix}$$

$$\left[ \sum_{i=1}^{8075} W_i' X_i X_i' \right]^{-1} = \begin{bmatrix} 0,0000000072 & -0,0000000147 \\ -0,0000000147 & 0,0000000606 \end{bmatrix}$$

$$\lambda = \left[ \sum_{i=1}^{n_R} W_i' x_i x_i' \right]^{-1} (X - \hat{X}) = \begin{bmatrix} 0,0000000072 & -0,0000000147 \\ -0,0000000147 & 0,0000000606 \end{bmatrix} \cdot \begin{bmatrix} -906404,81 \\ -767935,6653 \end{bmatrix}$$

$$\lambda = \begin{bmatrix} 0,004784209 \\ -0,033240449 \end{bmatrix}.$$

Ayarlama ağırlıkları  ${}_c W_i^{**} = W_i'(1 + x_i' \lambda)$  formülünden  $i=1, \dots, 8075$ 'a kadar tüm birimler için elde edilir. Ağırlıklar Çizelge 5.15'in son sütununda verilmektedir.

Çizelge 5.15.  $x_2$  ve  $x_3$  yardımcı değişkenlerin gözlem değerleri, düzeltilmiş örnekleme ağırlıkları ve ayarlama ağırlıkları

| $Birim no$<br>$i$ | $x_2$ | $x_3$ | $W'_i$  | ${}_cW_i^{**}$ |
|-------------------|-------|-------|---------|----------------|
| 1                 | 2     | 0     | 1708,8  | 1725,149       |
| 2                 | 2     | 0     | 1708,8  | 1725,149       |
| 3                 | 5     | 0     | 1708,8  | 1749,675       |
| .                 | .     | .     | .       | .              |
| .                 | .     | .     | .       | .              |
| .                 | .     | .     | .       | .              |
| 633               | 4     | 0     | 2602,12 | 2651,919       |
| 634               | 7     | 1     | 2602,12 | 2602,771       |
| 635               | 6     | 0     | 2602,12 | 2676,817       |
| .                 | .     | .     | .       | .              |
| .                 | .     | .     | .       | .              |
| .                 | .     | .     | .       | .              |
| 6932              | 7     | 0     | 2235,4  | 2310,24        |
| 6933              | 2     | 1     | 2235,4  | 2182,463       |
| .                 | .     | .     | .       | .              |
| .                 | .     | .     | .       | .              |
| .                 | .     | .     | .       | .              |
| 8074              | 3     | 0     | 899,22  | 912,1298       |
| 8075              | 6     | 1     | 899,22  | 895,1454       |

Mesela, Çizelge 15'deki  ${}_cW_1^{**}$  ağırlığı,

$${}_cW_1^{**} = 1708 \left\{ 1 + (2 \ 1) \begin{bmatrix} 0,004784209 \\ -0,033240449 \end{bmatrix} \right\} = 1725,149$$

şeklinde elde edilir.

${}_cW_i^{**}$  ayarlama ağırlıkları için varyans şişme faktörü,

$$VIF = 1 + CV^2({}_cW_i^{**}) = \left[ n_R \left( \sum_{i=1}^{n_R} {}_cW_i^{**2} \right) / \left( \sum_{i=1}^{n_R} {}_cW_i^{**} \right)^2 \right]$$

$$VIF({}_cW_i^{**}) = 1,129405511$$

olarak belirlenir.

$\bar{Y}$  yığın ortalamasının ayarlama tahmin edicisi Hájek tahmin edicisi yapısında,

$$\bar{y}_{C,w} = \frac{\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* W_i^{**} y_{ci}}{\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* W_i^{**}} = \frac{36743869,11}{14031012,76} = 2,61187$$

eşitliğinden elde edilir. Ayarlama tahmin edicisiyle Türkiye’de 15–49 yaşında evlenmiş kadınların canlı doğan çocuk sayısı ortalaması 2,61187 olarak tahmin edilir. Ayarlama tahmin edicisi hesaplamaları aşağıda Çizelge 5.16’da verilmektedir.

Çizelge 5.16. Ayarlama tahmin edicisi hesaplamaları

| Birim no | $W'_h$  | $y$ | $x_1$ | $W_c^*$ | $x_2$ | $x_3$ | $W_i^{**}$    | $W_{kl}^{**} W_c^*$ | $W_{kl}^{**} W_c^* y_{klchi}$ |
|----------|---------|-----|-------|---------|-------|-------|---------------|---------------------|-------------------------------|
| 1        | 1708,8  | 2   | 1     | 1,148   | 2     | 0     | 1725,149      | 1980,471492         | 3960,942984                   |
| 2        | 1708,8  | 1   | 1     | 1,148   | 2     | 0     | 1725,149      | 1980,471492         | 1980,471492                   |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| 633      | 2602,12 | 3   | 1     | 1,148   | 4     | 0     | 2651,919      | 3044,403108         | 9133,209323                   |
| 634      | 2602,12 | 4   | 1     | 1,148   | 7     | 1     | 2602,771      | 2987,980713         | 11951,92285                   |
| 635      | 2602,12 | 7   | 1     | 1,148   | 6     | 0     | 2676,817      | 3072,986241         | 21510,90369                   |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| 6932     | 2235,4  | 1   | 1     | 1,148   | 7     | 0     | 2310,24       | 2652,155958         | 2652,155958                   |
| 6933     | 2235,4  | 1   | 1     | 1,148   | 2     | 2     | 2108,158      | 2420,165147         | 4840,330294                   |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| .        | .       | .   | .     | .       | .     | .     | .             | .                   | .                             |
| 8074     | 899,22  | 5   | 2     | 1,065   | 3     | 0     | 912,1298      | 971,4182012         | 4857,091006                   |
| 8075     | 899,22  | 8   | 2     | 1,065   | 6     | 1     | 895,1454      | 953,3298436         | 7626,638748                   |
|          |         |     |       |         |       |       | <b>Toplam</b> | 14031012,76         | 36743869,11                   |

$\bar{y}_{C,w}$ 'nin varyansının tahmini için,

$$\hat{V}_{HT}(\bar{y}_{C,w}) = \sum_{c=1}^C \sum_{i=1}^{n_{cR}} [(W_c^* W_i^{**})^2 (y_{chi} - \bar{y}_{C,w})^2] = 0,000533$$

eşitliğinden elde edilebilir.

Örnekleme ve yanıtlanmama ağırlıklandırması yapılan veriden ayarlama tahmin edicisi yöntemiyle yanıtlayanlar örneğindeki gözlem değerlerine  $W_c^* W_i^{**}$  ağırlıkları atanmıştır. Bu  $W_c^* W_i^{**}$  ağırlıklarının varyans şişme faktörü değeri,

$$VIF = 1 + CV^2(W_c^* W_i^{**}) = \left[ n_R \left( \sum_{c=1}^C \sum_{i=1}^{n_{cR}} (W_c^* W_i^{**})^2 \right) / \left( \sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* W_i^{**} \right)^2 \right]$$

$$VIF(W_c^* W_i^{**}) = 1,316002669$$

elde edilir.

TNSA–2003 araştırmasıyla yığın ortalamasının yanıtlayıcı örneği üzerinden ve ağırlıklandırma aşamalarından tahminleri belirlenmeye çalışılmıştır. Bu tahmin aşamalarından elde edilen yığın ortalaması tahmin edicilerinin varyans ve varyans şişme faktörlerinin kıyaslamaları aşağıda Çizelge 5.17’de verilmektedir.

Tahmin edicilerin varyans tahmin değerlerine bakıldığında, tahmin edicilerin duyarlılıkları açısından aralarında çok büyük farklılıklar olmadığı görülmektedir. En iyi tahmin sonucunu ise en küçük varyansa sahip ( $\hat{V}_{HT}(\bar{y}_{C,w})=0,000533$ ) ayarlama tahmin edicisinin verdiği görülmektedir.

Ayarlama ağırlıklandırması yöntemlerinde ağırlıklardaki değişkenliğin bir ölçümü olan varyans şişme faktörü değerlerine bakılabilir. Tarama ağırlıklarıyla

ağırlıklandırma düzeltmelerindeki değişkenliğin daha da azaltılmış olduğu değişkenliğinin en az olduğu görülmektedir.  $VIF(W_{kl}^{**}) = 1,018$  en küçük değerlidir.

Çizelge 5.17. Ağırlıklandırma aşamaları sonuçlarının ve ağırlıklandırma yöntemlerinin kıyaslanması

|                                                                    | Ortalama                 | $\hat{V}_{HT}(\hat{\theta})$ | $VIF$                      |                                       |
|--------------------------------------------------------------------|--------------------------|------------------------------|----------------------------|---------------------------------------|
| Yığın değeri                                                       | $\bar{Y} = 2,590$        | -                            | -                          | -                                     |
| Yanıtlayıcı ortalaması                                             | $\bar{y}_R = 2,780$      | 0,000587                     | -                          | -                                     |
| Örnekleme ağırlıklandırması                                        | $\bar{y}_{HT} = 2,629$   | 0,000547                     | -                          |                                       |
|                                                                    | $\bar{y}_w = 2,629$      |                              | -                          | -                                     |
| Yanıtlanmama ağırlıklandırması<br>Ağırlıklandırma sınıf düzeltmesi | $\bar{y}_{w,a} = 2,617$  | 0,000540                     | -                          | -                                     |
| Sonradan tabakalama                                                | $\bar{y}_{w,st} = 2,626$ | 0,000604                     | $VIF(W_{kl}^{**}) = 1,068$ | $VIF(W_{st}^{**} W_c^* W_h') = 1,438$ |
| Tarama                                                             | $\bar{y}_{w,tr} = 2,624$ | 0,000575                     | $VIF(W_{kl}^{**}) = 1,018$ | $VIF(W_{tr}^{**} W_c^* W_h') = 1,378$ |
| Ayarlama tahmin edicisi                                            | $\bar{y}_{C,w} = 2,612$  | 0,000533                     | $VIF(W_i^{**}) = 1,129$    | $VIF(W_c^* W_i^{**}) = 1,316$         |

Ancak çalışmada ayarlama ağırlıklandırması veriye örnekleme ve yanıtlanmama ağırlıklandırmaları yapıldıktan sonra uygulanmıştır. Dolayısıyla tahmin ediciler için üç aşama ağırlığın birlikte değişimi de önemlidir.

Bu nedenle, birlikte değişim için varyans şişme faktörünün ayarlama ağırlıklandırmasıyla olduğu ( $VIF(W_c^* W_i^{**}) = 1,316$ ) görülmektedir. Ayarlama tahmin edicisi yönteminin yığın ortalamasının tahminindeki ağırlıklar açısından değişimin en aza indirgendiği görülmektedir.

#### 5.4. Aşamalı Ayarlama

TNSA–2003 doğurgan yaşlardaki 15-49 yaşında evlenmiş kadınlar verisi üzerinden,  $\bar{Y}$  : 15-49 yaşında evlenmiş kadınlar için kadın başına düşen ortalama canlı doğan çocuk sayısının tahmini aşamalı (koşullu) ayarlama ağırlıklandırmasıyla yapılacaktır.

Ayarlama tahmin edicisinde koşullu dağılım özelliklerinden yararlanılabilir. Ağırlıklandırma yöntemlerinden ayarlama tahmin edicisi, yardımcı değişkenlerin modele aşamalı eklenmesine olanak sağlayarak, koşullu dağılım özelliklerinden yararlanılmasına ve yığın parametre tahminlerinin daha iyi tahminlerini elde edilmesine olanak sağlayacaktır.

Yardımcı değişkeninin bilinen yığın toplamı  $X$ , ayarlama ağırlıklarına bağlı olarak aşağıdaki gibi tanımlanır:

$$X = \sum_{i=1}^N x_i = \sum_{i=1}^n {}^c W_i x_i$$

Bu ayarlama kısıtının sağlanmış olmasına dikkat edilmelidir [Kott, 2003; Särndal, 2007]. Aşamalı ayarlamamanın her aşamasında bu kısıtın sağlanmış olması durumu kontrol edilmelidir. Çalışma kapsamında üç aşamalı ayarlama tahmin edicisi yapılmış ve her aşamada ayarlama kısıtının sağlanmış olduğu görülmüştür.

$y$  canlı doğan çocuk sayısı değişkeniyle ( $y = f\{x_2, x_3, x_4\}$ ) ilişkili olduğu belirlenen ve yığın toplamı bilinen  $x_2$  : Yaş grubu,  $x_3$  : Eğitim durumu ve  $x_4$  : yaşayan çocuk sayısı yardımcı değişkenleri varken, bu değişkenler ayarlama tahmin modeline aşamalı olarak alınacaktır.

2000 Genel Nüfus Sayımından 15-49 yaşında evlenmiş kadınlar için  $x_2$ ,  $x_3$  ve  $x_4$  değişkenlerinin yığın toplamı bilgilerinin

$X=[X_2 \ X_3 \ X_4]'=[54319581 \ 15191989 \ 30398682]'$  şeklinde öğrenilmiştir.

### 1. Aşama:

Aşamalı ayarlama tahmin edicisinde öncelikli olarak  $x_3$ : Eğitim durumu yardımcı değişkenine göre ayarlama yapılmak istenilsin.  $y = f(x_3)$  tahmin modeli vardır. Eğitim durumuna göre verinin dağılımı incelenerek  $\bar{Y}$ 'nin tahmini yapılmak istenilmektedir.

$\lambda$ , Lagrange çarpanı,

$$\lambda = \left[ \sum_{i=1}^n W_i' x_i x_i' \right]^{-1} (X_3 - \hat{X}_3) = 0,0000000305 (-767935,67) = -0,02344$$

olarak elde edilmiştir.  ${}_c W_i^{**}$  i inci örnekleme birimi ayarlama ağırlıkları  $\lambda$ 'ya ve  $W_i' = W_h'$  düzeltilmiş örnekleme ağırlıklarına dayalı olarak,

$${}_c W_i^{**} = W_i'(1 + x_i' \lambda) \quad i = 1, \dots, 8075$$

aşağıdaki gibi Çizelge 5.16'daki gibi belirlenir.

Birinci aşama için ayarlama kısıtının,

$$X_3 = \sum_{i=1}^N x_{3i} = \sum_{i=1}^n {}_c W_i x_{3i} = 15191989$$

şeklinde sağlanmış olduğu çalışmada görülmüştür.

Çizelge 5.18. Birinci aşama ayarlama tahmin edicisi hesaplamaları

| Birim no | $W'_h$  | $y$ | $x_1$ | $W_c^*$ | $x_3$ | ${}_cW_i^{**}$ | ${}_cW_i^{**}x_3$ | ${}_cW_i^{**}W_c^*$ | ${}_cW_i^{**}W_c^*y_{ci}$ |
|----------|---------|-----|-------|---------|-------|----------------|-------------------|---------------------|---------------------------|
| 1        | 1708,8  | 2   | 1     | 1,148   | 0     | 1708,79888     | 0                 | 1961,701114         | 3923,402228               |
| 2        | 1708,8  | 1   | 1     | 1,148   | 0     | 1708,79888     | 0                 | 1961,701114         | 1961,701114               |
| .        | .       | .   | .     | .       | .     | .              | .                 | .                   | .                         |
| 633      | 2602,12 | 3   | 1     | 1,148   | 0     | 2602,1226      | 0                 | 2987,236841         | 8961,710523               |
| 634      | 2602,12 | 4   | 1     | 1,148   | 1     | 2541,1387      | 2541,1387         | 2917,227285         | 11668,90914               |
| 635      | 2602,12 | 7   | 1     | 1,148   | 0     | 2602,1226      | 0                 | 2987,236841         | 20910,65789               |
| .        | .       | .   | .     | .       | .     | .              | .                 | .                   | .                         |
| 6932     | 2235,4  | 1   | 1     | 1,148   | 0     | 2235,3787      | 0                 | 2566,214795         | 2566,214795               |
| 6933     | 2235,4  | 1   | 1     | 1,148   | 2     | 2130,6010      | 4261,2021         | 2445,930018         | 2445,930018               |
| .        | .       | .   | .     | .       | .     | .              | .                 | .                   | .                         |
| 8074     | 899,22  | 5   | 2     | 1,065   | 0     | 899,2235       | 0                 | 957,6730754         | 4788,365377               |
| 8075     | 899,22  | 8   | 2     | 1,065   | 1     | 878,1491       | 878,1491          | 935,2288334         | 7481,830667               |
|          |         |     |       |         |       | <b>Toplam</b>  | 15191989          | 13909595,14         | 36371985,71               |

$${}^1\bar{y}_{C,w} = \frac{\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* {}_cW_i^{**} y_{ci}}{\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* {}_cW_i^{**}} = \frac{36371985,71}{13909595,14} = 2,61488457$$

Buradan, Türkiye’de 15–49 yaşında evlenmiş kadın başına düşen ortalama canlı doğan çocuk sayısı 2,61 olarak tahmin edilmiştir.

${}^1\bar{y}_{C,w}$  tahmin edicisinde  $W_c^*$ ,  $c$  inci ağırlıklandırma sınıfı yanıtlamama ağırlığı ve  ${}_cW_i^{**}$ ,  $i$  inci birimin ayarlama ağırlığıdır,  $c = 1, 2, i = 1, \dots, 8075$ . Burada  ${}_cW_i^{**}$  ayarlama ağırlıkları,  $W'_i$  düzeltilmiş örnekleme ağırlıklarına ve  $x_2$  yardımcı değişkenine dayalı belirlenmiştir.

Ayarlama tahmin edicisiyle belirlenen  ${}_cW_i^{**}$  ağırlıkların değişkenliği için varyans şişme faktörü,

$$VIF = 1 + CV^2({}_c W_i^{**}) = \left[ n_R \left( \sum_{i=1}^{n_R} {}_c W_i^{**2} \right) / \left( \sum_{i=1}^{n_R} {}_c W_i^{**} \right)^2 \right]$$

$$VIF({}_c W_i^{**}) = 1,20844078$$

eşitliğiyle elde edilir. Ayrıca yanıtlanmama ve ayarlama ağırlıklandırması uygulanan veri için  $W_c^* {}_c W_i^{**}$  ağırlıklarının varyans şişme faktörü değeri,

$$VIF = 1 + CV^2(W_c^* {}_c W_i^{**}) = \left[ n_R \left( \sum_{c=1}^C \sum_{i=1}^{n_{cR}} (W_c^* {}_c W_i^{**})^2 \right) / \left( \sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* {}_c W_i^{**} \right)^2 \right]$$

$$VIF(W_c^* {}_c W_i^{**}) = 1,34746476$$

olarak belirlenir.

${}^1\bar{y}_{C,w}$  'in varyansının tahmini için,

$$\hat{V}_{HT}({}^1\bar{y}_{C,w}) = \sum_{c=1}^C \sum_{i=1}^{n_{cR}} [(W_c^* {}_c W_i^{**})^2 (y_{chi} - \bar{y}_{C,w})^2] = 0,022497488$$

değeri elde edilmiştir. Bu varyans eşitliğinde,  $\sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* {}_c W_i^{**} = 1$  olmasını sağlayacak şekilde  $W_c^* {}_c W_i^{**}$  ağırlıkları kullanılmıştır.

## 2. Aşama.

TNSA–2003’de  $x_2$  yaş grupları değişkenine dayalı ayarlama ağırlıklandırması yapılan verinin,  $x_3$  değişkenine göre koşullu dağılımı incelenecektir.  $y = f(x_3 | x_2)$  koşullu tahmin modeli vardır.  $x_3$ : Eğitim durumu değişkenine göre koşullu ayarlama tahmininde, ayarlama ağırlıklarını belirlerken temel ağırlıklar olarak birinci aşamada

elde edilen  ${}_cW_i^{**}$  ayarlama ağırlıkları kullanılacaktır. Buradan, ikinci aşama ayarlama ağırlıkları,

$$\lambda = \left[ \sum_{i=1}^n {}_cW_i^{**} x_i x_i' \right]^{-1} (X_3 - {}_w\hat{X}_3)$$

$${}_c\tilde{W}_i^{**} = {}_cW_i^{**} (1 + x_i' \lambda) \quad i = 1, \dots, 8075$$

$$\hat{X}_3 = 15111466$$

formüllerinden belirlenir.  ${}_c\tilde{W}_i^{**}$   $i$  inci birimin ikinci aşama koşullu ayarlama ağırlığıdır ve  ${}_cW_i^{**}$  ağırlıkları ile  $x_3$  yardımcı değişkenine dayalı belirlenir,  $i = 1, \dots, 8075$ .

Çizelge 5.19. İkinci aşama ayarlama tahmin edicisi hesaplamaları

| Birim no | $W_h'$  | $y$ | $x_1$ | $W_c^*$ | $x_2$ | ${}_cW_i^{**}$ | ${}_cW_i^{**} x_3$ | ${}_cW_i^{**} W_c^*$ | ${}_cW_i^{**} W_c^* y_{ci}$ |
|----------|---------|-----|-------|---------|-------|----------------|--------------------|----------------------|-----------------------------|
| 1        | 1708,8  | 2   | 1     | 1,148   | 4     | 1707,912       | 6831,646           | 1960,682             | 3921,365                    |
| 2        | 1708,8  | 1   | 1     | 1,148   | 1     | 1708,577       | 1708,577           | 1961,446             | 1961,446                    |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| 633      | 2602,12 | 3   | 1     | 1,148   | 3     | 2601,109       | 7803,328           | 2986,073             | 8958,220                    |
| 634      | 2602,12 | 4   | 1     | 1,148   | 6     | 2600,096       | 15600,577          | 2984,910             | 11939,64                    |
| 635      | 2602,12 | 7   | 1     | 1,148   | 3     | 2601,109       | 7803,328           | 2986,073             | 20902,51                    |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| 6932     | 2235,4  | 1   | 1     | 1,148   | 4     | 2234,218       | 8936,872           | 2564,882             | 2564,882                    |
| 6933     | 2235,4  | 1   | 1     | 1,148   | 6     | 2233,638       | 13401,827          | 2564,216             | 2564,216                    |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| .        | .       | .   | .     | .       | .     | .              | .                  | .                    | .                           |
| 8074     | 899,22  | 5   | 2     | 1,065   | 2     | 898,9901       | 1797,980           | 957,4244             | 4787,122                    |
| 8075     | 899,22  | 8   | 2     | 1,065   | 3     | 898,8734       | 2696,620           | 957,3001             | 7658,401                    |
|          |         |     |       |         |       | <b>Toplam</b>  | 15191989           | 13901689,77          | 36288733,01                 |

$x_2$  modeldeyken  $x_3$ 'e göre koşullu ayarlama ile  $\bar{Y}$  yığın ortalamasının ayarlama tahmin edicisi,

$${}^2\bar{y}_{C,w} = \frac{\sum_{c=1}^C \sum_{i=1}^{n_c} W_c^* \tilde{W}_i^{**} y_{ci}}{\sum_{c=1}^C \sum_{i=1}^{n_c} W_c^* \tilde{W}_i^{**}} = \frac{36288733,01}{13901689,77} = 2,61038$$

elde edilmiştir.

${}_c\tilde{W}_i^{**}$  ağırlıkları ve  $W_c^* {}_c\tilde{W}_i^{**}$  ağırlıkları için varyans şişme faktörü sırasıyla aşağıdaki gibi elde edilmiştir.

$$VIF({}_c\tilde{W}_i^{**}) = \left[ n_R \left( \sum_{i=1}^{n_R} {}_c\tilde{W}_i^{**2} \right) / \left( \sum_{i=1}^{n_R} {}_c\tilde{W}_i^{**} \right)^2 \right] = 1,13083278$$

$$VIF(W_c^* {}_c\tilde{W}_i^{**}) = \left[ n_R \left( \sum_{c=1}^C \sum_{i=1}^{n_{cR}} (W_c^* {}_c\tilde{W}_i^{**})^2 \right) / \left( \sum_{c=1}^C \sum_{i=1}^{n_{cR}} W_c^* {}_c\tilde{W}_i^{**} \right)^2 \right] = 1,20893128$$

$x_2$  modeldeyken  $x_3$ 'e göre koşullu ayarlamayla yığın ortalaması tahmin edicisinin varyansı,

$$\hat{V}_{HT}({}^2\bar{y}_{C,w}) = \sum_{c=1}^C \sum_{i=1}^{n_{cR}} [(W_c^* {}_c\tilde{W}_i^{**})^2 (y_{chi} - \bar{y}_{C,w})^2] = 0,000512579$$

değeri elde edilmiştir.

### 3. Aşama.

$x_4$ , yaşayan çocuk sayısına göre veride  $\bar{Y}$  yığın ortalamasının ayarlama tahmin edicisi incelenecektir. Ancak, veri aşamalı bir şekilde öncelikle  $x_2$ 'ye göre ayarlanmış, daha sonra da ayarlanmış veri  $x_3$ 'e göre yeniden ayarlanmıştır. İki

aşamalı ayarlama ağırlıklandırmasının yapıldığı veride  $x_4$  değişkenine göre tahmin incelemesi yapılması düşünülmüştür.

Buradan, üçüncü aşama ayarlama ağırlıkları,

$$\lambda = \left[ \sum_{i=1}^n {}_c\tilde{W}_i^{**} x_i x_i' \right]^{-1} (X_4 - {}_w\hat{X}_4)$$

$${}_c\tilde{W}_i^{**} = {}_c\tilde{W}_i^{**} (1 + x_i' \lambda) \quad i = 1, \dots, 8075$$

$$\hat{X}_4 = 31108861,5$$

formüllerinden belirlenir.  ${}_c\tilde{W}_i^{**}$  i inci birim üçüncü aşama ayarlama ağırlığı,  ${}_c\tilde{W}_i^{**}$  ikinci aşama ayarlama ağırlığı ile  $x_4$  yardımcı değişkenine dayalı belirlenir. 2000 Genel Nüfus Sayımından  $x_4$  yardımcı değişkeninin yığın toplamı  $X_4 = 30398682$  bilgisi edinilmiştir.

İlk aşamada  $x_2$  'ye, ikinci aşamada  $x_3$  'e göre koşullu ayarlanmış verinin ayarlama ağırlıklarının kullanılmasıyla  $x_4$ , yaşayan çocuk sayısına göre koşullu ayarlama ağırlıklandırması yapılan örnek verisinden  $\bar{Y}$ , yığın ortalamasının tahmin edicisi,

$${}^3\bar{y}_{c,w} = \frac{\sum_{c=1}^C \sum_{i=1}^{n_c} {}_cW_c^* {}_c\tilde{W}_i^{**} y_{ci}}{\sum_{c=1}^C \sum_{i=1}^{n_c} {}_cW_c^* {}_c\tilde{W}_i^{**}} = 2,6103227$$

elde edilir.

${}_c\tilde{W}_i^{**}$  ve  ${}_cW_c^* {}_c\tilde{W}_i^{**}$  ağırlıkları için varyans şişme faktörü değerleri sırasıyla aşağıdaki gibidir.

$$VIF(\check{\check{W}}_i^{**}) = 1,121311588$$

$$VIF(W_c^* \check{\check{W}}_i^{**}) = 1,2079722$$

İlk aşamada  $x_2$  'ye, ikinci aşamada  $x_3$  'e göre ayarlamayla yığın ortalaması tahmin edicisinin varyansı,

$$\hat{V}_{HT}(\bar{y}_{c,w}) = \sum_{c=1}^C \sum_{i=1}^{n_{hR}} [(W_c^* \check{\check{W}}_i^{**})^2 (y_{chi} - \bar{y}_{c,w})^2] = 0,00050389$$

elde edilmiştir.

Ayrıca üç yardımcı değişkenin tümünün modelde aynı anda olması durumunu inceleyelim.  $x_2$ , yaş grupları,  $x_3$ , eğitim durumu ve  $x_4$ , yaşayan çocuk sayısı değişkenleri modelde olsun. Bu değişkenlere göre ayarlama yapılmasıyla  $\bar{Y}$  yığın ortalamasının ayarlama tahmin edicisi belirlenmek istenilebilir.

$\bar{Y}$  yığın ortalamasının ayarlama tahmin edicisi,

$$\lambda = \left[ \sum_{i=1}^n W_i' x_i x_i' \right]^{-1} (X - \hat{X})$$

$${}_c W_i^{**} = W_i' (1 + x_i' \lambda)$$

$$\bar{y}_{c,w} = \frac{\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{nch} W_c^* {}_c W_i^{**} y_{chi}}{\sum_{c=1}^C \sum_{h=1}^H \sum_{i=1}^{nch} W_c^* {}_c W_i^{**}} = 2,643157$$

şeklinde elde edilmiştir.

Çizelge 5.20. Aşamalı ayarlama ağırlıklandırması sonuçlarının kıyaslanması

|                                                        | Ortalama                 | $\hat{V}_{HT}(\hat{\theta})$ | $VIF$                                      |                                                   |
|--------------------------------------------------------|--------------------------|------------------------------|--------------------------------------------|---------------------------------------------------|
| $X_2$ ve $X_3$ 'e göre ayarlama tahmin edicisi         | $\bar{y}_{C,w}=2,612$    | 0,000533                     | $VIF({}_C W_i^{**})=1,129$                 | $VIF(W_c^* {}_C W_i^{**})=1,316$                  |
| 1. Aşama ayarlama                                      | $^1\bar{y}_{C,w}=2,6148$ | 0,002497                     | $VIF({}_C W_i^{**})=1,208$                 | $VIF(W_c^* {}_C W_i^{**})=1,3474$                 |
| 2. Aşama ayarlama                                      | $^2\bar{y}_{C,w}=2,6104$ | 0,000512                     | $VIF({}_C \tilde{W}_i^{**})=1,1308$        | $VIF(W_c^* {}_C \tilde{W}_i^{**})=1,208$          |
| 3. Aşama ayarlama                                      | $^3\bar{y}_{C,w}=2,6103$ | 0,0005034                    | $VIF({}_C \tilde{\tilde{W}}_i^{**})=1,121$ | $VIF(W_c^* {}_C \tilde{\tilde{W}}_i^{**})=1,2079$ |
| $X_2$ , $X_3$ ve $X_4$ 'e göre ayarlama tahmin edicisi | $\bar{y}_{C,w}=2,6431$   | 0,000542                     | $VIF({}_C W_i^{**})=1,162$                 | $VIF(W_c^* {}_C W_i^{**})=1,216$                  |

Çizelge incelendiğinde 3. Aşama ayarlama tahmin edicisinin en düşük varyans tahminine sahip olduğu ve varyans şişme faktörü değerinin de en küçük olduğu görülmektedir. Koşullu dağılımın özelliklerini kullanarak ayarlama tahmini yapmanın yığın parametrelerinin tahminin doğruluğunda artış sağlayacağı söylenebilir.

### 5.5. İki Değişkenin Birbirine Oranının Tahmini

Çalışmada doğurgan yaştaki 15–49 yaşında evlenmiş kadınların yaşayan çocuk sayısının canlı doğum sayısına oranıyla ilgileniyor olsun. Buradan,  $y$ , yaşayan çocuk sayısını ve  $x$ , canlı doğum sayısını gösterir.  $R = Y/X = \text{Yaşayan çocuk sayısı/Canlı doğum sayısı}$  tahmin edilmek istenilmektedir.

2000 Genel Nüfus Sayımından 15–49 yaşında evlenmiş kadınların gerçek yığın oranı  $R = Y/X = 30398682/32713021=0,929253$  şeklindedir. Buna göre Türkiye’de canlı doğan çocuklardan yaklaşık %93’ünün yaşıyor olduğu söylenebilir.

TNSA–2003’ün örnek seçiminde ağırlıklı, çok aşamalı, tabakalı küme örnekleme yaklaşımı kullanılmıştır. Tabakalı bir tasarım olduğu için yığın oranının bileşik veya

ayrılmış oran tahmini yapılabilir. Çalışmada  $R = \text{Yaşayan çocuk sayısı} / \text{Canlı doğum sayısı}$  oranının tahmininde, bilinen  $r_b$  bileşik oran tahmin edicisi ve Ayhan (2003)'ün önerdiği  ${}_A r_b$  bileşik oran tahmin edicisi belirlenecek elde edilen sonuçlar kıyaslanacaktır. Tahminleri yapmada  $W_h$ , örnekleme ağırlıkları için Çizelge 5.2'deki  $W'_h$  düzeltilmiş örnekleme ağırlıkları kullanılır.

### 5.5.1. Bileşik oran tahmin edicisi

Tabakalı tesadüfî örneklemede  $Y$  ve  $X$  yığın toplamalarının tahmini yapıldıktan sonra, tahmin edicilerin oran tahmininde kullanılması bileşik oran tahmini olarak adlandırılır. Çalışmada  $R$ , yığın oranının bileşik oran tahmin edicisi,

$$r_b = \frac{\sum_{h=1}^H W'_h y_h}{\sum_{h=1}^H W'_h x_h} = \frac{30466444}{33214885} = 0,917253$$

elde edilir. Buna göre, 15–49 yaşında evlenmiş kadınların canlı doğan çocuklarından %91,7'sinin yaşadığı tahmini yapılır.  $W'_h$ ,  $h$  incı tabaka düzeltilmiş örnekleme ağırlığıdır.

Yığın oranının tahmin için önerilen genel bir varyans tahmini,

$$\hat{V}(r_{tb}) = \frac{1}{X^2} \sum_{h=1}^H W_h^2 \frac{1}{n_h} (s_{yh}^2 - 2r_{tb} s_{yxh} + r_{tb}^2 s_{xh}^2)$$

eşitliğiyle verilmişti. Buna dayalı olarak, yığın oranının bileşik oran tahmin edicisinin varyans tahmini için,

$$\hat{V}(r_{tb}) = \hat{V}(r_b) = \frac{1}{X^2} \sum_{h=1}^H W_h'^2 \frac{1}{n_h} (s_{yh}^2 - 2r_b s_{yxh} + r_b^2 s_{xh}^2)$$

ile elde edilebilir.

Çizelge 5.21. Tabaka birim varyansları

| Tabaka | $W'_h$  | $S_{yh}^2$ | $S_{xh}^2$ | $S_{yxh}$ | Tabaka | $W'_h$  | $S_{yh}^2$ | $S_{xh}^2$ | $S_{yxh}$ |
|--------|---------|------------|------------|-----------|--------|---------|------------|------------|-----------|
| 1      | 280,51  | 3,157      | 5,763      | 1,88      | 21     | 1405,17 | 3,147      | 4,130      | 1,89      |
| 2      | 443,16  | 4,064      | 6,010      | 2,22      | 22     | 1419,55 | 1,510      | 1,867      | 1,32      |
| 3      | 516,83  | 1,758      | 2,628      | 1,38      | 23     | 1428,91 | 0,757      | 0,973      | 0,83      |
| 4      | 553,94  | 1,762      | 2,119      | 1,22      | 24     | 1450,24 | 1,181      | 1,316      | 1,33      |
| 5      | 730,72  | 3,243      | 4,145      | 1,9       | 25     | 1474,9  | 1,952      | 2,820      | 2,37      |
| 6      | 748,78  | 1,330      | 1,719      | 1,24      | 26     | 1576,19 | 1,051      | 1,384      | 1,20      |
| 7      | 776,14  | 9,318      | 12,35      | 1,21      | 27     | 1601,88 | 2,861      | 3,538      | 3,03      |
| 8      | 816,28  | 4,600      | 6,186      | 3,3       | 28     | 1625,4  | 3,151      | 4,250      | 1,76      |
| 9      | 817,22  | 5,255      | 7,000      | 2,2       | 29     | 1891,85 | 1,421      | 2,106      | 1,32      |
| 10     | 956,36  | 6,562      | 9,490      | 2,5       | 30     | 1929,41 | 1,028      | 1,835      | 1,23      |
| 11     | 1044,15 | 1,340      | 1,590      | 1,18      | 31     | 2119,52 | 1,080      | 1,467      | 1,25      |
| 12     | 1090,77 | 2,747      | 2,857      | 1,67      | 32     | 2245,95 | 1,713      | 2,475      | 1,44      |
| 13     | 1126,95 | 2,248      | 2,935      | 1,55      | 33     | 2273,7  | 1,874      | 3,713      | 1,66      |
| 14     | 1151,26 | 1,480      | 1,949      | 1,28      | 34     | 2360,39 | 1,282      | 1,653      | 1,19      |
| 15     | 1164,75 | 1,837      | 2,978      | 1,67      | 35     | 2541,93 | 2,691      | 3,078      | 1,7       |
| 16     | 1181,52 | 7,153      | 10,17      | 3,03      | 36     | 2960,91 | 1,595      | 2,120      | 1,33      |
| 17     | 1318,64 | 0,952      | 0,941      | 1,97      | 37     | 2971,38 | 1,869      | 2,272      | 1,44      |
| 18     | 1355,77 | 2,294      | 3,509      | 1,76      | 38     | 3753,62 | 2,056      | 3,400      | 1,63      |
| 19     | 1370,21 | 2,966      | 3,833      | 1,88      | 39     | 4465,31 | 3,188      | 4,027      | 1,89      |
| 20     | 1372,67 | 2,589      | 3,367      | 1,71      | 40     | 4910,24 | 1,341      | 1,737      | 1,21      |

Buradan gerekli hesaplamalar yapıldığında,

$$\hat{V}(r_b) = \frac{1}{X^2} \sum_{h=1}^H W_h^2 \frac{1}{n_h} (s_{yh}^2 - 2r_{tb} s_{yxh} + r_{tb}^2 s_{xh}^2)$$

$$\hat{V}(r_b) = 0,00000000322$$

olarak belirlenir.  $W'_h$  ağırlıklarının birlikte değişimi için,  $VIF$  değeri belirlenir.

$$VIF(W'_h) = \left[ H \left( \sum_{h=1}^H (W'_h)^2 \right) \right] / \left( \sum_{h=1}^H W'_h \right)^2 = 1,387289$$

### 5.5.2. Ayhan (2003) bileşik oran tahmin edicisi

Ayhan (2003) tarafından önerilen bileşik oran tahmin edicisi,

$${}_A r_b = \frac{\sum_{h=1}^H {}_A W_h y_h}{\sum_{h=1}^H {}_A W_h x_h} \quad {}_A W_h = \frac{P_0}{P_h}$$

formülüyle tanımlanmıştır.  ${}_A W_h$ , önerilen h ıncı tanım kümesi (veya tabaka) örnekleme ağırlığıdır.  ${}_A W_h$  ağırlıkları şöyle tanımlanmıştır:

$${}_A W_h = P_0 / P_h = \left[ \sum_{h=1}^H x_h / \sum_{h=1}^H (x_h / P_h) \right] / [(X/x) p_h]$$

$$P_0 = \sum_{i=1}^I x_h / \sum_{i=1}^I (x_h / P_h)$$

$$P_h = (X/x) p_h$$

$P_0$ : Ağırlıklandırılmış ve ağırlıklandırılmamış örnek çaplarının eşit olmalarının sağlanması için hesaplanan bir sabit

$p_h$ : h. tanım kümesindeki birimlerin örneğe seçilmesi olasılığı

Burada  $p_h = W_h'$  alınır.

$$X=32713021, x=22443 \quad X/x=1457,605$$

Çizelge 5.22.  $P_h$ ,  $x_h / P_h$  hesaplamaları ve  ${}_A W_h$  bileşik örnekleme ağırlıkları

| Tabaka | Tüm Örnekteki Kadın (Örnekleme) Ağırlığı $W'_h$ | $x_h$   | $P_h = (X / x)p_h$ | $x_h / P_h$ | ${}_A W_h = P_0 / P_h$ |
|--------|-------------------------------------------------|---------|--------------------|-------------|------------------------|
| 1      | 1708,8                                          | 1442,7  | 2490699,324        | 0,000579    | 0,688                  |
| 2      | 2602,12                                         | 969,84  | 3792777,695        | 0,000255    | 0,452                  |
| 3      | 324,996                                         | 123     | 473705,1158        | 0,000259    | 3,618                  |
| 4      | 1209,74                                         | 508,41  | 1763283,357        | 0,000288    | 0,972                  |
| 5      | 2455,64                                         | 362,34  | 3579272,524        | 0,000101    | 0,478                  |
| 6      | 5688,92                                         | 186,12  | 8292011,47         | 0,000022    | 0,206                  |
| 7      | 1680,23                                         | 197,58  | 2449056,487        | 0,000080    | 0,699                  |
| 8      | 1333,83                                         | 448     | 1944153,488        | 0,000234    | 0,881                  |
| 9      | 1655,51                                         | 102,12  | 2413025,303        | 0,000042    | 0,710                  |
| 10     | 5173,43                                         | 135,16  | 7540647,592        | 0,000017    | 0,227                  |
| 11     | 2235,38                                         | 135,2   | 3258227,678        | 0,000041    | 0,526                  |
| 12     | 598,79                                          | 125,95  | 872779,6397        | 0,000144    | 1,964                  |
| 13     | 2734,71                                         | 590,2   | 3986037,189        | 0,000148    | 0,430                  |
| 14     | 3442,6                                          | 205,92  | 5017837,953        | 0,000041    | 0,341                  |
| 15     | 846,61                                          | 650,7   | 1233995,175        | 0,000527    | 1,389                  |
| 16     | 1587,51                                         | 1425,92 | 2313910,396        | 0,000616    | 0,740                  |
| 17     | 1305,67                                         | 840,35  | 1903108,255        | 0,000441    | 0,900                  |
| 18     | 867,53                                          | 787,65  | 1264487,585        | 0,000622    | 1,355                  |
| 19     | 1349,47                                         | 371,28  | 1966949,916        | 0,000188    | 0,871                  |
| 20     | 641,78                                          | 930,1   | 935440,6673        | 0,000994    | 1,832                  |
| 21     | 513,44                                          | 1025,28 | 748375,855         | 0,001370    | 2,290                  |
| 22     | 945,73                                          | 953,47  | 1378469,728        | 0,000691    | 1,243                  |
| 23     | 946,82                                          | 2144,38 | 1380058,482        | 0,001553    | 1,242                  |
| 24     | 1527,76                                         | 78,88   | 2226820,459        | 0,000035    | 0,769                  |
| 25     | 1826,15                                         | 354,2   | 2661745,418        | 0,000133    | 0,644                  |
| 26     | 3430,47                                         | 449,55  | 5000157,602        | 0,000089    | 0,342                  |
| 27     | 4348,88                                         | 171,84  | 6338806,459        | 0,000027    | 0,270                  |
| 28     | 2191,87                                         | 85,02   | 3194808,713        | 0,000026    | 0,536                  |
| 29     | 2945,04                                         | 152     | 4292608,344        | 0,000035    | 0,399                  |
| 30     | 1263,75                                         | 68,04   | 1842006,83         | 0,000036    | 0,930                  |
| 31     | 1644,66                                         | 107,07  | 2397210,645        | 0,000044    | 0,715                  |
| 32     | 1570,77                                         | 254,66  | 2289510,638        | 0,000112    | 0,748                  |
| 33     | 1628,01                                         | 1092,52 | 2372942,069        | 0,000460    | 0,722                  |
| 34     | 1883,16                                         | 605,79  | 2744841,608        | 0,000220    | 0,624                  |
| 35     | 1590,35                                         | 602,82  | 2318049,901        | 0,000260    | 0,739                  |
| 36     | 2634,27                                         | 207,09  | 3839638,641        | 0,000053    | 0,446                  |
| 37     | 1855,92                                         | 375,48  | 2705137,342        | 0,000138    | 0,633                  |
| 38     | 1108,02                                         | 418,27  | 1615019,116        | 0,000258    | 1,061                  |
| 39     | 1368,89                                         | 783,9   | 1995255,968        | 0,000392    | 0,859                  |
| 40     | 899,22                                          | 1974,7  | 1310678,047        | 0,001506    | 1,307                  |
| Toplam |                                                 | 22443,5 |                    | 0,013092    |                        |

$$P_0 = \sum_{h=1}^H x_h / \sum_{h=1}^H (x_h / P_h) = 22443,5 / 0,013092223 = 1714262,026$$

Yanıtlanmama için ağırlıklandırma:

Örnekte yanıt alınamayan birimleri olduğu için örnekleme ağırlıklarında yanıtlanmama için ağırlıklandırma düzeltmesi yapılmalıdır.

$${}^A W_h^* = \frac{R_0}{R_h} \quad R_0 = \frac{\sum_{h=1}^H {}^A W_h x_h}{\sum_{h=1}^H ({}^A W_h x_h / R_h)} \quad h=1, \dots, H$$

$R_h$ , h ıncı tabaka yanıt oranıdır ve örnek için yanıt oranları aşağıdaki gibi olmuştur.

Çizelge 5.23.  $R_h$  yanıt oranları ve  $R_0$  denklemi için hesaplamalar

| Tabaka | $R_h$ | $n_{hR}$ | ${}^A W_h x_h$ | ${}^A W_h x_h / R_h$ | Tabaka | $R_h$ | $n_{hR}$ | ${}^A W_h x_h$ | ${}^A W_h x_h / R_h$ |
|--------|-------|----------|----------------|----------------------|--------|-------|----------|----------------|----------------------|
| 1      | 0,82  | 630      | 2465286        | 3007711,9            | 21     | 0,775 | 260      | 526419,8       | 678937,77            |
| 2      | 0,736 | 449      | 2523640        | 3427234,7            | 22     | 0,969 | 99       | 901725,2       | 930813,09            |
| 3      | 0,891 | 50       | 39974,51       | 44872,971            | 23     | 0,917 | 107      | 2030342        | 2213915,5            |
| 4      | 0,971 | 34       | 615043,9       | 633133,44            | 24     | 0,946 | 270      | 120509,7       | 127423,38            |
| 5      | 0,9   | 269      | 889776,6       | 988509,07            | 25     | 0,853 | 557      | 646822,3       | 758053,41            |
| 6      | 0,958 | 115      | 1058822        | 1105702,2            | 26     | 0,899 | 286      | 1542168        | 1716072,1            |
| 7      | 0,809 | 183      | 331979,8       | 410348,86            | 27     | 0,929 | 343      | 747311,5       | 804735,05            |
| 8      | 0,932 | 94       | 597555,8       | 641451,46            | 28     | 0,958 | 159      | 186352,8       | 194499,83            |
| 9      | 0,939 | 135      | 169060,7       | 180042,87            | 29     | 0,898 | 267      | 447646,1       | 498410,29            |
| 10     | 0,932 | 89       | 699240,8       | 750154,29            | 30     | 0,967 | 153      | 85985,55       | 88960,829            |
| 11     | 0,94  | 48       | 302223,4       | 321514,23            | 31     | 0,964 | 136      | 176093,7       | 182744,35            |
| 12     | 0,766 | 200      | 75417,6        | 98419,969            | 32     | 0,983 | 59       | 400012,3       | 406792,16            |
| 13     | 0,904 | 46       | 1614026        | 1786188,6            | 33     | 0,935 | 355      | 1778633        | 1901711,9            |
| 14     | 0,963 | 26       | 708900,2       | 736165,58            | 34     | 0,908 | 126      | 1140799        | 1256089,7            |
| 15     | 0,846 | 62       | 550889,1       | 650900,51            | 35     | 0,958 | 384      | 958694,8       | 1000319,5            |
| 16     | 0,974 | 38       | 2263662        | 2323232,3            | 36     | 0,951 | 151      | 545531         | 573688,93            |
| 17     | 0,945 | 65       | 1097220        | 1161364,9            | 37     | 0,91  | 371      | 696860,8       | 765865,46            |
| 18     | 1     | 21       | 683310         | 683310               | 38     | 0,909 | 195      | 463451,5       | 510077,56            |
| 19     | 0,929 | 55       | 501031,2       | 539481,08            | 39     | 0,926 | 742      | 1073073        | 1158541,5            |
| 20     | 0,977 | 43       | 596919,6       | 610801,43            | 40     | 0,956 | 403      | 1775690        | 1857464,1            |
| Toplam |       |          |                |                      |        |       | 8075     | 34028101       | 37725657             |

$$R_0 = \frac{\sum_{h=1}^H {}^A W_h x_h}{\sum_{h=1}^H ({}^A W_h x_h / R_h)} = \frac{34028101}{37725657} = 0,901988$$

Çizelge 5.24.  ${}_A W_h^*$  yanıtlamama için bileşik ağırlıklar

| Tabaka | ${}_A W_h^* = R_0 / R_h$ | ${}_A W_h {}_A W_h^*$ | Tabaka | ${}_A W_h^* = R_0 / R_h$ | ${}_A W_h {}_A W_h^*$ |
|--------|--------------------------|-----------------------|--------|--------------------------|-----------------------|
| 1      | 1,1004                   | 1880,447              | 21     | 1,1633                   | 597,2943              |
| 2      | 1,2249                   | 3187,459              | 22     | 0,9310                   | 880,5547              |
| 3      | 1,0125                   | 329,0642              | 23     | 0,9835                   | 931,2369              |
| 4      | 0,9285                   | 1123,265              | 24     | 0,9537                   | 1457,079              |
| 5      | 1,0020                   | 2460,737              | 25     | 1,0570                   | 1930,422              |
| 6      | 0,9419                   | 5358,535              | 26     | 1,0037                   | 3443,17               |
| 7      | 1,1149                   | 1873,316              | 27     | 0,9712                   | 4224,055              |
| 8      | 0,9682                   | 1291,477              | 28     | 0,9414                   | 2063,474              |
| 9      | 0,9605                   | 1590,252              | 29     | 1,0042                   | 2957,633              |
| 10     | 0,9676                   | 5006,144              | 30     | 0,9331                   | 1179,33               |
| 11     | 0,9595                   | 2144,986              | 31     | 0,9360                   | 1539,491              |
| 12     | 1,1770                   | 704,8326              | 32     | 0,9172                   | 1440,83               |
| 13     | 0,9982                   | 2729,789              | 33     | 0,9644                   | 1570,06               |
| 14     | 0,9366                   | 3224,615              | 34     | 0,9931                   | 1870,249              |
| 15     | 1,0657                   | 902,2662              | 35     | 0,9411                   | 1496,759              |
| 16     | 0,9257                   | 1469,597              | 36     | 0,9485                   | 2498,724              |
| 17     | 0,9547                   | 1246,549              | 37     | 0,9913                   | 1839,783              |
| 18     | 0,9019                   | 782,5019              | 38     | 0,9927                   | 1099,969              |
| 19     | 0,9712                   | 1310,616              | 39     | 0,9738                   | 1333,067              |
| 20     | 0,9229                   | 592,3403              | 40     | 0,9435                   | 848,4382              |

Sonradan tabakalama için bileşik ağırlıklandırma:

TNSA–2003 araştırmasında  $x_2$ :Yaş grubu kategorik yapıdaki yardımcı değişiyile sonradan tabakalama yapılacağına karar verilmiştir,  $x_2$ ,  $K=7$  düzeylidir,  $k=1,...,7$ . Amaç eğitim durumuna göre örnek dağılımının yığın dağılımına uygunluğunu sağlamaktır.

Sonradan tabakalama bağımsız alt sınıflarda ilave ağırlıklar hesaplanmasını gerektirir. Ayhan (2003) çalışmasında,  $k$  inci sonradan tabaka için bileşik ağırlıklandırmada,

$${}_A W_k^{**} = \frac{X_k}{X} \quad k=1,...,7$$

sonradan tabaka ağırlıkları formülü önerilmiştir. Buradan,  $R$  yığın oranı tahmin edicisi,

$${}_A r_b = \frac{\sum_{k=1}^K {}_A W_k^{**} \sum_{h=1}^H W'_h {}_A W_h^* y_{khR}}{\sum_{k=1}^K {}_A W_k^{**} \sum_{h=1}^H W'_h {}_A W_h^* x_{khR}} \quad {}_A W_h = \frac{P_0}{P_h} \quad {}_A W_h^* = \frac{R_0}{R_h} \quad {}_A W_k^{**} = \frac{X_k}{X}$$

formülüyle hesaplanır.  ${}_A W_k^{**}$ , sonradan tabaka ağırlıkları aşağıda hesaplanacaktır.

Öncelikli olarak  ${}_A r_b$  tahmin edicisinin pay ve paydasındaki  $\sum_{h=1}^H {}_A W_h {}_A W_h^* y_{hkR}$  ve  $\sum_{h=1}^H {}_A W_h {}_A W_h^* x_{hkR}$  toplam değerleri Çizelge 5.20'deki gibi belirlenir.

Çizelge 5.25. Yaş gruplarına göre  $n_{kR}$  ve  $N_k$  dağılımı ile  $\sum_{h=1}^H {}_A W_h {}_A W_h^* y_{hkR}$  ve

$\sum_{h=1}^H {}_A W_h {}_A W_h^* x_{hkR}$  ağırlıkları

| $x_2$ : Yaş grubu | $N_k$   | $n_{kR}$ | $\sum_{h=1}^H W'_h {}_A W_h^* x_{khR}$ | $\sum_{h=1}^H W'_h {}_A W_h^* y_{khR}$ |
|-------------------|---------|----------|----------------------------------------|----------------------------------------|
| 15-19             | 453511  | 240      | 180,5941                               | 194,5517                               |
| 20-24             | 1727365 | 1080     | 1749,078                               | 1842,871                               |
| 25-29             | 2378665 | 1516     | 4214,355                               | 4002,468                               |
| 30-34             | 2244391 | 1506     | 5734,09                                | 5362,331                               |
| 35-39             | 2282957 | 1410     | 6394,855                               | 5852,954                               |
| 40-44             | 1922351 | 1297     | 6846,528                               | 6141,907                               |
| 45-49             | 1621270 | 1026     | 5581,468                               | 4919,209                               |

Çalışmada  $y$ , yaşayan çocuk sayısını ve  $x$ , canlı doğum sayısını gösteriyorken  $R = Y / X = \text{Yaşayan çocuk sayısı} / \text{Canlı doğum sayısı}$  oranı için  ${}_A r_b$  tahmin edicisi belirlenecektir. Bunun için, yaş gruplarına göre sonradan tabakalanmış örnek için  $x$ , canlı doğum sayısı değişkeninin yığın toplamalarının biliniyor olması istenir.

2000 Genel Nüfus Sayımından  $X_k$ ,  $k$  ıncı sonradan tabaka yığın toplamı değerleri öğrenilmiştir.  ${}_A W_k^{**} = X_k / X$ ,  $k$  ıncı sonradan tabaka ağırlıkları aşağıdaki gibi belirlenir.

Çizelge 5.26.  $X_k$  yığın toplamaları ve  ${}_A W_k^{**}$  sonradan tabaka ağırlıkları

| $x_2$ :Yaş grubu | $X_k$    | ${}_A W_k^{**} = X_k / X$ | $W_k^{**} \sum_{h=1}^H {}_A W_h {}_A W_h^* y_{khR}$ | $W_k^{**} \sum_{h=1}^H {}_A W_h {}_A W_h^* x_{khR}$ |
|------------------|----------|---------------------------|-----------------------------------------------------|-----------------------------------------------------|
| 15-19            | 294628   | 0,0359                    | 7883,072                                            | 8452,682                                            |
| 20-24            | 2078364  | 0,1368                    | 263369,562                                          | 275931,954                                          |
| 25-29            | 4522719  | 0,1883                    | 848192,363                                          | 886883,534                                          |
| 30-34            | 5700038  | 0,1777                    | 1033509,385                                         | 1099511,889                                         |
| 35-39            | 7036619  | 0,1807                    | 1173253,904                                         | 1279513,638                                         |
| 40-44            | 6707033  | 0,1522                    | 1011205,47                                          | 1126042,687                                         |
| 45-49            | 6394157  | 0,1284                    | 717786,4869                                         | 820766,657                                          |
| Toplam           | 32733558 | -                         | 5055200,244                                         | 5497103,043                                         |

$${}_A r_b = \frac{\sum_{k=1}^K {}_A W_k^{**} \sum_{h=1}^H W_h' {}_A W_h^* y_{khR}}{\sum_{k=1}^K {}_A W_k^{**} \sum_{h=1}^H W_h' {}_A W_h^* x_{khR}} = \frac{5055200,244}{5497103,043} = 0,919612$$

Buradan, canlı doğan çocuklardan %91,9'unun yaşıyor olduğu tahmini yapılmıştır.

Çizelge 5.27.  $X_k$  yığın toplamaları ve  ${}_A W_k^{**}$  sonradan tabaka ağırlıkları

| $x_2$ :Yaş grubu | $W_k^{**} {}_A W_h {}_A W_h^*$ | $s_{yk}^2$ | $s_{xk}^2$ | $s_{yxk}$ |
|------------------|--------------------------------|------------|------------|-----------|
| 15-19            | 14040,19763                    | 0,50       | 0,25       | 1,32      |
| 20-24            | 226920,2164                    | 0,92       | 0,84       | 1,7       |
| 25-29            | 448066,179                     | 1,63       | 2,65       | 2,11      |
| 30-34            | 415966,1629                    | 1,44       | 2,091      | 1,33      |
| 35-39            | 399255,382                     | 3,36       | 11,33      | 6,7       |
| 40-44            | 303439,722                     | 1,50       | 2,25       | 1,8       |
| 45-49            | 212837,3363                    | 1,39       | 1,95       | 2,7       |

${}_A r_b$  tahmin edicisinin varyans tahmini için,

$$\hat{V}({}_A r_b) = \frac{1}{X^2} \sum_{k=1}^K W_k^2 \frac{1}{n_h} (s_{yk}^2 - 2r_b s_{yxk} + r_b^2 s_{xk}^2)$$

$$\hat{V}({}_A r_b) = 0,0000000029$$

ile elde edilebilir. Burada,  $W_k = W_k^{**} W_h W_h^*$  şeklindedir.  $VIF(W_k^{**} W_h W_h^*) = 1,238408$  olarak hesaplanmıştır.

Çizelge 5.28. Bileşik oran tahmin edicileri sonuçlarının kıyaslanması

|                                                      | Ortalama           | $\hat{V}_{HT}(\hat{\theta})$ | $VIF$                                |
|------------------------------------------------------|--------------------|------------------------------|--------------------------------------|
| Bileşik oran tahmin edicisi                          | $r_b = 0,917$      | 0,0000000032                 | $VIF(W_h') = 1,387289$               |
| Ayhan (2003)'ün önerdiği bileşik oran tahmin edicisi | ${}_A r_b = 0,919$ | 0,0000000029                 | $VIF(W_k^{**} W_h W_h^*) = 1,238408$ |

$r_b$  bileşik oran tahmin edicisi ile Ayhan (2003)'ün önerdiği bileşik oran tahmin edicisinin varyans tahmini değerlerine bakıldığında, aynı duyarlılıkta sonuçlar verdikleri görülmektedir. Ağırlıkların değişkenlik düzeylerine bakıldığında ise  ${}_A r_b$  tahmin edicisindeki ağırlıkların daha az değişken olduğu görülmektedir.

## 6. SONUÇ

Örnekleme ve örnekleme dışı hataları düzeltmek ve örnek araştırmalarındaki tahminlerin etkinliğini arttırmak amacıyla ağırlıklandırma yapılması ve ağırlıklandırma yöntemlerinin kullanımı yaygın bir uygulamadır. Ağırlıklandırma gerektiren veri ağırlıklandırılmamışsa parametre tahminleri için yanlış sonuçlar elde edileceği söylenebilir.

Veride uygun ağırlıklandırmalar yapılarak her bir örnekleme biriminin seçim olasılığındaki farklılıkları denkleştirilmeye ve yığının daha iyi temsil edilmesi amacıyla parametre tahminlerinde düzeltmeler yapılmaya çalışılması hiç bir zaman yapılan ‘yanlış’ bir işlem olmayacaktır.

Ağırlıklandırma adımsal olarak sırasıyla örnekleme ağırlıklandırması, yanıtlanmama ağırlıklandırması ve ayarlama ağırlıklandırması şeklinde veride yürütülebilir. Örnekleme ağırlıklandırması, örnekleme yöntemleri farklı seçim olasılıklarıyla birimlerin örneğe seçilmesine yol açtığında veriye uygulanır. Örnekleme yöntemlerinde eşit olmayan seçim olasılıklarını telafi etmek amacıyla gözlem değerlerine ağırlıklar atanır.

Yanıtlanmama ve örnek dağılımının yığın dağılımına uyumunu sağlamak amacıyla yanıtlanmama ve ayarlama ağırlıklandırmalarına başvurulabilir. Yanıtlanmama ve ayarlama ağırlıklandırmalarını yapmak amacıyla ağırlıklandırma yöntemleri geliştirilmiştir.  $\theta$  parametresinin tahmininde ağırlıklandırma yöntemlerinin kullanımı,  $\theta$  için daha iyi tahminler sağlanması açısından önerilebilir.

Ağırlıklandırma yöntemlerinin uygulanması yardımcı bilgi (değişken) olması durumuna bağlıdır. Yardımcı bilgi yoksa yanıtlanmama veya ayarlama ağırlıklandırması yapılamaz. Ağırlıklandırma düzeltmelerinde, ilgilenilen çalışma değişkeniyle ilişkili yardımcı değişkenler tahmin işlemine dahil edilerek parametre tahminlerinin etkinliği artırılabilir. Ağırlıklandırma yöntemleriyle hanelerin

demografik özellikleri, yerleşim bölgesi, sosyal ekonomik göstergeleri, aile mesleği, eğitim düzeyi, ev koşulları gibi yardımcı bilgilere dayalı örnek dağılımı incelenir.

Araştırmalarda tahmin aşamasında kullanılacak yardımcı değişkenlerin seçimi önemlidir. Ağırlıklandırma düzeltmeleri yapmada kullanmak için çok fazla sayıda yardımcı değişken olduğunda, yalnızca bazıları seçilebilir. Yardımcı değişkenlerin seçiminde yanıtlanmama düzeltmelerinde esas olarak göze yanıt olasılıklarının tahminini veriyor olması ve ayarlama ağırlıklandırmalarında yardımcı değişkenlerin araştırma değişkenleriyle ilişkili olması durumları dikkate alınmalıdır.

Bu çalışmada Horvitz-Thompson ve Hájek tahmin edicileri, ağırlıklandırma aşamaları ve ağırlıklandırma yöntemleri ele alınmıştır. Literatürde Hájek tahmin edicilerine dayalı başvuru kaynağının az olduğu görülmüştür. Çalışmada ağırlıklandırma aşamaları ve ağırlıklandırma yöntemlerinde Horvitz-Thompson ve Hájek tahmin edicilerinin yapıları verilmiştir. Ağırlıklandırma yöntemlerinden; ağırlıklandırma sınıf düzeltmesi, sonradan tabakalama, tarama, ayarlama ve genelleştirilmiş regresyon tahmin edicisi incelenmiştir.

Horvitz-Thompson tahmin edicisinde ağırlıklar toplamının yığın çapına eşit olması koşulu aranır. Örneklem ağırlıklandırması yapılan veride yanıtlanmama ve ayarlama ağırlıklandırması yapılması durumlarında, gözlem değerlerine atanan ağırlıkların toplamı  $N$  yığın çapına eşit olmayacaktır. Horvitz-Thompson tahmin edicisi hesaplamaları yapılamayacak, Hájek tahmin edicisiyle istenen tahminler belirlenebilecektir.

Horvitz-Thompson tahmin edicisi hesaplamalarında ağırlıklar toplamı  $N$ 'i verecek şekilde yeniden tanımlama yapıldığında ise Horvitz-Thompson tahmin edicisi yapısındaki hesaplamalar yürütülebilir. Ancak uygulamalı araştırmalarda gerçek seçim olasılıklarının tam olarak belirlenemediği, yanıtlanmama ve ayarlama ağırlıklandırmasının yapılacağı durumlarda Hájek yapısındaki tahmin edicilerini kullanımları önerilebilir.

Çalışmanın uygulama kısmında 2003 Türkiye Nüfus ve Sağlık Araştırması verisiyle örnekleme ağırlıklandırması, yanıtlanmama ağırlıklandırması ve ayarlama ağırlıklandırması incelenerek elde edilen tahmin sonuçları kıyaslanmıştır.

TNSA–2003 verileri için Horvitz-Thompson ve Hájek tahmin edicileriyle ağırlıklandırma incelenmesi yapılması amaçlanmıştır. Ancak, yanıtlanmama ve ayarlama ağırlıklandırması aşamalarında gözlemlere atanana ağırlıklar toplamı  $N$  yığın çapına eşit olmayacağından Horvitz-Thompson tahmin edicisi hesaplamalarına başvurulmamıştır. İstenen parametre tahminleri Hájek tahmin edicisi yapılarında belirlenmiştir.

Uygulama çalışmasında ağırlıklandırma yöntemlerinden ağırlıklandırma sınıf düzeltmesi, sonradan tabakalama, tarama ve ayarlama tahmin edicisi incelenmiştir. Uygulamada ayarlama tahmin edicisinin tüm ağırlıklandırma yöntemleri içinde en iyi tahmin sonucunu sağlayan ağırlıklandırma yöntemi olduğu belirlenmiştir.

Kalton ve Flores-Cervantes (2003), göze ağırlıklandırması, tarama, genelleştirilmiş regresyon tahmin edicisi, lojistik regresyon ağırlıklandırması, yöntemlerin karışımları ve düzeltme türlerini kısıtlayan yöntemleri kıyaslamıştır. Kalton ve Flores-Cervantes'in çalışmasında ayarlama tahmin edicisi incelenmemiştir. Ancak tarafımızca aynı çalışmanın verisinde ayarlama tahmin edicisi incelendiğinde ve Kalton'la yapılan müzakerelerde ayarlama tahmin edicisinin en iyi tahmin sonucunu verecek ağırlıklandırma yöntemi olduğu görülmüştür.

Yardımcı bilgileri tahmin aşamasında kullanan en yaygın yöntemler regresyon ve oran tahmin edicisidir. Ayarlama tahmin edicisi regresyon tahmin edicisi yöntemine benzeyen ve kullanımı gittikçe yaygınlaşan bir yöntemdir. Ayarlama tahmin edicisinin genel bir ağırlıklandırma yöntemi olduğu, diğer ağırlıklandırma yöntemlerinin ayarlama ağırlıklandırması altında sınıflandırılabileceği görülmektedir.

Ağırlıklandırma sınıf düzeltmesi, sonradan tabakalama, tarama yöntemlerinde yardımcı değişkenlere göre sınıflama yaparak, bu sınıflarda ağırlıkları yığın ve örnek çaplarına dayalı belirlenir. Ayarlama tahmin edicisinde ise ağırlıkları hesaplamada yardımcı değişkenin örnekten gözlenen değerleri ve yardımcı değişken yığın toplamı bilgileri kullanılır. Ayarlama ağırlıklarında yardımcı değişkenlere ilişkin daha çok bilgi kullanılmış olur.

Ayhan (2003)'ün önerdiği bileşik oran tahmin edicisinde de ağırlıklar, yığın ve örnek çaplarına, hem de yardımcı değişken bilgisine dayalı tanımlanmıştır. Uygulama çalışmasında ayarlama ve Ayhan (2003)'ün önerdiği bileşik oran tahmin edicisinin parametrelerin daha iyi tahminlerini sağlayan tahmin ediciler olduğu görülmüştür. Buradan, ağırlıkların hesaplamalarında yardımcı değişkenlerin değerlerini de kullanan ağırlıklandırma yöntemlerinin iyi tahminler sağlayacağı söylenebilir.

Araştırmacılar yardımcı değişkenler için mükemmel tahminler sağlayan ağırlıkların, ilgilenilen değişken için de iyi tahminler sağlayacağına inanmaktadır. Ağırlıkların uygulanmasıyla yansız tahminler elde edilecektir denilemez, ama yansızlığa çok yakın olunacağı söylenebilir [Dewille ve Särndal, 1992]. İlgilenilen değişkenle ilişkili olduğu bilinen yardımcı değişken veya değişkenlere göre ağırlıklandırılmış veriyi kullanmak bilgide kazanca yol açacağı söylenebilir.

Ayarlama tahmin edicisi yardımcı değişkenlerin modele aşamalı eklenmesine olanak sağlayarak aşamalı ağırlıklandırmayla parametrelerin daha iyi tahminlerinin elde edildiği belirlenmiştir. Diğer ağırlıklandırma yöntemlerinin yapısı gereği aşamalı ağırlıklandırma mümkün değildir. Aşamalı ayarlama tahmin edicisi TNSA-2003 araştırmasında uygulanarak sonuçları incelenmiş ve aşamalı ayarlama yapmanın tahminlerin doğruluğunda artış sağlayacağı görülmüştür.

Pek çok araştırmada genellikle bir değişkene göre dağılımı incelenen örnek verisinin, başka değişkenlere dayalı incelemeleri de yapılmak istenilir. Böyle durumlarda koşullu dağılım kullanımının daha iyi tahminlerin yapılmasına katkı sağlayacağı

açıktır. Bilinenden yola çıkarak bilinmeyene ulaşmak gerçek sonuçlara biraz daha yaklaşmamızı sağlayacaktır.

## KAYNAKLAR

Ardilly, P. And Tillé, Y., “Sampling methods: exercises and solutions”, Translated from French by Leon Jang, **Springer Science+Business Media**, United States of America, 209-292 (2006).

Ayhan, H.Ö., “Türkiye doğurganlık araştırmasında (1978) yanıtlamama kaynakları ve yanlılığı”, **H. Ü. Nüfus Etütleri Enstitüsü Nüfus Bilim Dergisi**, 2 (3):104-148 (1981).

Ayhan, H.Ö., “Combined weighting procedures for post-survey adjustment in complex sample surveys,” **Bulletin International Stat. Inst.**, 60 (1): 53-54 (2003).

Basu, D., “An essay on the logical foundations of survey sampling”, **Foundations of Statistical Inference**, Toronto, 203-233 (1971).

Bethlehem, J.G., Keller, W.J., “Linear weighting of the sample survey data”, **J. Of Official Statistics**, 3 (2): 141-153 (1987).

Bethlehem, J.G. And Kersten, H.M.P., “On the treatment of nonresponse in sample surveys”, **J. Of Official Statistics**, 1 (3): 287-300 (1985).

Carlson, B. And Williams, S., “A comparison of two methods to adjust for non-response: propensity modeling and weighting class adjustments”, **American Stat. Ass. Section on Survey Research Methods**, Atlanta, 5-9 (2001).

Cassel, C.M., Särndal, C.E. And Wretman, J.H., “Some results on generalized difference estimation and generalized regression estimation for finite populations”, **Biometrika**, 63: 615-620 (1976).

Cho, C. And Chiang, A., “A comparison between ratio estimation and post-stratification”, **American Stat. Ass. Section on Survey Research Methods**, Washington, 2825-2830 (2006).

Churchill, G.A., “Basic marketing research”, **The Dryden Press**, South-Western, 70-99 (1996).

Cochran, W.G., “Sampling techniques”, **John Wiley & Sons Inc.**, New York, 18-64, 189-203 (1977).

Creedy, J., “Survey reweighting for tax microsimulation modelling”, **J A Bishop and Y Amiel Eds.**, 12: 229-49 (2003).

Creel, D.V. And Fahimi, M., “Multidimensional control totals for poststratified weights”, **American Stat. Ass. Section on Survey Research Methods**, Minneapolis, 2923-2927 (2005).

Crockett, A., "Weighting the social surveys", **UK Data Archive and Inst. for Social and Econ. Research**, United Kingdom, 1-24 (2008).

Çingil, H., "Örnekleme kuramı", **H.Ü. Fen Fakültesi Basımevi**, Ankara, 1-50 (1994).

Deming, W.E. And Stephen, F.F., "On a least square adjustment of a sampled frequency table when the expected marginal totals are known", **Annals of Mathematical Statistics**, 11(4): 427-444 (1940).

Deville, J.C. And Särndal, C.E., "Calibration estimators in survey sampling", **J. of The American Statistical Ass.**, 87: 376-382 (1992).

Deville, J.C., Särndal, C.E., Sautory, O., "Generalized raking procedures in survey sampling", **Journal of The American Statistical Association**, 88: 1013-1020 (1993).

DHS, "Designing Household Survey Samples: Practical Guidelines", **Economic and Social Affairs Statistics Division**, New York, 1-45 (2005).

Dorfman, A.H. And Valliant, R., "The Hájek Estimator Revisited", **American Stat. Ass. Section on Survey Research Methods**, Washington, 760-765 (1997).

Einbeck, J., Augustin, T. And Singer, J.M., "Smoothing, sampling, and Basu's elephants", **Proceedings of the 22nd International Workshop on Statistical Modelling**, Barcelona, 245-248 (2007).

Esin, A., Aydın, C., Bakır, M.A. Ve Gürbüzsel, E., "Temel örnekleme yöntemleri", Taro Yamane Çevirisi, **Literatür Yayıncılık**, Ankara, 45-50, 229-250 (2001).

Estevao, V.M. And Särndal, C.E., "A functional form approach to calibration", **Journal Of Official Statistics**, 16 (4): 379-399 (2000).

Fuller, W.A., "Regression estimation for survey samples", **Survey Methodology**, 28 (1): 5-23 (2002).

İnternet : Gelman, A. And Carlin, B.C., "Poststratification and weighting adjustments", The Pennsylvania State University, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.38.5290> (2000).

Gelman, A., "Struggles with survey weighting and regression modeling", **Statistical Science**, 22 (2) :153-164 (2007).

Guy, P.W., "Small sample theory for poststratification", Ph.D. Thesis, **University of Texas ASM**, Texas, 1-110 (1979).

Hájek, J., "Comment on a paper of D. Basu, in Godambe, V.P., Sprott, D.A. Édés.", **Foundations of Statistical Inference**, Toronto, 236-237 (1971).

Hansen, M.H. And Hurwitz, W.N., "On the theory of sampling from a finite population", **Ann. Math. Stat.**, 14: 333-362 (1943).

Hansen, M.H., Hurwitz, W.N. And Madow, W.G., "Sample survey methods and theory I", **John Wiley and Sons**, New York, 15-90 (1953).

Internet: Hardin, D. And Gedif, K., "'Variance estimation in complex surveys", Sampling: Design and Analysis, Southern Methodist University, <http://faculty.smu.edu/slstokes/stat6380/Variance%20Estimation%20in%20Complex%20Surveys.ppt> (2006).

Holt, D. And Elliot, D., "Methods of weighting for unit non-response", **The Statistician**, 40: 333-342 (1991).

Holt, D. And Smith, T.M.F., "Post stratification", **Journal of The Royal Statistical Society**, Series A (General), 142 (1): 33-46 (1979).

Horvitz, D.G. And Thompson, D.J., "A generalization of sampling without replacement from a finite universe", **J. of The American Stat. Association**, 47: 663-685 (1952).

Hulliger, B., "Simple and Robust Estimators for Sampling", **American Stat. Ass. Section on Survey Research Methods**, 54-63 (1999).

Jayasuriya, B. And Valliant, R., "An application of regression and calibration estimation to post-stratification in a household survey", **American Stat. Ass. Section on Survey Research Methods**, 902-907 (1995).

Kalton, G., "Introduction to survey sampling", **SAGE Publications**, 35, Beverly Hills, 1-22 (1983).

Kalton, G., "Discussion: Little, R.J.A., Survey nonresponse adjustments", **American Stat. Ass. Section on Survey Research Methods**, 1-10 (1984).

Kalton, G. And Flores-Cervantes, I., "Weighting methods", **Journal of Official Statistics**, 19 (2): 81-97 (2003).

Kalton, G. And Maligalig, D.S., "A comparison of methods of weighting adjustment for nonresponse", Proceedings of The 1991 Annual Research Conference, **U.S. Bureau of the Census**, 409-428 (1991).

Kish, L., "Survey sampling", **John Wiley & Sons Inc.**, New York, 1-35 (1965).

Kish, L., "Weighting: why, when, and how". **American Stat. Ass. Section on Survey Research Methods**, 121-130 (1990).

Kish, L., "Weighting for unequal  $\pi_i$ ", **J. of Official Statistics**, 8 (2): 183-200 (1992).

Kott, P.S., "On calibration weighting", Research and Science, Reports, **National Agricultural Statistics Service**, NASS, 1-40 (2003).

Little, R.J.A., "Robust model based inferences about a finite population mean from weighted samples", **American Stat. Ass. Section on Survey Research Methods**, 613-617 (1981).

Little, R.J.A., "Models for nonresponse in sample surveys, **Journal of The American Statistical Association**, 77: 237-250 (1982).

Little, R.J.A., "Survey nonresponse adjustments", **American Stat. Ass. Section on Survey Research Methods**, 1-10 (1984).

Little, R.J.A., "Survey nonresponse adjustments for estimates of means", **International Statistical Review**, 54: 139-157 (1986).

Little, R.J.A., "Post-stratification: a modeler's perspective", **Journal of The American Statistical Association**, 88: 1001-1012 (1993).

Little, R.J.A. And Vartivarian, S., "Does weighting for nonresponse increase the variance of survey means?", **American Stat. Ass. Section on Survey Research Methods**, Minneapolis, 3897-3904 (2005).

Internet: Lu, H. And Gelman, A., "Sampling variances for surveys with weighting, poststratification, and raking", The Pennsylvania State University, <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.27.3738> (2000).

Lu, H. And Gelman, A., "A method for estimating design-based sampling variances for surveys with weighting, poststratification, and raking", **Journal of Official Statistics**, 19 (2): 133-151 (2003).

Internet: Norman, P., "Putting iterative proportional fitting on the researcher's desk", University of Leeds, [www.geog.leeds.ac.uk/wpapers/99-3.pdf](http://www.geog.leeds.ac.uk/wpapers/99-3.pdf) (1999).

Oh, H.L. And Scheuren, F.J., "Multivariate raking ratio estimation in the 1973 exact match study **American Stat. Ass. Section on Survey Research Methods**, Washington, 716-722 (1978).

Oh, H.L. And Scheuren, F.J., "Weighting adjustment for unit nonresponse", *Incomplete Data in Sample Surveys 2*, Madow, W., Olkin, I., Rubin, D., **Academic Press**, New York, 143-184 (1983).

Oh, H.L. And Scheuren, F.J., "Modified Raking Ratio Estimation", **Survey Methodology**, 13 (2): 209-219 (1987).

İnternet: Opsomer, J.D., “Alternative Approaches to Inference from Survey Data” Colorado State University Department of Statistics, [http://www.stat.colostate.edu/~jopsomer/papers/Opsomer\\_Inference\\_Intro2.pdf](http://www.stat.colostate.edu/~jopsomer/papers/Opsomer_Inference_Intro2.pdf) (2008).

Oral, E., “Recent advances in survey sampling”, Recent Advances in Statistics, Ayhan H.Ö., Batmaz İ., **Turkish Statistical Institute Press.**, Ankara, 151-165 (2007).

ORC Macro, “Sampling manual”, DHS-III Basic documentation Calverton, Maryland: **Macro International Inc.**, Maryland, 1-55 (1996).

Osier, G. And Museux J.M., “Variance Estimation for EU-SILC Complex Poverty Indicators using Linearization Techniques”, **European Conference on Quality in Survey Statistics**, Luxembourg, 1-11 (2006).

Özdamar, K., “Paket programlar ile istatistiksel veri analizi 1”, **Kaan Kitabevi**, Eskişehir, 80-120 (1999).

Rao, J.N.K., “Conditional inference in survey sampling”, **Survey Methodology**, 11: 15-31 (1985).

Rao, J.N.K., “Estimating totals and distribution functions using auxiliary information at the estimation stage”, **Journal of Official Statistics**, 10 (2):153-165 (1994).

Rao, R.S., “Inference for multi-wave surveys with nonresponse”, Ph.D. Thesis, **University of Boston Graduate School**, Boston, 1-50 (2001).

Rizzo, L., Kalton, G. And Brick, J.M., “Weighting adjustments for panel nonresponse in the SIPP”, **U.S. Bureau of the Census**, Westat, 2-3 (1994).

Rubin, D.B., “Multiple imputation for nonresponse in surveys”, **John Wiley & Sons Inc.**, New York, 1-42 (1987).

Särndal, C.E., “On  $\pi$ -inverse weighting versus best linear unbiased weighting in probability sampling”, **Biometrika**, 67: 639-650 (1980).

Särndal, C.E., “The calibration approach in survey theory and practice”, **Survey Methodology**, 33 (2): 99-119 (2007).

Särndal, C.E., Swenson, B. And Wretman, J., “Model assisted survey sampling”, **Springer-Verlag Inc.**, New York, 24-58, 70-114, 162-207 (1992).

Shahbaz, M.Q., “Sampling with unequal probabilities and without replacement”, Ph.D. Thesis, **University of National College of Business Adm. & Econ**, Pakistan, 1-55 (2003).

Shahbaz, M.Q. And Hanif, M., “A simple procedure for unequal probability sampling without replacement sample size 2”, **Pak J. Statist.**, 19 (1): 151-160 (2003).

Siegel, P., Chromy, J.R. And Copello, E., “Propensity models versus weighting class approaches to nonresponse adjustment: a methodological comparison”, **American Stat. Ass. Section on Survey Research Methods**, Minneapolis, 3560-3565 (2005).

Singh, S., “Advanced sampling theory with applications 1”, **Kluwer Academic Publishers**, London, 729-733 (2003).

Smith, T.M.F., “Poststratification”, **The Statistician**, 40: 315-323 (1991).

Internet: Tillé, Y., “Teoría de Muestreo”, Université de Neuchâtel, [http://www2.unine.ch/webdav/site/statistics/shared/documents/corso\\_teoria\\_de\\_muestreo.pdf](http://www2.unine.ch/webdav/site/statistics/shared/documents/corso_teoria_de_muestreo.pdf) (2005).

Internet: TMSD, “Training Manual on Sample Design for Surveys” Survey Research Training Materials, International Programs Center, <http://www.aec.msu.edu/fs2/survey/> (2006).

Ünüvar, T.G., “Data weighting procedures in sample surveys”, Msc. Thesis, **The Middle East Technical University**, Institute of Science & Technology, Ankara, 40-100 (1997).

Internet: Vanderhoeft, C., “Generalized calibration at statistics Belgium, SPSS® module g-calib-s and current practices”, Official Stat. Institution In Belgium, <http://statbel.fgov.be/studies/paper03.pdf> (2001).

Vartivarian, S.L. And Little, R.J.A., “Weighting adjustments for unit nonresponse with multiple outcome variables”, **American Stat. Ass. Section on Survey Research Methods**, San Francisco, 4358-4363 (2003).

Vartivarian, S., “On the formation of weighting class adjustments for unit nonresponse in sample surveys”, Ph.D. Thesis, **University of Michigan**, Michigan, 1-100 (2004).

Vaughn, B.J., “An empirical comparison of weighting class adjustment and propensity modeling adjustment for nonresponse”, Msc. Thesis, **University of North Carolina**, Chapel Hill, 1-50 (2001).

Verma, V., “Sampling methods”, **Manual for Statistical Trainers Statistical Institute for Asia and The Pacific**, Tokyo, 2-7 (2002).

Verma, V., “Recent advances in survey sampling”, Recent Advances in Statistics, Ayhan H.Ö., Batmaz İ., **Turkish Stat. Inst. Press.**, Ankara, 77-101 (2007a).

Verma, V., “Cross-sectional and longitudinal weighting in a rotational household panel: applications to EU-SILC”, *Statistics in Transition*, 8 (1): 5-50 (2007b).

Verma, V., Betti, G. And Ghellini, G., “Cross-sectional and longitudinal weighting in a rational household panel: applications to EU-SILC”, *Statistics in Transition*, New Series, 8 (1): 5-50 (2007).

Verma, V., “Sampling for household-based surveys of child labour”, International Labour Organization, *International Training Centre of the ILO*, Italy, 50-110 (2008).

Wu, C. And Sitter, R.R., “A model-calibration approach to using complete auxiliary information from survey data”, *J. of The American Stat. Ass.*, 96: 185-193 (2001).

Wu, C., “Optimal calibration estimators in survey sampling”, *Biometrika*, 90(4): 937-951 (2003).

Internet: Wun, L.M., Ezzati-Rice, T.M., Baskin, R., Greenblatt, J., Zodet, M., Potter, F., Diaz-Tena, N. And Touzani, M., “Using propensity scores to adjust weights to compensate for dwelling unit level nonresponse in the medical expenditure panel survey”, Agency for Healthcare Research and Quality, <http://www.ahrq.gov> (2004).

Wywiał, J., “On conditional sampling strategies”, *SpringerLink*, 49: 397-419 (2003).

Zhang, L.C., “Post-stratification and calibration-a synthesis”, *The American Statistician*, 54 (3): 178-184 (2000).

Zieschang, K.D., “Sample weighting methods and estimation af totals in the consumer expenditure survey”, *J. The American Stat. Ass.*, 85: 986-1001 (1990).

## ÖZGEÇMİŞ

### Kişisel Bilgiler

Soyadı, adı : ALKAYA, Aylin  
 Uyruğu : T.C.  
 Doğum tarihi ve yeri : 14.09.1974 Zonguldak  
 Medeni hali : Bekar  
 Telefon : 0 (384) 215 20 07  
 e-mail : [aylin@nevsehir.edu.tr](mailto:aylin@nevsehir.edu.tr).

### Eğitim

| Derece        | Eğitim Birimi                           | Mezuniyet tarihi |
|---------------|-----------------------------------------|------------------|
| Yüksek lisans | Gazi Üniversitesi /İstatistik Böl.      | 2003             |
| Lisans        | Hacettepe Üniversitesi/ İstatistik Böl. | 1998             |
| Lise          | T.E.D. Zonguldak Koleji                 | 1992             |

### İş Deneyimi

| Yıl       | Yer                   | Görev               |
|-----------|-----------------------|---------------------|
| 2007-     | Nevşehir Üniversitesi | Araştırma Görevlisi |
| 2000-2007 | Erciyes Üniversitesi  | Araştırma Görevlisi |

### Yabancı Dil

İngilizce